
S

Safety

► [Officer Safety, Health, and Wellness](#)

Sample Selection Models

Brian Johnson
Department of Criminology and Criminal
Justice, University of Maryland, College Park,
MD, USA

Overview

Analytical issues involving sample selection are pervasive in criminological research. This entry provides an introductory overview of the sample selection problem followed by discussion of several common statistical models designed to address sample selection bias in criminological research. Examples of research involving sample selection models are discussed along with the different types of sample selection that can occur and their appropriate modeling strategies. The entry concludes with a brief overview of common criticisms and concerns surrounding sample selection models in the field and discusses future directions for development in this area.

Introduction

Social scientists have long recognized the ubiquitous threat of sample selection. Whether

relying on surveys, interviews, official records, or even experiments, sample selection issues are pervasive in criminological research. Whenever information is available on a nonrandom subsample of the population of interest, or when observations are selected through a process that is not independent of the outcome of interest, sample selection bias represents a key threat to empirical validity (Bushway et al. 2007). Examples of selection bias in criminological work are wide ranging. From examinations of police discretion (Worrall 2002; Lundman and Kaufman 2003; D'Alessio and Stolzenberg 2003; Kingsnorth et al. 1999) to sentencing disparity (Berk 1983; Hagan and Palloni 1986; Klepper et al. 1983), from publication bias in meta-analysis (Wilson 2010) to the effects of adolescent work on crime (Apel et al. 2008), and from tests of self-control theory (Paternoster and Brame 1998) to examinations of life-course criminology (Laub and Sampson 2003), selection bias often rears its ugly head in criminological research. Indeed, it can be difficult to identify any criminological research question that is not threatened to some degree by sample selection. This entry briefly reviews analytical issues surrounding sample selection bias, including why sample selection leads to biased and inconsistent estimates, how sample selection has been dealt with in criminological research, and how important conceptual distinctions in different types of selection can be classified and understood. The focus is on statistical models that are designed specifically for situations involving censoring or truncation of

the dependent variable, which are collectively referred to here as “sample selection models.”

Sample Selection in Criminological Research

The threat of selection bias is omnipresent in social research, and it is a fundamental aspect of many criminological inquiries – it can result from sundry causes, including natural human behaviors, specific research designs or implementations, or a host of other unforeseen individual, organizational, and environmental causes. Sample selection can result from instrument bias, such as when a survey instrument is designed to collect certain information from only a subset of respondents. It can be caused by sampling strategies that systematically omit certain populations, such as the homeless or incarcerated populations (Mosher et al. 2002). It can also arise through sample attrition, such as when different types of subjects are more likely to drop out of a study than others. Not only is sample selection an important source of bias, but sample selectivity itself may at times be of substantive interest to the researcher as well. If the selection mechanisms affecting the sample are fully known, they can be accounted for through sampling weights or related mechanisms, but rarely are all selection processes easily identifiable or explicitly defined in criminological research.

Various different treatments of sample selection issues have conceptualized the problem in terms of nonresponse, sample attribution, model misspecification, and/or omitted variable bias (Gondolf 2000; Maxwell et al. 2002; Berk 1983; Heckman 1979; Little and Rubin 1987). Selection bias can be understood as a sampling issue (the researcher is limited to information from a nonrandom subsample of the population of interest), a selection problem (observations are selected in such a way that they are not independent of the outcome of interest), or as a statistical problem (a correlation exists between the error term and the independent variable in the model). However it is conceived, though, the basic problem is the same – a researcher is limited to

information on the dependent variable that comes from some subsample of observations that is not randomly drawn from the population of interest and therefore is not representative of the larger population to which inferences are to be made. Because statistical inferences are grounded in the principles of random sampling, generalizations based on nonrandom subsamples are likely to lead to faulty causal inferences.

The prototypical example comes from economics, where researchers are often interested in estimating wages, but wages are only observed for employed individuals (e.g., Heckman 1974; Moffitt 1999; Hagan 1990). Hagan (1990), for instance, examined this issue in the context of gender stratification in the justice system, examining income inequality among male and female lawyers. He found that both men and women who were more committed to work roles were more likely to be employed full time, and because commitment to work affected earnings, comparisons based only on full-time employees led to biased estimates of income differentials in the legal profession.

Similar examples are common in other areas of criminological research. Scholars interested in criminal punishment, for instance, only observe severity of charges for arrested offenders, sentence lengths for incarcerated offenders, and recidivism outcomes for offenders who have been released from jail or prison. Of course the same factors associated with being arrested, incarcerated or reoffending are also likely to affect initial charge severity, sentence lengths, and the likelihood of release. The problem, then, is that the process that determines which cases will be observed is fundamentally related to the outcome of interest to the researcher. This means that at any given stage of criminal case processing, the observed sample is unlikely to be representative of the larger population of all offenders from which it originated. Selective attrition through different stages of the justice system produces nonrandom samples of offenders, so estimates based on these subsamples provide biased inferences about relationships in the larger population of interest. Kingsnorth et al. (1999), for instance,

investigated the role that victim characteristics and offender-victim relationships played in the severity of case processing outcomes from initial intake to final disposition in a sample of sexual assaults. As they recognized, “Unobserved factors may influence whether a case is continued to the next step,” therefore “Failure to take these unobserved influences into account. . . leads to the problem of sample selection bias” (Kingsnorth et al. 1999, p. 284). Out of the 467 cases in their sample, only 41 % were fully prosecuted and even fewer received terms of incarceration. To address this, these authors implemented sample selection models in an attempt to control for selection bias in their analysis of sentence lengths.

Another source of potential selection bias can arise in offending or victimization surveys that have skip patterns built into them (Winship and Mare 1992). For instance, a researcher might ask survey respondents to report whether or not they have ever been arrested, and then ask additional questions of arrested respondents, such as how the arrest affected their level of satisfaction with police. Analyses based on these questions, though, would only be observed for arrested respondents, so inferences drawn from them would be subject to important selection biases. Negative attitudes towards police, for example, may affect both the probability of arrest and postarrest satisfaction levels, leading to biased inferences for the population. Early research on police satisfaction that relied on police-citizen encounters was limited in this way because it relied on data from only citizens who came in contact with the police (Reisig and Parks 2000). Estimates of satisfaction with the police from this type of selected sample will be biased if there are any differences in satisfaction between individuals who have had contact with the police and those that have not.

Even research designs that incorporate random experiments are often threatened by selection bias. Although experiments represent the quintessential tool for randomizing observations, controlling for unmeasured variables, and drawing improved causal inferences (Weisburd 2010), they are far from catholicons for sample selection. As Sampson (2010, pp. 490–491) recently

argued, randomization is “the trump card that the experimentalist plays in the causal inference game” though too often scholars “focus on the benefits of randomization rather than the selection mechanisms that produce the samples in the first place.” His point is that randomization can only address selection bias if the researcher begins with a random sample from the population of interest, and often this is not the case. Even when a random sample is available, though, noncompliance, information diffusion or contagion, or sample attrition, interference, and missing data problems may jeopardize causal inferences in experimental research (Berk 2005; Sampson 2010). For instance, differential nonparticipation or implementation across experimental sites can often reintroduce the types of selection bias that experimentation is designed to overcome.

In the majority of research, though, scholars are limited to nonexperimental or observational data. Spohn and Holleran (2002), for example, compared recidivism outcomes in a sample of drug-involved offenders, some of whom were sentenced to probation and others who received prison. They rightly recognized that “offenders sentenced to prison” are “qualitatively different from offenders placed on probation” and that “offenders deemed more threatening and more likely to return to a life of crime will, all else being equal, face higher odds of incarceration” (Spohn and Holleran 2002, p. 340). The problem is that offenders are not randomly assigned to prison and probation so such analyses risk comparing apples to oranges. The authors attempt to control for these differences by including “each offender’s predicted probability of incarceration” in their models of recidivism (341). Although this is not explicitly a sample selection model, it follows the same general logic. Similarly, D’Alessio and Stolzenberg (2003, p. 1390) investigated the impact of race on arrest using NIBRS data for four violent crimes. To simplify the analysis, they excluded cases involving multiple offenders or multiple victims. The problem is that race may not exert the same influence in these cases as in cases with a single offender and single victim. They therefore include a “logit-based ‘hazard

rate' variable to account for the exclusion of crime incidents with multiple offenders and/or victims" – this is essentially a sample selection model to correct for selection bias.

Smith and Paternoster (1990) explore the role of selection bias more explicitly in their examination of the effect of juvenile placement on future offending. From the perspective of labeling theorists, formal sanctions can trigger secondary deviance, in which an individual's self-identity is transformed to accept the deviant label, leading to "deviance amplification" or increased involvement in future deviant behaviors (Lemert 1951). As these authors point out, though, if youth who are more likely to recidivate are also more likely to receive a sentence to juvenile court, then the relationship between the two could be caused by selection bias. In this case the measure of court referral would be capturing unobserved correlates of recidivism that are not included in the statistical model (e.g., predisposition to delinquency), and the result is biased and inconsistent estimates of the effect of placement on recidivism. If there are any omitted variables that affect both placement and future offending, then it may be that youth who are assigned more severe court dispositions would have been more delinquent regardless of their placement in juvenile court. Using data on juvenile offenders in Florida, Smith and Paternoster initially found a positive effect of juvenile court placement on subsequent referral; however, when they employed sample selection models to control for selection bias, this effect disappeared entirely. They interpret this as evidence that the deviance amplification thesis is largely the result of unaccounted-for selection effects in juvenile case processing.

There are numerous other examples of potential selection effects in criminological research. For instance, Brame (2000) discusses the bias that can characterize missing outcome data in randomized experiments and offers an example of this in the context of police response to domestic disputes. Western and Pettit (2005) use data from inmate surveys to impute missing wages for incarcerated offenders who would otherwise be excluded from earnings estimates. They find that

the apparent decrease in the black-white wage gap in recent years is largely the product of increasing selection bias due to greater numbers of young African American males being incarcerated. Rhodes et al. (2001) provides a useful discussion of the use of instrumental variables to address selection bias in the context of nonlinear survival models, which offer an alternative approach for examining such things as treatment effects in offender recidivism studies. Paternoster and Brame (1998) examine select propositions rooted in self-control theory by investigating the association between criminal and analogous behaviors using data from the Cambridge Study. They estimate a bivariate probit model that quantifies the magnitude of the association between criminal and analogous behaviors. Similarly, Sampson et al. (2002, p. 466) in their review of neighborhood-level research suggest that "the issue of selection bias is probably the biggest challenge facing neighborhood-level research." They point out that selection bias can arise from residential mobility decisions, when individuals self-select into neighborhoods. In short, these brief examples serve to demonstrate that sample selection issues characterize a dynamic range of criminological research questions. Sample selection models offer one useful analytical approach for addressing selection bias, but before discussing these models, it is necessary to first understand the different ways that sample selection can occur.

Varieties of Sample Selection

In order to know whether or not sample selection models are appropriate, one must first know what type of selection is present. Sample selection on the dependent variable can occur in two primary ways, which directly affects the type of selection model that should be estimated. Some types of sample selection involve what is referred to as *explicit selection*, which occurs when the selection process is "explicitly" tied to the dependent variable; that is, whether or not a case is observed is a direct function of the score on the dependent

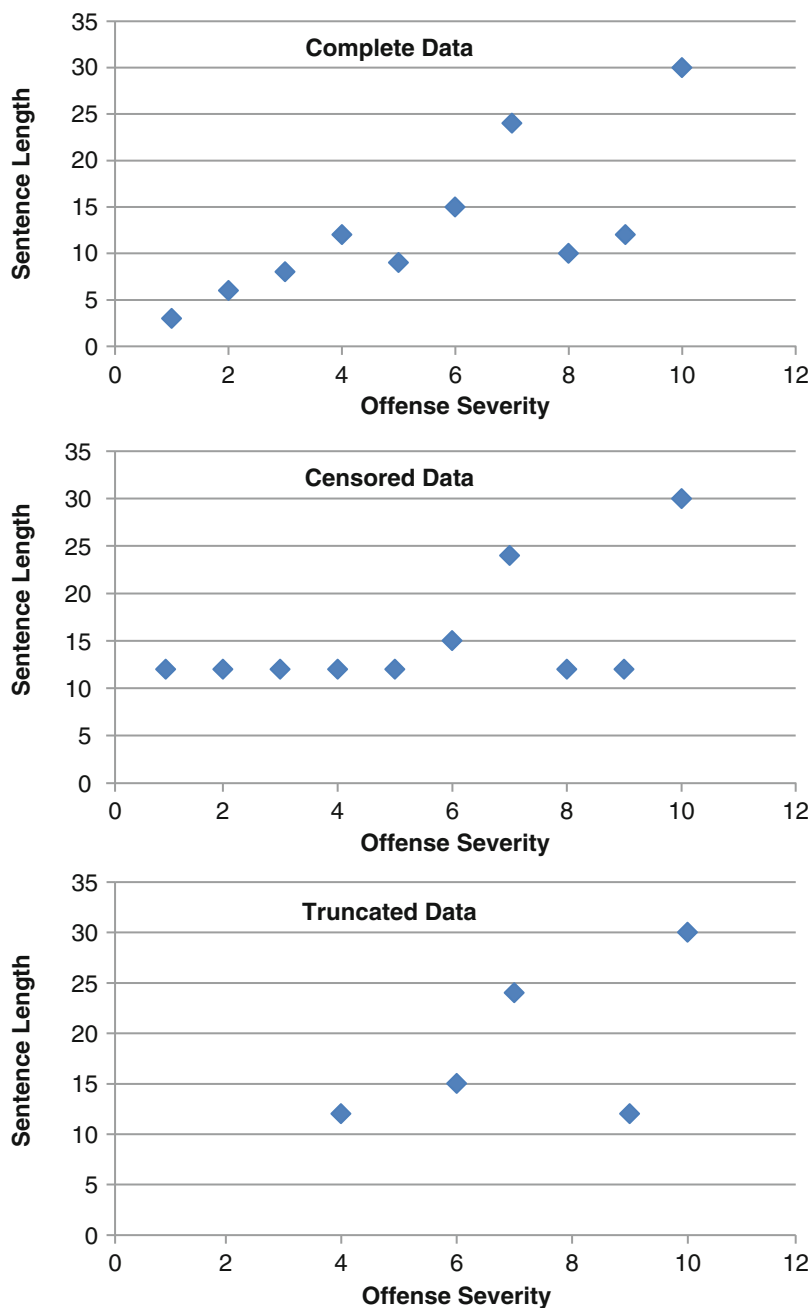
variable. Put another way, Y is known only if some criterion defined in terms of Y is met. The researcher does not observe the value of Y for specific values of the dependent variable. Other types of sample selection are characterized by what is referred to as *implicit* or *incidental selection*, which occurs when the selection process results from some external, stochastic function related to the dependent variable. With incidental selection, the likelihood that an observation is excluded from the sample is determined by a separate probabilistic function, and each observation on the dependent variable has some non-zero chance of inclusion. Put another way, Y is observed only if criteria defined in terms of some other random variable is met. Implicit selection processes necessitate different types of selection models than explicit selection processes, so before the researcher can make a decision about how to best account for selection bias, this basic distinction must be understood.

Regardless of whether one is faced with explicit or implicit selection, though, the sample itself can also be characterized by an important distinction between *truncated data* and *censored data*. Truncation occurs when excluded observations provide no data on either the dependent or independent variable. Unobserved cases are unknown for all variables in the data. That is, X is observed only if Y is observed. Censoring, on the other hand, occurs when information is only missing or incomplete on the dependent variable. Information for the other variables in the data is recorded for all cases. That is, X is observed for the entire sample, regardless of whether Y is observed or not. Typically, censored data involves the clustering of observations at some threshold in the data, such that exact values above or below some cutoff are unknown and therefore set to the threshold value. With censored data we have all of the observations on the independent variables, but we don't know the "true" values on the dependent variable if they fall below or above the threshold. These basic distinctions are visually represented in Fig. 1 using the example of sentence lengths that are censored or truncated at 1 year, where each data point represents the sentence of one offender in the data.

To make the discussion more concrete, imagine that you want to estimate the effect of a criminal record on one's future earnings (e.g., Nagin and Waldfogel 1998). In scenario 1, you use official data that are only collected for a sample of individuals with earnings above the poverty line. This results in explicit selection because the dependent variable, income, determines which cases are observed, and it represents a truncated sample, because no information is known about either prior record or income for omitted cases. In scenario 2, income and criminal history data are collected for all offenders, but incomes below the poverty line are only recorded as "below the poverty line." Exact income values are therefore unknown for offenders who fall below this threshold. In this case, you still have explicit selection based on the dependent variable, but the sample is now censored. Data on criminal history are available for all offenders, but exact information on income is unknown for subjects below the poverty line. In scenario 3, you collect survey data in which income values are missing for a nonrandom subsample of offenders. Sample selection in this case is not an explicit function of the dependent variable in which every value above or below some cutoff value is excluded. Instead, it results from a stochastic process in which the *probability* of being unobserved is nonrandom. For instance, offenders with low incomes may be less likely to answer the survey at all, in which case you have incidental selection with truncation, or perhaps they respond to the survey but refuse to answer the income question, in which case you have incidental selection with censoring. Although different selection processes necessitate different modeling solutions, in general, similar biases result from cases involving both explicit and incidental selection processes.

These basic distinctions are further illustrated in Table 1, which uses the example of criminal sentencing data. Imagine that a researcher is interested in estimating the effect of crime severity on length of imprisonment. The first question is whether there is an explicit or implicit selection process at work. If data on the dependent variable are only collected for prisoners (i.e., offenders sentenced to at least 1 year of incarceration),

Sample Selection Models, Fig. 1 Censored and truncated data distributions



then the researcher is faced with explicit selection. Whether or not the dependent variable is observed is determined fully by the length of imprisonment. The second question then is whether the data are truncated or censored. If data on the independent variable, crime severity, is only collected for prisoners, then the data are

truncated – they are missing entirely for offenders who receive less than 12 months of incarceration, as in the upper left-hand box of [Table 1](#). If, however, crime severity data are collected on all convicted offenders, but the dependent variable is only recorded as “less than 1 year” for non-prisoners in the data, then the data are censored,

Sample Selection Models, Table 1 Typology of types of sample selection

		Type of selection	
		Explicit selection	Incidental selection
Type of sample	Truncated	Exact prison lengths are only observed for sentences of 1 year or more. No data are collected from offenders sentenced to less than 1 year of imprisonment	The probability that prison length is observed is a function of other factors related to the likelihood of imprisonment and no data are collected from non-prisoners
	Censored	Exact prison lengths are only observed for sentences of 1 year or more. Data on offense severity are collected from all offenders but sentences are only recorded as “less than 1 year” for non-prisoners	The probability that prison length is observed is a function of other factors related to the likelihood of imprisonment. Data on independent variables are collected from all offenders but prison lengths are only known for prisoners

as in the lower left-hand box in [Table 2](#). Information on the independent variable would be available for all offenders, but exact sentence lengths would be unobserved for offenders sentenced to less than 1 year of incarceration. The alternative to explicit selection is implicit selection. If the likelihood that any given prison length is observed is a function of other factors related to sentencing, then implicit selection is at work. If no data are collected from certain offenders, then the sample is once again truncated as in the upper right-hand box of [Table 1](#). If, however, prison lengths are missing as a function of other sentencing factors related to imprisonment, but offense severity is known for all offenders, then the sample is characterized by incidental selection with censoring as in the lower right-hand box of [Table 1](#).

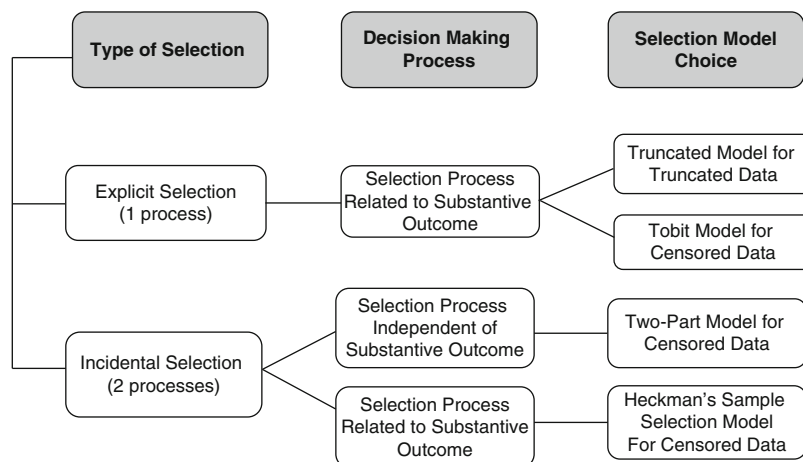
Sample Selection Models, Table 2 Censored, truncated and sample selected models

	Y	X	Example
Censored (explicit, censored)	Y is known only if some criterion defined in terms of y is met	X is observed for the entire sample, regardless of whether Y is observed or not	Exact prison lengths are recorded only for terms of 1 year or greater. Shorter sentences are recorded as “less than 1 year”
Truncated (explicit, truncated)	Y is known only if some criterion defined in terms of y is met	X is observed only if Y is observed	Prison lengths are only recorded for terms of 1 year or greater. No data are collected for offenders sentenced to less than 1 year of incarceration
Sample selected (incidental, censored)	Y is observed only if criteria defined in terms of some other random variable (Z) are met	X and W (the determinants of whether $Z = 1$) are observed for the entire sample, regardless of whether Y is observed or not	Prison lengths are recorded only for offenders who are imprisoned and the probability of imprisonment is a function of other factors, such as offense severity

Adapted from Breen (1996)

[Figure 2](#), adapted from Bushway et al. (2007), provides a schematic heuristic outlining the decision-making process involved in selecting an appropriate selection model depending on the type of sample selection that is present. Detailed discussion of these models follows below, but the point here is that it is essential to understand the type of selection one is faced with before deciding how to attempt to correct for selection bias. In the face of explicit selection, the sample selection process is inherently tied to the dependent variable. When data are truncated, truncated

Sample Selection Models, Fig. 2 The process of selecting a sample selection model (Note: Heuristic adapted from Bushway et al. 2007)



regression models can be estimated. When data are censored, Tobit regression models may be appropriate (Tobin 1958). These two approaches are similar except that the Tobit model is able to incorporate information on the independent variables of censored cases in the parameter estimates of the substantive equation, whereas this information is unavailable in truncated samples. When incidental selection is present, one must make a determination about whether or not the selection process is independent of the outcome of interest. If the selection process is unrelated to the outcome, then by definition selection bias is not a problem and two separate, independent models can be estimated, known as the two-part model (discussed below). If, however, the selection process is related to the outcome, then selection bias becomes a concern, and Heckman's two-part sample selection model may provide improved model estimates.

The astute reader will notice that regardless of which type of selection process is operative, under incidental selection, both model choices apply only to censored data. The reason for this is because with truncated data, no information is available on the characteristics of the excluded cases, so it becomes impossible to model the selection process in the presence of incidental selection involving truncated data. Theoretically, it is possible that an exogenous selection process determines which cases are omitted, but if no data is collected from the omitted cases, this process

cannot be formally modeled with traditional selection models. For this reason, some treatments (e.g., Breen 1996) of selection bias collapse the four boxes in Table 1 into the three categories summarized in Table 2. Once a researcher knows what type of sample selection is occurring, then it becomes possible to choose an appropriate statistical model. For censored samples, the fundamental logic involves a two-stage modeling procedure that attempts to first capture the selection process and then correct for it in the substantive model of interest.

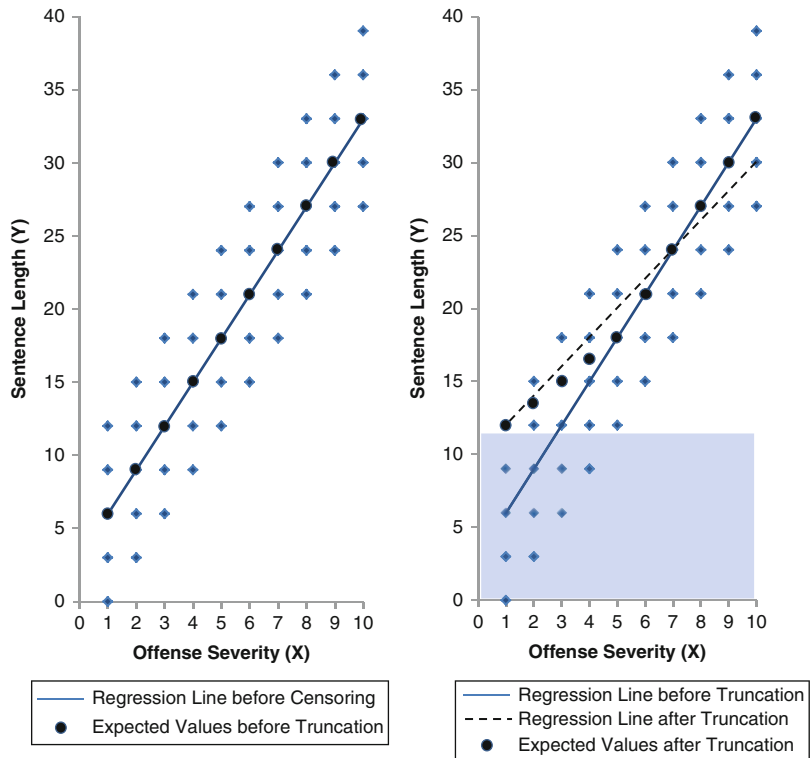
The Sample Selection Problem

Before turning to an elaboration of common types of statistical models that can be employed when faced with sample selection, it is instructive to first consider the limitations of the ordinary least squares regression model and how analyses conducted on selected subsamples can lead to biased statistical inferences. For illustrative purposes, the ordinary regression model is presented in Eq. 1:

$$Y = \beta X + \varepsilon \quad (1)$$

In this model, Y is a continuous outcome, such as sentence length; X is an exogenous independent variable, such as crime severity; and ε is a normally distributed error term with an

Sample Selection Models, Fig. 3 The effects of truncation on the ordinary regression line and expected values



expected value of zero for each value of X . We assume that the functional form between X and Y is linear and additive, that the conditional distribution of Y is normal with uniform standard deviations, that observations are independent, and that X is uncorrelated with the error term ϵ . When these assumptions are met, the OLS estimator will provide unbiased and consistent estimates of the model parameters. Selection processes typically introduce an association between the error term and the exogenous variable in the model; in this context, then, selection bias can be understood as a violation of traditional OLS model assumptions.

Imagine that you want to examine the relationship between crime severity and sentence length in a random sample of 50 convicted offenders. The sample selection problem is that not all sentence lengths are observed. If the unobserved cases are comprised of a random subset of cases, then selection bias is not a problem. Typically, though, prison sentences are reserved for confinement sentences of 1 year or more, with shorter

sentences being served in county jail. In an analysis of sentence lengths that relies on prison data, then, shorter terms of confinement (and concomitantly less serious crimes) will be systematically excluded from observation. In this example, the researcher would be faced with explicit selection in a truncated sample, though it is just as easy to conceive of related examples that involve implicit selection and/or censored samples and the biases introduced are comparable. This scenario is demonstrated in Fig. 3, which reproduces a similar example discussed in Berk (1983).

The shaded area represents sentence lengths of less than 12 months, which are omitted from the prison data. The solid line is the regression line based on all of the values if sentences of less than 12 months are included, and the dashed line is an approximation of the regression line you might expect based on only the non-truncated observations. Recall that one of the key assumptions of OLS is that the expected value for the disturbance term is zero for all values of X . That is, the fitted

values from the model should match the means of the observed data. In the complete data shown in the left-hand panel above, the expected values fall directly on the solid regression line. For instance, when offense severity is one, the expected sentence based on the five data points would be 6 months. In the truncated sample, however, the expected values no longer conform to the original regression line, nor do they align with the adjusted line after truncation. For offense severity scores below five (where the truncation occurs), the expected values fall systematically above the original regression line. Sample selection in this case results in expected values that are shifted upward. For instance, when offense severity is one in the truncated sample, the expected sentence is shifted from 6 to 12 months. When offense severity is equal to two, the expected sentence shifts from 9 to 13.5 months and so on. The regression line no longer coincides with the expected values at each value of X , and the errors associated with the regression line will no longer have expected values of zero in this range. The regression line estimated on the truncated sample systematically falls on or above the expected values when offense severity is low, and it falls below the expected values when offense severity is high. The truncation of low values on Y therefore introduces a correlation between the exogenous variable, offense severity, and the error term in the regression model, which violates OLS assumptions and leads to bias in the estimated parameters.

Several observations are readily apparent from Fig. 3. First, the regression line for the truncated observations is different from the regression line for the complete data. Compared to the true relationship for all offenders, the truncated relationship is attenuated. The effect of selection bias in this instance is to underestimate the relationship between offense severity and sentence length. Of course other types of censoring in other scenarios may just as easily lead to overestimates of relationships in the population. Second, the expected values (the black circles) after censoring no longer conform to the expected linear relationship between X and Y .

This reflects the fact that the degree of censoring is increased at lower values of the exogenous variable. The regression line, in a sense, is flattened as it nears the censoring floor. This violates another assumption of ordinary least squares, namely, that the relationship between the independent and dependent variable is linear. In this context, then, the problem of selection bias can be understood as a specification error regarding the wrong functional form of the relationship (Berk 1983).

There are several implications to the violations of these model assumptions. First, external validity is compromised. The estimated relationship in the censored sample does not represent the true relationship between offense severity and sentence length for all sentenced offenders. The regression slope is a biased estimate of the true relationship in the population. Second, internal validity is also compromised. This is seldom understood. It is common for researchers to claim an interest in only the selected subsample as a means of deflecting concerns over selection bias, but because the regression line from the truncated sample will systematically overestimate the relationship between X and Y for lower values of offense severity and underestimate it for higher values, the error term will be positively correlated with X , which will produce biased and inconsistent parameter estimates, *even if one claims an interest in only the subsample of observed cases*. In sum, the systematic exclusion of nonrandom observations from a sample is likely to inadvertently induce a relationship between observed variables and the error term, which produces biased and inconsistent parameter estimates. This unintended correlation is at the heart of the selection bias problem in diverse criminological research.

To further demonstrate, imagine an ordinary regression equation where Y is a measure of the number of future arrests, X is whether or not an offender went through some treatment program expected to lower the number of arrests, and β estimates the effect of program participation on future arrests. The least squares solution to β is provided in Eq. 2.1, and the expected value for $\hat{\beta}$ after substitution is given in Eq. 2.2:

$$\hat{\beta} = (\hat{X}X)^{-1}\hat{X}Y \quad (2.1)$$

$$E(\hat{\beta}) = \beta + E[(\hat{X}X)^{-1}\hat{X}\varepsilon] \quad (2.2)$$

If the covariance between X and ε is 0, then the second term on the right-hand side of the equation is zero and the expected value of $\hat{\beta}$ is equal to its true value β . If however X and ε are correlated, then this term does not drop out of the equation and the estimated value for β is no longer equal to its true value. This reflects the influence of selection bias on the program effect in our example. Heckman (1976, 1979) demonstrates that this type of selection problem can be understood in terms of a specification error where the second parameter in Eq. 2.2 represents an omitted variable that is not included in the ordinary regression model. This omitted variable will produce biased estimates of our program effect in the current example. By systematically excluding certain observations, one introduces the need for an additional regressor to capture the selection process. Heckman's solution is the use of a two-step estimator that first models the sample selection process and then introduces an additional regressor to capture the selection process in the outcome equation. As others have shown, the degree of bias introduced by sample selection will depend on a number of factors, including the degree and location of censoring and the strength of the correlation between the exogenous variable and the error terms in the two equations (Stolzenberg and Relles 1990, 1997; Bushway et al. 2007).

State of the Art: Statistical Models for Addressing Sample Selection

The Simple Two-Part Model

To be clear, statistical inferences based on subsets of data are not inherently problematic. If the missing observations comprise a random subset of cases from the population, OLS will continue to provide unbiased and consistent estimates with only a loss of efficiency. In the case of incidental selection with randomly missing data, for example, a simple two-part model (TPM) can be

estimated (Duan et al. 1983, 1984). The researcher can model the selection process using a binary outcome model, such as a probit or logit, followed by a separate OLS model for the continuous outcome of interest, which is estimated on the selected subset of observed cases. This two-part model is presented in Eqs. 3.1 and 3.2:

$$P = \gamma Z + \delta \quad (3.1)$$

$$Y = \beta X + \varepsilon \quad (3.2)$$

In this model, P is a dummy variable equal to 1 if the subject is selected and 0 otherwise, and Y is the substantive outcome of interest. Z represents an exogenous independent variable related to the probability of being selected, and all the parameters in Eq. 3.2 are equivalent to Eq. 1. The assumption of the two-part model is that the error term in Eq. 3.2 is uncorrelated with the error term in Eq. 3.1. That is, we must assume that there are no unobserved factors that are related to both the selection process and the substantive outcome of interest. Statistically speaking, we assume that the conditional expectation of ε is zero for all values of X . Under these conditions, there is no selection bias and the simple two-part model above will provide unbiased and consistent estimates of the coefficients of interest. When these assumptions are violated, however, parameter estimates from this "uncorrected" model will be biased and inconsistent. When that is the case, a variety of statistical models can be estimated in an attempt to address the selection effect.

Tobit and Truncated Regression

As Fig. 3 suggests, the Tobit model is appropriate when a researcher is faced with explicit selection with a continuous outcome that is censored or bounded by a limiting value. In some sense, this is the simplest model for addressing sample selection because it involves a single selection process based on the dependent variable (i.e., explicit selection). This model, first proposed by Tobin (1958), accounts for the fact that all values falling above (or below) a certain threshold on the

dependent value are observed as part of a continuous distribution, whereas values falling below (or above) some threshold are censored. These censored cases remain in the data and provide information on the independent, but information on the dependent variable is treated as unknown. The model assumes a normal error distribution and a linear relationship between the exogenous predictor and a latent continuous variable, which differs from the observed values because it includes the range of values that fall below the censoring threshold. The relationship between the observed and latent variables is summarized in Eq. 4:

$$\begin{aligned} Y &= Y^* \text{ if } Y^* > c \\ Y &= c \text{ if } Y^* \leq c \end{aligned} \quad (4)$$

Y represents the observed dimension of the dependent variable, Y^* the latent dimension, and c the censoring value given censoring occurs from below. In practice, the censoring threshold can take on different values under different theoretical frameworks and may involve multiple censoring values at either or both ends of the distribution. Scores for the observed values of the dependent variable, then, are equivalent to scores on the latent variable when they are above the censoring value, and when they fall at or below the censoring value, they equal the cutoff value. Conceptually, the Tobit model treats the censored values on the outcome as part of an underlying latent continuum, the values which are treated as unknown for observations that fall below the censoring value. For example, Y^* might represent the latent propensity to offend, which is only realized as actual offending, Y , for offenders who have been arrested in the data. Although numerous observations may share a value of 0 in the data, in theory they have different values on the underlying latent variable. The explanatory variables in the model, then, are linearly related to this latent dimension of the dependent variable Y^* as shown in Eq. 5, which can be rewritten in terms of the observed variable, Y , in Eq. 6:

$$Y^* = \beta X + \varepsilon \quad (5)$$

$$\begin{aligned} Y &= \beta X + \varepsilon \text{ if } Y > c \\ Y &= c \text{ if } Y \leq c \end{aligned} \quad (6)$$

The model assumes the errors ε are independent and normally distributed with a constant variance and can be estimated using maximum likelihood procedures that combine a standard probit model for the censoring of values with an ordinary linear model for the observed values (Osgood et al. 2002). The expected value of Y , then, is equal to the conditional probability of censoring times the conditional expectation of Y given it falls above the censoring threshold. The same explanatory variable is used to characterize both the probability of censoring and the linear relationship of X to Y for uncensored observations. The Tobit model assumes that the error terms from the probit and linear models are a singular normal density (i.e., that they are equal) and that the independent variables and their associated coefficients from the two models are equivalent. Heckman (1976) shows that the Tobit model can be understood as a special case of the two-step sample selection model (discussed below) in which the error terms, explanatory variables, and regression coefficients are constrained to be equal in the selection and substantive equations. Equivalency among the explanatory factors is a key assumption of the Tobit model, though alternative formulations have been proposed that relax these assumptions. For instance, Cragg (1971) offers a reformulation of the model that allows the effects of explanatory variables to vary between the selection and substantive equations, and Smith and Brame (2003) discuss the advantages of this model. Useful elaborations of Tobit regression that are relatively accessible are provided in Breen (1996), Long (1997), and Osgood et al. (2002).

The Tobit model is appropriate when there is explicit selection with censored data. As Table 1 suggests, however, explicit selection may also occur with truncated samples in which case information is missing on both the independent and dependent variables in omitted cases. In such a scenario, truncated regression models can be employed to correct model estimates

(represented by the upper left-hand box in Table 1). Because implicit selection processes require information on the explanatory variables to model the stochastic selection process, there is no straightforward solution to sample selection involving truncated samples with implicit selection (represented by the upper right-hand box in Table 1). The truncated regression model is similar to the Tobit model in that the dependent variable Y is only observed if some censoring value c is exceeded. In the Tobit model, however, the selection process is captured using a probit model, whereas in a truncated model the requisite information for the probit equation is now missing. However, the log-likelihood function for the truncated regression model can be estimated using maximum likelihood. It is similar to the Tobit log-likelihood function for the uncensored observations, but the density function for the truncated normal distribution includes an additional term that captures the truncation point. In essence, the model assumes that the conditional distribution of the dependent variable is normal and that the excluded values, if observed, would also follow a normal distribution. The truncated regression model has not been used extensively in criminological research but Hausman and Wise (1977) and Breen (1996) provide useful elaborations of it.

Heckman's Two-Part Sample Selection Model

Perhaps the most widely used approach for addressing sample selection bias in criminology has been Heckman's two-part sample selection model. Figure 3 makes it clear that the model is appropriate when the researcher is faced with incidental selection in a censored sample. Unlike the Tobit model, which assumes a single process characterizes both the selection process and the outcome of interest, Heckman's sample selection model provides for a separate stochastic process that determines selection into the sample. The first step in calculating this model is to estimate a probit model for the selection process, which is then used to control for the probability of being excluded from the sample in the substantive equation of interest. In essence, the two-part model calculates a predicted probability of

being excluded from the observed sample, and then it controls for that probability in the substantive equation of interest. The statistical model that summarizes Heckman's two-part solution for selection bias is represented by Eqs. 7.1 and 7.2 below:

$$P = \gamma Z + \delta \quad (7.1)$$

$$Y = \beta X + \sigma \rho \lambda(\beta X) + \varepsilon \quad (7.2)$$

Comparison of Eqs. 7.1 and 7.2 with 3.1 and 3.2 makes it clear that the essential difference between the simple uncorrected two-part model and Heckman's two-part sample selection model is that the latter includes the additional regressor $\sigma \rho \lambda(\beta X)$, where $\lambda(\beta X)$ is equal to $\varphi(\beta X)/(1 - \Phi(\beta X))$, or the inverse Mills' ratio, which represents "a monotone decreasing function of the probability that an observation is selected into the sample" (Heckman 1979, p. 156). This is also sometimes referred to as the "hazard rate" because it captures the instantaneous probability of exclusion conditional on being at risk for each observation. The ρ in Eq. 3.2 represents the correlation between the two error terms, δ and ε , which is assumed to follow a bivariate normal distribution, and the σ parameter represents the effect of selection on the outcome Y . If the correlation between the two errors is zero ($\rho = 0$), then the extra parameter drops out of the equation and there is no selection bias.

To calculate Heckman's sample selection model, the probit equation is first estimated and the predicted probabilities are retained and then subsequently used to calculate the inverse Mills' ratio described above. Using the probit equation, expected values for the residuals for each individual in the data are saved, conditional on the observed variables Z , and whether or not an individual was selected ($P = 1$) or not ($P = 0$) (i.e., $E[\delta|X, P]$) (see Heckman 1978, p. 938). If the case is selected (i.e., if $P = 1$), then $E[\delta|X, P = 1]$ is equal to $\frac{\varphi(\hat{P})}{\Phi(\hat{P})}$ where φ and Φ represent the standard normal density and cumulative distribution functions, respectively. If the

case is not selected, (i.e., if $P = 0$), then $E[\delta/X, P = 0]$ is equal to $\frac{-\varphi(\hat{P})}{[1-\Phi(\hat{P})]}$. In essence, the conditional residuals are used to capture information for each individual on unobservable factors that are related to the selection process. To illustrate, consider two individuals, one whose predicted value is 0.5 and one whose predicted value is -0.5 . Calculating the probabilities of being selected into the sample, we get a value of 0.69 for the first individual and a value of 0.31 for the second. Now imagine that both individuals are selected, even though the first individual had a much higher predicted probability. This suggests that the individual with the lower predicted probability who was nevertheless selected into the sample was more likely to have had unobserved factors that were positively associated with selection. Among cases that are selected, then, those with lower probabilities should have larger (i.e., more positive) residuals. If we plug the hypothetical probabilities above into the expectation for $E[\delta/X, P]$, we get a value of 1.14 for the case with the lower probability of selection and a value of only 0.51 for the case with the higher probability of selection. The values of the conditional residuals from the selection equation therefore serve to help capture the unobserved factors that are related to selection. In a sense, the Heckman model uses the conditional residuals from the selection equation to create an additional regressor that helps to control for unobserved heterogeneity.

This additional regressor is included in the substantive equation of interest in Eq. 3.2 to correct for selection bias. As Heckman (1979) shows, the least squares estimator will be unbiased but inefficient. Inefficiency is introduced by the censoring of the data, which produces heteroskedasticity in the error terms and leads to downwardly biased estimates of the population variance. This necessitates a correction for the standard errors in the model. According to Bushway et al. (2007), standard errors on average can be underestimated by between 10 % and 30 %. These authors identify a number of common errors that characterize prior attempts to use Heckman's model in criminological applications,

such as the miscalculation of the inverse Mills' ratio and failure to properly correct for misestimated standard errors. These issues are now easily addressed by canned routines for Heckman's model that are now readily available in standard statistical packages and properly calculate lambda and correct the standard errors. In Stata, for instance, the basic two-step sample selection model can be estimated as follows: "*heckman Y X, twostep select(Z)*" where Y is the dependent variable, X is the vector of independent variables included in the substantive equation of interest, and Z is the vector of independent variables included in the probit selection equation. All that the researcher has to do is to specify a sample selection model and then identify which exogenous variables to include in the selection and substantive equations, respectively. In addition to the two-step procedure outlined above, the Heckman model can also be estimated using maximum likelihood. In general, the maximum likelihood estimator will be more efficient than the two-stage estimator though it may be more sensitive to violations of model assumptions (Bushway et al. 2007).

Importantly, when Z and X contain the same set of independent variables, the Heckman sample selection model is only identified by the nonlinearity introduced by the inverse Mills' ratio. Prior research demonstrates that this can lead to severe inflation of standard errors, which results from collinearity between the correction term and the other regressors in the model (Stolzenberg and Relles 1990). When the same predictors are used in the selection and substantive equations, the sigma term in Eq. 7.2 will be correlated with X . Therefore, whenever possible, *exclusion restrictions* should be included in the sample selection model. Exclusion restrictions are similar to instrumental variables – they are predictors that are related to the selection process but not related to the substantive outcome. For example, Albonetti (1991) included evidentiary measures in her selection equation for the likelihood of conviction, but not her substantive equation for sentence severity. Similarly, Smith and Paternoster (1990) included the caseload and number of diversionary programs in their

selection model for juvenile placement but not for their substantive equation examining recidivism. The key is to identify factors that are theoretically related to the selection process but not to the outcome of interest, and whenever possible, researchers should search for and include valid exclusion restrictions when estimating this model.

The Bivariate Probit Model

Heckman's sample selection model applies to the case in which the substantive outcome of interest is a continuous variable. When the dependent variable is dichotomous, as is often the case in criminological research, then an alternative specification can be employed. The bivariate probit model involves estimation of two probit equations allowing their error terms to be correlated. This is useful when there is a binary selection process and a binary outcome in the substantive equation of interest (Heckman 1976). Whenever two dichotomous dependent variables are jointly determined, or when there is reason to believe that error terms of two binary outcomes are correlated, then the bivariate probit model can offer a useful approach for modeling this dependence. The basic model is summarized in Eqs. 8.1 and 8.2:

$$P_2 = \gamma Z + \delta \quad (8.1)$$

$$P_1 = \beta X + \varepsilon \quad (8.2)$$

If the two error terms in Eqs. 8.1 and 8.2 are independent (i.e., $Cov(\delta, \varepsilon) = 0$), then selection bias is not a problem and the two equations can be estimated separately. If, however, the two errors are related (e.g., $(\delta = \eta + u_1)$ and $(\varepsilon = \eta + u_2)$), so that part of the error (i.e., (η)) is common to both equations, then the same type of selection bias discussed above will bias model estimates. When this is the case, interest shifts to the joint probability of P_2 and P_1 . The model assumes the error terms in each equation are normally distributed and that their joint distribution is bivariate normal, and as with the Heckman model above, the association between errors is represented by the correlation coefficient ρ . The bivariate

probit can be estimated using maximum likelihood procedures that are readily available in common statistical packages. In Stata, for instance, the bivariate probit model can be estimated as follows: "*heckprob P₁ X, select(P₂)*" where P_1 is the second probit outcome, X is the vector of independent variables included in that equation, and Z is the vector of independent variables included in the first probit equation.

The bivariate probit can be used in basic applications to allow the two outcomes to be correlated, or it can be used equivalently to the Heckman selection model to capture a binary selection process that determines inclusion in a binary substantive outcome of interest. Paternoster and Brame (1998), for instance, investigated whether or not criminal behavior was associated with analogous noncriminal behavior (Paternoster and Brame 1998) where both criminal and analogous behaviors were measured as dichotomous outcomes. They used a bivariate probit to model the joint distribution of the two outcomes in order to estimate the strength of their association. Berinsky (2004), on the other hand, utilized the bivariate probit to explicitly correct for sample selection, where the first model was whether or not a respondent answered a survey question and the second was a dichotomous measure of racial attitudes. For the bivariate probit, it is essential that valid exclusion restrictions be included. Dubin and Rivers (1990) discuss application of the binary selection model, and Sartori (2003) offers an alternative estimator that may be preferable if exclusion restrictions are unavailable.

Contemporary Controversies

A number of scholars both in and outside of criminology have discussed at length the potential strengths and weaknesses of sample selection models (Bushway et al. 2007; Winship and Mare 1992; Duan et al. 1983, 1984; Manning et al. 1987; Leung and Yu 1996; Jones 2000; Norton et al. 2008). As Breen (1996, p. 58) summarized, "it is now clear that these models need to be approached with a degree of care." Sample

selection models work the best when there are strong exclusion restrictions and when distributional assumptions are clearly met. As others have shown, routinized application of sample selection models without careful circumspection can produce estimates that are no better and sometimes worse than those obtained through ordinary regression techniques obtained from simple two-part models (Duan et al. 1983, 1984). A number of potential concerns need to be considered when deciding whether or not it is advisable to employ a sample selection model.

First, sample selection models have been shown to be sensitive to violations of their distributional assumptions, including heteroskedasticity of errors and violations of normality. Whereas OLS estimates are consistent but not efficient in the presence of heteroskedasticity, estimates from sample selection models are neither consistent nor efficient (Amemiya 1985). Alternative functional forms can be specified for error distributions to address the first issue (Maddala 1983), and a semi-parametric or non-parametric methods that relax model assumptions might be used to address the latter (Heckman et al. 1999; Vella 1999; Moffitt 1999; Chay and Powell 2001). Winship and Mare (1992) discuss some of these approaches, though applications of them remain rare in criminological work. Still, the first step in deciding whether or not sample selection models are appropriate is to examine the validity of distributional assumptions regarding homoskedasticity and normality – Chesher and Irish (1987) provide a useful discussion of how this can be approached. Even if we assume that distributional assumptions are met, though, the performance of sample selection models is affected by additional factors that include the strength of the correlation (ρ) between the error terms in the selection and substantive equations, the degree of overlap in the variables included in the selection and substantive equations, and the size of the sample and degree of censoring in the data. The stronger the correlation is between errors, the larger the threat of sample selection bias and, in general, the greater need for sample selection corrections. Therefore, a researcher needs to have strong theory regarding the

association between the selection process and the outcome of interest to help guide the decision to use sample selection models.

Second, it is important for researchers to carefully search for, identify, and incorporate theoretically valid exclusion restrictions into their sample selection models. A number of commentators have expressed reservations about identifying the model using only the nonlinear functional form (e.g., Berk and Ray 1982). Doing so typically leads to high variance estimates and often results in problematic levels of multicollinearity between the correction term and the other independent variables in the model. The problem tends to be exacerbated when the selection model poorly fits the data and has low levels of explained variance. If we do a poor job of capturing the selection process, then the correction for selection will have limited utility.

Third, the performance of sample selection models is also tied to the degree of censoring in the data. In samples with minimal censoring, the sample selection models may not be needed. Low degrees of censoring will introduce less bias into parameter estimates than high degrees of censoring. However, in samples with high degrees of censoring, these models can also be problematic. Stolzenberg and Relles (1990, p. 406), for instance, examined the effectiveness of sample selection corrections in small samples with severe censoring and concluded that “Heckman’s method reduced the accuracy of coefficient estimates as often as it improved them.” They concluded that “there is considerable evidence that the method can easily do more harm than good, and that its careless and mechanical application runs much danger” (Stolzenberg and Relles 1990, pp. 408–409). To abandon sample selection models altogether, however, would be to throw out the baby with the bathwater. In general, larger samples are preferable to smaller samples when implementing sample selection corrections, whereas the simple uncorrected two-part model is likely to be preferred when explained variance is low or when the estimate of the correlation between error terms is near zero. When sample selection models are deemed appropriate, the maximum likelihood estimators

will be more efficient than Heckman's original two-step approach. In practice, the effects of the size of ρ , the overlap in the predictors in the two equations, and the sample size and degree of censoring are often intertwined and can be difficult to disentangle. The researcher is therefore often faced with the unhappy decision to either estimate the simple two-part model, which will knowingly produce biased estimates, or alternatively to implement a sample selection model that can under certain conditions make estimates worse. Luckily, there are several approaches that can help assess the feasibility of implementing sample selection models to correct for selection bias.

Lueng and Yu (1996, 2000) discuss the conditions under which sample selection models can suffer from multicollinearity problems, and they discuss some practical approaches for dealing with this problem. Specifically, they recommend researchers calculate the "condition number" as a measure of collinearity, and they recommend that condition numbers over 20 indicate Heckman's correction is likely to perform worse than simple uncorrected two-part models. Other diagnostic tests are also available. For instance, Manski (1990) outlines a useful approach for bounding estimates of selection bias (see also Nagin and Manski 1999). The interval of the bounds varies inversely with the degree of censoring so lower proportions of censored cases result in narrower bounding ranges. Winship and Mare (1992) discuss his method and related approaches. More recently, Stolzenberg and Relles (1997) have outlined a useful methodology for estimating the amount of bias associated with specific parameter estimates in a model, and Bushway et al. (2007) demonstrate the utility of this approach by applying it to the context of criminal sentencing. The approach allows one to quantify the amount of bias introduced by the selection process, providing for better informed decisions regarding the necessity of using sample selection models in place of ordinary regression. Before deciding whether or not to implement sample selection models, then, it is instructive to first examine distributional assumptions, to carefully search for valid exclusion restrictions,

and to conduct exploratory analyses that help assess the degree of the problem and the necessity of addressing it.

Conclusion

Sample selectivity has long been a topic of applied theoretical and methodological work in sociology and economics (Heckman 1974; Winship and Mare 1992; LaLonde 1986; Berk 1983), and in recent years it has become a popular analytical technique in criminology and criminal justice (Bushway et al. 2007). Although sample selection models can provide a very useful approach for addressing important selection biases that are pervasive in criminological research, "Infallible models for sample selection bias do not exist" (Winship and Mare 1992, p. 347) and "whatever happens to be the methodological tool du jour is too often oversold" (Berk 2005, p. 2). This suggests the need for careful consideration of the potential strengths and weaknesses of sample selection corrections in applied criminological work. As Winship and Mare (1992, p. 329) argue, "models for selection bias are only as good as their assumptions about the way that selection occurs, and estimation strategies are needed that are robust under a variety of assumptions."

Numerous extensions of these basic models are continuously being developed. Outcome measures are at times available for both the selected and unselected samples (e.g., Spohn and Holleran 2002). Moreover, it is not uncommon for certain applications to include multiple selection points across multiple outcomes of interest. Research on the criminal justice system, for instance, inherently involves multiple censoring points across successive stages of criminal case processing (Kingsnorth et al. 1999). In some applications, multiple censoring or truncation points may even be different for different subjects in a sample, and in general, the basic models discussed herein can readily accommodate such complexities without drastically increasing the difficulty of their application (Breen 1996; Maddala 1983).

The challenge for criminologists faced with sample selection then is to be diligent and circumspect, carefully formulating a theoretical model of the selection process before selecting an appropriate modeling strategy. Importantly, there are a number of different approaches that can be useful for addressing different forms of selection bias, such as experiments, statistical matching procedures, and instrumental variable approaches (e.g., Bushway and Apel 2010; Loughran and Mulvey 2010; Taxman and Rhodes 2010), and in many ways models for sample selection share considerable affinity with other quasi-experimental approaches for addressing selection bias (Lalonde 1986; Heckman and Robb 1986). Researchers should therefore investigate sample selection concerns through a variety of complementary methods whenever possible. Smith and Paternoster (1990) offer an excellent example of this type of research. They rightfully argue that “there is no one generic cure for selectivity bias and that applications of these models must make substantive sense in the context of specific applications” (Smith and Paternoster 1990, p. 1125). Similarly, Winship and Mare (1992, p. 328) conclude that “The large number of available methods and the difficulty of modeling selection indicate that researchers should be explicit about the assumptions behind their methods and should present results that derive from a variety of methods.” The sample selection models discussed in this entry offer one useful approach for investigating and potentially correcting for selection bias when researchers are faced with censored or truncated data on their dependent variable of interest.

Related Entries

► [Sample Selection Problems](#)

Recommended Reading and References

- Albonetti CA (1991) An integration of theories to explain judicial discretion. *Soc Probl* 38:247–266
- Amemiya T (1985) *Advanced econometrics*. Harvard University Press, Cambridge, MA
- Angrist J (2001) Estimation of limited dependent variable models with dummy endogenous regressors: simple strategies for empirical practice. *J Bus Econ Stat* 19:2–16
- Apel R, Bushway S, Paternoster R, Brame R, Sweeten G (2008) Using state child labor laws to identify the causal effect of youth employment on deviant behavior and academic achievement. *J Quant Criminol* 24:337–362
- Berinski AJ (2004) *Public opinion and political participation in America*. Princeton University Press, Princeton
- Berk R (1983) An introduction to sample selection bias in sociological data. *Am Socio Rev* 48:386–398
- Berk R (2005) Randomized experiments as the bronze standard. *J Exp Criminol* 1:417–433
- Berk R, Ray S (1982) Selection bias in sociological data. *Soc Sci Res* 11:352–398
- Brame R (2000) Investigating treatment effects in a domestic violence experiment with partially missing outcome data. *J Quant Criminol* 16:283–314
- Breen R (1996) *Regression models: censored, sample selected, or truncated data*. Sage, Thousand Oaks, CA
- Bushway SD, Apel RJ (2010) Instrumental variables in criminology and criminal justice. In: Piquero AR, Weisburd D (eds) *Handbook of quantitative criminology*. Springer, New York, pp 595–612
- Bushway S, Johnson BD, Slocum LA (2007) Is the magic still there? The use of Heckman two-step correction for selection bias in criminology. *J Quant Criminol* 23:151–178
- Chay KY, Powell JL (2001) Semiparametric censored regression models. *J Econ Perspect* 15:29–42
- Chesher A, Irish M (1987) Residual analysis in the grouped and censored normal linear model. *J Econometrics* 34:33–61
- Cragg J (1971) Some statistical models for limited dependent variables with applications to the demand for durable goods. *Econometrica* 25:241–262
- D’Alessio S, Stolzenberg L (2003) Race and the probability of arrest. *Soc Forces* 84:1381–1398
- Duan N, Manning WG, Morris CN, Newhouse JP (1983) A comparison of alternative models for the demand for medical care. *J Bus Econ Stat* 1:115–126
- Duan N, Manning WG, Morris CN, Newhouse JP (1984) Choosing between the sample-selection model and the multi-part model. *J Bus Econ Stat* 2:283–289
- Dubin JA, Rivers D (1990) Selection bias in linear regression, logit and probit models. In: Fox J, Long JS (eds) *Modern methods of data analysis*. Sage, Thousand Oaks, pp 411–443
- Gondolf EW (2000) A 30-month follow-up of court-referred batterers in four cities. *Int J Offender Ther Comp Criminol* 44:111–128
- Greene WH (1993) *Econometric analysis*, 2nd edn. Macmillan, New York
- Hagan J (1990) The gender stratification of income inequality among lawyers. *Soc Forces* 68:835–855

- Hagan J, Palloni A (1986) 'Club Fed' and the sentencing of white-collar offenders before and after Watergate. *Criminology* 24:603–622
- Hagan J, Parker P (1985) White-collar crime and punishment: the class structure and legal sanctioning of security violations. *Am Sociol Rev* 50:302–316
- Hausman J, Wise DA (1977) Social experimentation, truncated distributions, and efficient estimation. *Econometrica* 45:919–938
- Heckman J (1974) Shadow prices, market wages, and labor supply. *Econometrica* 42:679–694
- Heckman JJ (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables. *Ann Econ Soc Meas* 5:475–492
- Heckman J (1978) Dummy endogenous variables in a simultaneous equation system. *Econometrica* 46:931–959
- Heckman JJ (1979) Sample selection bias as a specification error. *Econometrica* 47:153–161
- Heckman J, Robb R (1986) Alternative methods for solving the problem of selection bias in evaluating the impact of treatments on outcomes. In: Wainer H (ed) *Drawing inference from self-selected samples*. Springer, New York, pp 63–107
- Heckman JJ, LaLonde R, Smith J (1999) The economics and econometrics of active labor market programs. In: Ashenfelter O, Card D (eds) *Handbook of labor economics*, vol 3A. North-Holland, Amsterdam, pp 1865–2097
- Jones AM (2000) Health econometrics. In: Culyer AJ, Newhouse JP (eds) *Handbook of health economics*, 1st edn, vol 1, chapter 6. Elsevier, pp 265–344
- Kingsnorth R, MacIntosh R, Wentworth J (1999) Sexual assault: the role of prior relationship and victim characteristics in case processing. *Justice Q* 16:275–302
- Klepper S, Nagin D, Tierney LJ (1983) Discrimination in the criminal justice system: a critical appraisal of the literature. In: Blumstein A, Cohen J, Martin S, Tonry M (eds) *Research on sentencing: the search for reform*, vol II. National Academy, Washington, DC
- LaLonde RJ (1986) Evaluating the econometric evaluations of training programs with experimental data. *Am Econ Rev* 76:604–620
- Laub H, Sampson RJ (2003) *Shared beginnings, divergent lives: delinquent boys to age 70*. Harvard University Press, Cambridge
- Lemert EM (1945) *Social pathology: systematic approaches to the study of sociopathic behavior*. McGraw-Hill, New York
- Leung SF, Yu S (1996) On the choice between sample selection and two-part models. *J Econ* 72:197–229
- Leung SF, Yu S (2000) Collinearity and two-step estimation of sample selection models: problems, origins, and remedies. *Comput Econ* 15:173–199
- Little RJA, Rubin DB (1987) *Statistical analysis with missing data*. J Wiley & Sons, New York
- Long JS (1997) *Regression models for categorical and limited dependent variables*. Sage, Thousand Oaks
- Loughran T, Mulvey E (2010) Estimating treatment effects: matching quantification to the question. In: Weisburd D, Piquero A (eds) *Handbook of quantitative criminology*. Springer, New York, pp 163–181
- Lundman R, Kaufman R (2003) Driving while black: effects of race, ethnicity, and gender on citizen self-reports of traffic stops and police actions. *Criminology* 41:195–220
- Maddala GS (1983) *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press, Cambridge
- Manning WG, Newhouse JP, Duan N, Keeler E, Benjamin B, Leiboqitz A, Marquis AS, Zwanziger J (1987) Health insurance and the demand for medical care: evidence from a randomized experiment. Publ. No. R-3476-HHS. Rand Corporation, Santa Monica
- Manski CF (1990) Nonparametric bounds on treatment effect. *Am Econ Rev* 80:319–323
- Maxwell C, Garner J, Fagan J (2002) The preventative effects of arrest on intimate partner violence. *Criminol Public Policy* 2:51–80
- Moffitt R (1999) Econometric methods for labor market analysis. In: Ashenfelter O, Card D (eds) *Handbook of labor economics*, vol 3A. North-Holland, Amsterdam
- Mosher CJ, Miethe TD, Hart TC (2011) *The mismeasure of crime*, 2nd edn. Sage, Thousand Oaks, CA
- Nagin D, Manski C (1999) Bounding disagreements about treatment effects: a case study of sentencing and recidivism. *Sociol Methodol* 28:99–138
- Nagin D, Waldfoegel J (1998) The effect of conviction on income through the life cycle. *Int Rev Law Econ* 18:25–40
- Norton EC, Dow WH, Do YK (2008) Specification tests for sample selection and two-part models. *Health Serv Outcomes Res Methodol* 8:201–208
- Osgood DW, Finken L, McMorris B (2002) Analyzing multiple item measures of crime and deviance II: Tobit regression analysis of transformed scores. *J Quant Criminol* 18:859–866
- Paternoster R, Brame R (1998) The structural similarity of processes generating criminal and analogous behaviors. *Criminology* 36:633–669
- Peterson RD, Hagan J (1984) Changing conceptions of race: toward an account of anomalous findings of sentencing research. *Am Sociol Rev* 49:56–70
- Puhani PA (2000) The Heckman correction for sample selection and its critique. *J Econ Surv* 14:53–68
- Reisig MD, Parks RB (2000) Experience, quality of life, and neighborhood context: a hierarchical analysis of satisfaction with police. *Justice Q* 17:607–630
- Rhodes W (1991) Federal criminal sentencing: some measurement issues with application to pre-guideline sentencing disparity. *J Crim Law Criminol* 81:1002
- Rhodes W, Pelissier B, Gaes G, Saylor B, Camp S, Wallace S (2001) Alternative solutions to the problem of selection bias in an analysis of federal residential drug treatment programs. *Eval Rev* 25:331–369

- Sampson RJ (2010) Gold standard myths: observations on the experimental turn in quantitative criminology. *J Quant Criminol* 25:489–500
- Sampson RJ, Morenoff JD, Gannon-Rowley T (2002) Assessing “neighborhood effects”: social processes and new directions in research. *Annu Rev Sociol* 28:443–78
- Sartori AE (2003) An estimator for some binary-outcome selection models without exclusion restrictions. *Polit Anal* 11:111–138
- Smith DA, Brame R (2003) Tobit models in social science research: some limitations and a more general alternative. *Sociol Method Res* 31:364–388
- Smith DA, Paternoster R (1990) Formal processing and future delinquency: deviance amplification as selection artifact. *Law Soc Rev* 24:1109–1131
- Spohn C, Holleran D (2002) The effect of imprisonment on recidivism rates of felony offenders: a focus on drug offenders. *Criminology* 40:329–358
- Stolzenberg RM, Relles DA (1990) Theory testing in a world of constrained research design: the significance of Heckman’s censored sampling bias correction for nonexperimental research. *Sociol Method Res* 18:395–415
- Stolzenberg RM, Relles DA (1997) Tools for intuition about sample selection bias and its correction. *Am Sociol Rev* 62:494–507
- Taxman FS, Rhodes A (2010) Multisite trials in criminal justice settings: trials and tribulations of field experiments. In: Piquero A, Weisburd D (eds) *Handbook of quantitative criminology*. Springer, New York, pp 519–543
- Tobin J (1958) Estimation of relationships for limited dependent-variables. *Econometrica* 26:24–36
- Ulmer JT, Johnson BD (2004) Sentencing in context: a multilevel analysis. *Criminology* 42:137–178
- Vella F (1999) Estimating models with sample selection bias: a survey. *J Hum Resour* 33:127–169
- Weisburd D (2010) Justifying the use of non-experimental methods and disqualifying the use of randomized controlled trials: challenging folklore in evaluation research in crime and justice. *J Exp Criminol* 6:209–227
- Western B, Pettit B (2005) Black-white wage inequality, employment rates, and incarceration. *Am J Sociol* 111(2):553–578
- Wilson DB (2010) Meta-analysis. In: Weisburd D, Piquero A (eds) *Handbook of quantitative criminology*. Springer, New York, pp 181–208
- Winship C, Mare RD (1992) Models for selection bias. *Annu Rev Sociol* 18:327–350
- Worrall JL (2002) If you build it, they will come: consequences of improved citizen complaint review procedures. *Crime Delinq* 48:355–379
- Zatz MS, Hagan J (1985) Crime, time, and punishment: an exploration of selection bias in sentencing research. *J Quant Criminol* 1:103–126

Sample Selection Problems

Robert Apel

School of Criminal Justice, Rutgers University,
Newark, NJ, USA

Overview

In a seminal series of contributions that would culminate in his being awarded the Nobel Prize in Economics in 2000, James Heckman proposed a deceptively simple solution to the problem posed by sample selection bias (Heckman 1976, 1979). In the classic formulation of the problem, the analyst desires to estimate a wage or earnings function from a collection of individuals sampled randomly from the population. While one wishes to generalize to the population of all wage offers, it is obvious that wages are only observed for individuals who are employed. The problem that arises is that these individuals are a self-selected (i.e., nonrandom) subset of the population, and consequently their wages cannot necessarily be used to estimate what the wages of nonworking individuals would have been, had they chosen to work. So a regression of (log) wages on a determinant such as schooling, limited to the subsample of workers, will potentially yield biased estimates of the impact of schooling on wages for the population of interest.

In criminology, a similar challenge is encountered in studies of the determinants of criminal sentencing. For example, one might be interested in studying the impact of legally irrelevant factors (e.g., gender, race) on sentencing decisions once legally relevant factors (e.g., offense seriousness, criminal history) have been netted out. However, sentence lengths are only observed for convicted offenders who are sent to prison. The determinants of sentence length among these individuals are not necessarily representative of what the determinants of sentence length would have been among non-incarcerated individuals, had they instead been incarcerated. Whereas the population of interest entails all criminal

sentences rather than strictly carceral ones, the latter are a decidedly nonrandom subset of all convicted offenders. Many other examples that have criminological relevance are provided in Berk (1983), Winship and Mare (1992), Bushway et al. (2007).

Heckman (1976, 1979) formulated the problem as one of missing data giving rise to specification error. If the process that determines whether an individual is “uncensored” with respect to the outcome of interest (e.g., employed, incarcerated) is correlated with the process that determines the mean outcome, then the least squares estimator is biased and inconsistent. Heckman’s insight was that exclusion of censored units from a sample will lead to bias and inconsistency, even if one is only interested in generalizing to the population of uncensored units (Berk 1983).

This chapter is devoted to an investigation of the sample selection problem, a presentation of the various solutions that have been proposed, and a discussion of the limitations of these solutions. An empirical illustration will then be provided, using data from the National Longitudinal Survey of Youth 1997 to examine the relationship between criminal behavior and legal earnings.

The Sample Selection Problem

The sample selection problem is one of estimating the relationship between two or more variables for a population, when information on the dependent variable is only observed for a nonrandom subset of a sample from that population. Consider estimating a model of earnings from employment. The population equation of interest is the following:

$$\begin{aligned}
 Y_i^* &= \beta_0 + \beta_1 X_{1i} + \dots + \beta_j X_{ji} + e_i \\
 &= \mathbf{X}\boldsymbol{\beta} + e_i
 \end{aligned}
 \tag{1}$$

where the coefficients and regressors are consolidated into $\mathbf{X}\boldsymbol{\beta}$ for notational convenience. Here, Y_i^* is a random variable representing *latent earnings*. The sample selection problem arises

because Y_i^* is only partially observed – earnings are unobserved or censored among the subsample of individuals who are not employed. A second random variable, which can be thought of as *employment propensity*, determines whether earnings are observed or censored. This is also a latent variable, denoted D_i^* , that is parameterized in the following manner:

$$\begin{aligned}
 D_i^* &= \alpha_0 + \alpha_1 Z_{1i} + \dots + \alpha_k Z_{ki} + u_i \\
 &= \mathbf{Z}\boldsymbol{\alpha} + u_i
 \end{aligned}
 \tag{2}$$

where the coefficients and regressors are consolidated into $\mathbf{Z}\boldsymbol{\alpha}$ for notational convenience. Concerning the residuals from these two equations, e_i is assumed to be distributed normal and u_i is assumed to be distributed standard normal. Their joint distribution is assumed to be bivariate normal:

$$\begin{pmatrix} e_i \\ u_i \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_e^2 & \sigma_{e,u} \\ \sigma_{e,u} & 1 \end{pmatrix} \right]$$

As an aside, the correlation between the unobserved determinants of the two latent response variables Y_i^* and D_i^* is provided by the equation:

$$\rho_{e,u} = \frac{\sigma_{e,u}}{\sigma_e}$$

Berk (1983) refers to Eq. 1 as the “substantive equation,” and to Eq. 2 as the “selection equation.” Of course, neither Y_i^* nor D_i^* are fully observed. Instead, the two latent variables have observed counterparts. Namely, latent earnings (Y_i^*) are only observed when employment propensity (D_i^*) exceeds a defined threshold:

$$Y_i = \begin{cases} 0 & \text{if } D_i^* < 0 \\ Y_i^* & \text{if } D_i^* > 0 \end{cases}$$

Note that the choice of 0 as the threshold is a convenient normalization. If we substitute the linear predictor in Eq. 2 for D_i^* and rearrange terms, it can be shown that latent earnings are only observed if:

$$u_i > -\mathbf{Z}\boldsymbol{\alpha}$$

Taking the expectation of the observed Y_i for the subsample with nonzero earnings, we now have the following:

$$E(Y_i|W_i, D_i^* > 0) = \mathbf{X}\boldsymbol{\beta} + E(e_i|u_i > -\mathbf{Z}\boldsymbol{\alpha}) \quad (3)$$

The last term in Eq. 3 illustrates the source of the sample selection problem and is the key insight provided by Heckman (1976, 1979). Namely, a least squares regression model of earnings, fitted only for the subsample with nonzero earnings, can lead to omitted variables bias because of the exclusion of a potentially relevant regressor – the conditional mean of the latent earnings residual for the uncensored subsample. If this last term is 0, there is no omitted variables bias, and inferences based on the subsample with nonzero earnings are valid for the general population, including those with censored earnings. Obviously, the sample selection problem arises when the conditional mean is not 0.

Consider the implications of the above for the question of interest in this chapter – the relationship between arrest and employment earnings. It is plausible that this relationship, estimated only from the subsample of individuals with nonzero earnings, is rendered biased and inconsistent because of the fact that these individuals are not a random subset of the sample. Specifically, the unobservables which determine whether or not someone works are likely to be correlated with the observables which determine someone’s mean earnings, conditional on working. This means that regressors which are truly correlated with earnings might be statistically insignificant when the model is estimated from individuals with nonzero earnings. Similarly, regressors which are truly uncorrelated with earnings might be statistically significant.

Heckman’s Two-Step Solution

To overcome the sample selection problem, Heckman (1976, 1979) proposed an elegant, two-step solution – the so-called Heckman

selection correction. His solution entails direct estimation of the selection term in Eq. 3 and its inclusion as a regressor in the equation of substantive interest. The first step of this procedure is to estimate a model of the probability of being uncensored, in this case, the probability of having nonzero earnings. This begins by linking employment propensity (D_i^*), which is unobserved, to a binary response variable according to the following rule:

$$D_i = \begin{cases} 0 & \text{if } D_i^* < 0 \\ 1 & \text{if } D_i^* > 0 \end{cases}$$

This binary variable is then treated as the dependent variable in a probit regression model:

$$\begin{aligned} \Phi^{-1}[Pr(D_i = 1)] &= \alpha_0 + \alpha_1 Z_{1i} + \dots \\ &\quad + \alpha_k Z_{ki} \\ &= \mathbf{Z}\boldsymbol{\alpha} \end{aligned} \quad (4)$$

The second step of Heckman’s procedure requires inserting a transformation of the linear predictor from Eq. 4 into the substantive equation, estimated from the subsample with positive earnings:

$$\begin{aligned} (Y_i|D_i = 1) &= \beta_0 + \beta_1 X_{1i} + \dots + \beta_j X_{ji} \\ &\quad + \beta_{j+1} \lambda(-\mathbf{Z}\boldsymbol{\alpha}) + e_i \end{aligned} \quad (5)$$

The new term, $\lambda(\cdot)$, is known as the “inverse Mills ratio” (IMR) of the argument:

$$\lambda(-\mathbf{Z}\boldsymbol{\alpha}) = \frac{\phi(-\mathbf{Z}\boldsymbol{\alpha})}{1 - \Phi(-\mathbf{Z}\boldsymbol{\alpha})} = \frac{\phi(\mathbf{Z}\boldsymbol{\alpha})}{\Phi(\mathbf{Z}\boldsymbol{\alpha})}$$

The numerator $\phi(\cdot)$ represents the standard normal density function (p.d.f.) evaluated at the argument, whereas the denominator $\Phi(\cdot)$ represents the standard normal distribution function (c.d.f.) evaluated at the same argument. In empirical applications, of course, the true IMR is unknown. Instead, the estimate of the linear predictor from the probit regression model in Eq. 4 is substituted into the ratio:

$$\hat{\lambda}(-\mathbf{Z}\boldsymbol{\alpha}) = \frac{\phi(\hat{\alpha}_0 + \hat{\alpha}_1 Z_{1i} + \dots + \hat{\alpha}_k Z_{ki})}{\Phi(\hat{\alpha}_0 + \hat{\alpha}_1 Z_{1i} + \dots + \hat{\alpha}_k Z_{ki})}$$

The IMR has a number of important properties. For one, it is nonnegative and is therefore bounded by 0 and ∞ . In addition, because the denominator yields the estimated probability that an individual is uncensored (i.e., has positive earnings), individuals with a small probability of being uncensored will tend to have a large $\lambda(\cdot)$. Secondly, the parameter β_{j+1} is formally identified in Eq. 5, even when the same set of regressors is included in both the first and second steps of the procedure. This is because $\lambda(\cdot)$ is a nonlinear transformation of the parameter estimates from the first step. Finally, the significance of the estimated coefficient for $\lambda(\cdot)$ is informative about the degree to which nonrandom sample selection introduces bias into the substantive equation. To be precise, the significance of the parameter conforming to $\lambda(\cdot)$ is a test of the significance of the correlation between the unobserved determinants of Y_i^* and D_i^* . Specifically, it can be shown that:

$$\beta_{j+1} = \rho_{e,u} \sigma_e = \sigma_{e,u}$$

As a final technical note on the two-step estimator, it is important to point out that least squares regression of Eq. 5 will yield consistent estimates of the coefficients in $\boldsymbol{\beta}$. The covariance estimator, on the other hand, is inconsistent and will produce misleading standard errors. This is partly because the residuals in the two-step model are heteroscedastic and partly because the selection term represents an estimate rather than the actual selection term. Therefore, robust estimation of the covariance matrix is typically recommended, or else bootstrapping.

Limitations of the Two-Step Procedure

The appeal of Heckman’s solution to the sample selection problem derives from the fact that it is an intuitive and tractable method, requiring only estimation of a probit regression model followed by a linear regression model (although see

Bushway et al. 2007, for examples of misapplication of the method in criminology). The model should not be treated as a panacea to the sample selection problem, however, for it has well-known limitations. Several of these are described below (see also Stolzenberg and Relles 1990, 1997).

One limitation of Heckman’s model in many research settings is the often mechanical application of the procedure. This has been especially pronounced in sociology and criminology (see Bushway et al. 2007). This is largely a consequence of the absence of formal selection models in these disciplines, leading many researchers to conceive of the sample selection problem as a purely technical rather than theoretical issue. In economics, on the other hand, there are often well-defined theories that inform researchers about the proper specification of the selection and substantive models. Theory can provide helpful guidance about regressors which can be included in the selection model but excluded from the substantive model. Such regressors, called “instrumental variables” (IV’s), can improve the empirical performance of the two-step estimator.

A second limitation of Heckman’s model is that the selection term tends to be highly collinear with the regressors in the substantive equation. This is because the same regressors appear in both the selection and substantive equations. High collinearity can make the standard errors of the two-step estimator so inflated, in fact, that it loses its appeal altogether. The trade-off between consistency and efficiency is an important one. Under some circumstances, the inconsistency of the uncorrected least squares estimator is preferred to the inefficiency of the selection-corrected least squares estimator. Although the two-step model is formally identified even when the regressors in the two equations are identical, the estimator has better asymptotic properties when a researcher imposes one or more exclusion restrictions which can be justified a priori. Exclusion restrictions cannot only strengthen the theoretical grounding of the model (as mentioned in the previous paragraph) but can improve the statistical performance of the

two-step estimator. As a matter of empirical practice, then, the use of instrumental variables in the selection model, along with collinearity diagnostics for the substantive model, is recommended.

A third limitation of Heckman's two-step estimator is that it is comparatively less efficient than other estimators. Heckman's model is what is known as a limited information maximum likelihood (LIML) estimator. The two-step approach had practical utility at the time that Heckman (1976, 1979) proposed the model, simply because the alternatives were computationally burdensome. With today's computing capability, however, asymptotic efficiency can be achieved by employing a full information maximum likelihood (FIML) estimator, which simultaneously models the selection and substantive equations. The FIML estimator maximizes the following log likelihood function:

$$\sum_{D_i=0} \ln \Phi(-\mathbf{Z}\boldsymbol{\alpha}) + \sum_{D_i=1} \ln \left[\frac{1}{\sigma_e} \phi \left(\frac{Y_i - \mathbf{X}\boldsymbol{\beta}}{\sigma_e} \right) \right] \\ + \sum_{D_i=1} \ln \Phi \left[\frac{\mathbf{Z}\boldsymbol{\alpha} + \rho_{e,u}(Y_i - \mathbf{X}\boldsymbol{\beta})/\sigma_e}{\sqrt{1 - \rho_{e,u}^2}} \right]$$

An Empirical Illustration of the Two-Step Procedure

Data from the National Longitudinal Survey of Youth 1997 (NLSY97) are used to provide an empirical illustration of the two-step solution to the sample selection problem. The NLSY97 is representative of all youth in the United States born in 1980–1984. This dataset is used to explore the relationship between arrest and earnings from employment. The analysis is limited to the 4,599 males who participated in the survey, among whom 85 % ($N = 3,895$) contributed non-missing data on the variables of interest. At the second interview – the interview from which the outcome variable is measured – the respondents range in age from 13 to 19, with a mean age of 15.9 years. Descriptive statistics for all measures included in the analysis are provided in Table 1.

The outcome in the substantive model is total weekly earnings from employment. This measure is taken from the second wave of the survey (1998) and represents earnings from “formal employment” or “paycheck jobs.” Sample selection bias is a potential concern here, because only half of the sample (50.9 %) is employed between the first and second interviews. Among the subsample that is employed, mean earnings are 147 dollars per month. The earnings distribution among these subsamples is shown in Fig. 1.

The independent variables that are used to model employment earnings include cohort, race (white), schooling (highest grade completed), family structure (live with both biological parents, household size), residential location (central city residence, local unemployment rate), previous work experience (percentage of weeks worked, number of jobs), variety of delinquent behaviors, and total arrests. To ensure proper temporal ordering vis-à-vis earnings, these measures are taken from the first wave of the survey (1997). For this empirical exercise, the regressor of interest is arrest. The question that guides the analysis is, “What impact does arrest have on earnings?”

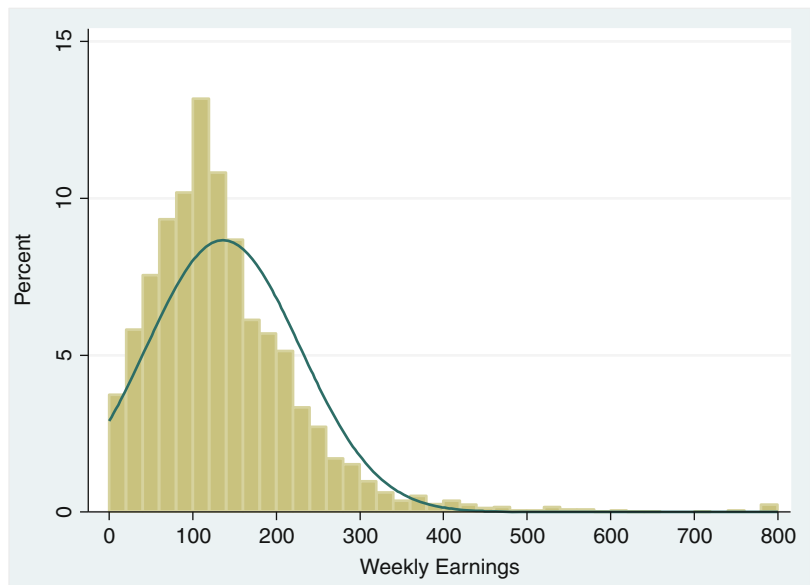
Three instrumental variables are included in the selection model and excluded from the substantive model. These represent variables that are thought to influence whether a respondent is employed, but are not believed to otherwise have any direct influence on earnings among those who are employed. The instrumental variables that are chosen represent provisions of the child labor law which prevails in the state in which a respondent resides as of the second interview (details on the laws are provided in Apel et al. 2008). Three such provisions impose limits on how intensively young people are allowed to work during the school year – the number of hours per week, the number of hours per weekday, and the work curfew on weeknights. For convenience, a binary indicator is constructed for each provision, denoting whether or not a respondent is allowed to work 40 h or more per week (53.6 %), 8 h or more per weekday (57.8 %), and later than 10 PM on weeknights (47.4 %). Interactions of these variables are also included in the selection model.

Sample Selection Problems, Table 1 Descriptive statistics and the selection model of employment

Variable	Min, max	Mean (std. dev.)	Probit model of employment	
			Coeff. (std. err.)	Marginal effect
Response variables (wave 2)				
Employed in a formal job	0, 1	50.9 %	–	–
Weekly earnings if employed	0.5, 800.0	147.0 (103.7)	–	–
Regressors (wave 1)				
Cohort	12, 16	14.0 (1.4)	0.275 (0.039)***	0.109
White	0, 1	52.1 %	0.350 (0.053)***	0.138
Highest grade completed	0, 12	7.6 (1.6)	0.168 (0.030)***	0.066
Live with both biological parents	0, 1	50.5 %	–0.090 (0.052) ⁺	–0.036
Household size	2, 16	4.5 (1.5)	–0.030 (0.016) ⁺	–0.012
Live in central city	0, 1	31.1 %	–0.036 (0.054)	–0.014
Local unemployment rate	1.5, 17.5	5.5 (3.0)	–0.025 (0.009)**	–0.010
Percentage of weeks employed	0.0, 100.0	12.4 (27.8)	0.012 (0.002)***	0.005
No. of different jobs	0, 7	0.4 (0.8)	0.293 (0.063)***	0.116
No. of different delinquent behaviors	0, 7	1.4 (1.7)	0.075 (0.016)***	0.030
No. of arrests	0, 33	0.3 (1.4)	–0.039 (0.018)*	–0.016
Instrumental variables (wave 2)				
(1) 40+ hours per week allowed	0, 1	53.6 %	0.223 (0.138)	0.088
(2) 8+ hours per weekday allowed	0, 1	57.8 %	0.405 (0.135)**	0.160
(3) work past 10 PM allowed	0, 1	47.4 %	1.144 (0.411)**	0.428
(1) × (2)	0, 1	49.7 %	–0.697 (0.206)***	–0.271
(1) × (3)	0, 1	41.7 %	–1.117 (0.485)*	–0.423
(2) × (3)	0, 1	45.7 %	–1.089 (0.444)*	–0.413
(1) × (2) × (3)	0, 1	40.6 %	1.378 (0.521)**	0.495

Note: $N = 3,895$. Estimates are unweighted. Means of binary variables are shown as percentages. Pseudo (McFadden's) R-square for the probit regression model is 0.336. The marginal effects are all evaluated at the sample means.
⁺ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ (two-tailed tests)

Sample Selection Problems,
Fig. 1 Distribution of employment earnings



Sample Selection Problems, Table 2 Comparative estimates of the substantive model of weekly earnings

Regressor	Uncorrected models		Selection-corrected models	
	(1) Least squares regression coeff. (std. err.)	(2) Truncated regression coeff. (std. err.)	(3) Heckman's two-step coeff. (std. err.)	(4) FIML coeff. (std. err.)
Cohort	12.4 (3.3)***	27.0 (6.7)***	17.8 (4.2)***	37.7 (3.5)***
White	-10.6 (5.0)*	-20.4 (9.9)*	-6.1 (5.5)	13.0 (5.5)*
Highest grade completed	6.3 (2.8)*	12.6 (5.4)*	8.4 (3.0)**	17.0 (3.0)***
Live with both biological parents	-7.8 (4.8)	-15.0 (9.4)	-8.6 (4.8) ⁺	-12.3 (5.2)*
Household size	3.2 (1.6)*	5.9 (3.1) ⁺	2.9 (1.6) ⁺	0.9 (1.7)
Live in central city	-9.1 (5.2) ⁺	-18.0 (10.4) ⁺	-9.7 (5.3) ⁺	-9.4 (5.6) ⁺
Local unemployment rate	1.2 (0.8)	2.1 (1.6)	0.7 (0.9)	-1.4 (0.9)
Percentage of weeks employed	-0.6 (0.1)***	-1.2 (0.2)***	-0.5 (0.1)***	-0.1 (0.1)
No. of different jobs	17.5 (3.1)***	30.3 (5.7)***	19.1 (3.3)***	24.0 (3.8)***
No. of different delinquent behaviors	2.7 (1.3)*	5.1 (2.6)*	3.6 (1.4)*	6.8 (1.5)***
No. of arrests	-1.3 (1.4)	-2.5 (2.8)	-1.7 (1.5)	-2.8 (1.7) ⁺
Lambda $\lambda(-Z\alpha)$	-	-	28.4 (14.1)*	121.8 (2.7)***
Sigma σ_e	99.3	139.1	100.6	125.2
Rho $\rho_{e,u}$	-	-	0.282	0.973

Note: $N = 1,981$. Estimates are unweighted. The dependent variable, weekly earnings, is measured from the second interview, while all regressors are measured from the first interview
⁺ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ (two-tailed tests)

The first step in the analysis is to model the probability that a respondent is uncensored, that is, has nonzero earnings. This is a probit regression model, the coefficients, standard errors, and marginal effects from which are provided in the last set of columns in Table 1. By way of a brief summary, youth who are more likely to work include those who are older, are white, have more schooling, have less local unemployment, and have more prior work experience. Youth who are more heavily involved in delinquent behavior are also more likely to work, although youth who have accumulated an arrest history are less likely to work. It is important to point out that the instrumental variables are jointly significant ($\chi^2 = 36.4$; $df = 7$; $p < 0.00001$), and six out of the seven main and interaction effects are statistically significant.

In Table 2, results from a variety of specifications of the earnings function are provided. All models are estimated from the subsample of male youth with nonzero earnings, but they differ as to

whether a selection correction is made. The first two columns provide coefficients from models in which no correction is made for sample selection – coefficients from the least squares regression model and the truncated regression model are provided. The coefficients from the two are of the same direction and almost identical significance. The difference is that the coefficients from least squares regression are about half the size of the coefficients from truncated regression. This is because the latter model explicitly accounts for the truncation in earnings at 0 (notice that the estimate of the model standard deviation, represented by σ_e , is 40 % larger in the truncated regression model) and is the typical starting point for analyses of this sort. As an aside, the truncated regression model represents the second part of the two-part model advocated by Cragg (1971; for criminological applications, see Smith and Brame 2003).

As for the question more immediately at hand, the number of arrests is unrelated to weekly

earnings in the uncorrected models. Although the coefficient is negative, it is not statistically significant. On the other hand, more extensive delinquent involvement is positively and significantly correlated with earnings. This is consistent with other research findings that high-risk youth tend to be more extensively involved in the labor market (Apel et al. 2008). A test of an interaction between delinquency and arrest is not significant.

The last two columns provide results from models which are corrected for sample selection – coefficients from Heckman’s two-step estimator and the full information maximum likelihood (FIML) estimator are provided. While the coefficient for arrest is not significant in the two-step model, it is marginally significant in the FIML model. The latter model suggests that each arrest reduces earnings by \$2.80 per week. Considering that the mean wage at the second interview is \$5.72 per hour, each arrest as of the first interview lowers earnings by the equivalent of about one-half hour’s worth of work per week between the first and second interviews.

Notice that the selection correction entails inclusion of a new regressor in the earnings model – $\lambda(-\mathbf{Z}\alpha)$, or lambda, the inverse Mills ratio. The coefficient that conforms to this regressor is $\sigma_{e,u}$, or the covariance of the residuals in the models at the first and second steps. The models also provide an estimate of the correlation between the residuals at the first and second steps – $\rho_{e,u}$, or rho. The relationship between the estimates of these quantities can be expressed as follows:

$$\hat{\rho}_{e,u} = \frac{\hat{\sigma}_{e,u}}{\hat{\sigma}_e} = \frac{\hat{\beta}_{\lambda(-Z\alpha)}}{\hat{\sigma}_e}$$

Importantly, the coefficient on lambda is meaningful. Since it is positive and statistically significant, it indicates that there is positive selection, meaning that the unobservables which make a young male more prone to work (e.g., motivation, work orientation) also tend to yield him higher mean earnings. So, selection into the non-zero earnings subsample produces higher expected earnings compared to someone drawn at random from the population, with the same set

of characteristics, who chooses not to work. Therefore, models which do not perform a selection correction will produce estimates which are biased and inconsistent. Upon first glance, then, the selection-corrected models are preferred to the truncated regression model.

By the same token, there are circumstances in which the cure for the sample selection problem is oftentimes worse than the disease. Recall that high collinearity between the regressors and the selection term, lambda, can produce inefficiency, which can undermine the consistency advantage of the two-step and FIML estimators. Unfortunately, high collinearity cannot be ruled out in the present analysis, even with exclusion restrictions for the child labor law provisions. A regression of the inverse Mills ratio on all of the covariates yields a squared multiple correlation (R-square) of 0.875. And this is only marginally better than an inverse Mills ratio which omits the instrumental variables, producing an R-square of 0.907. This suggests that the selection-corrected models might be plagued by high collinearity between the regressors and the inverse Mills ratio and that the instrumental variables which are chosen might not be sufficiently strongly related to employment to justify their use as instruments. On second glance, therefore, the truncated regression model might be preferred to the selection-corrected models after all.

Conclusion

The sample selection problem which led to the formulation of Heckman’s two-step estimator (Heckman 1976, 1979) is a very general problem. Casual consideration leaves one with the impression that sample selection is pervasive in social scientific analysis. Namely, any restriction that leads to the elimination of a potentially nonrandom subset of a sample that is intended to be representative of a larger population can give rise to the sample selection problem. Viewed in this light, Heckman’s two-step estimator, and the growing class of models to which it belongs, is more than a peculiar method which is strictly of interest to economists. It is a tool that should be readily available in any researcher’s toolbox.

To be sure, Heckman's two-step solution to the sample selection problem is not free of its detractors. It has well-known limitations, some of which have been described in this chapter. In fact, Heckman (1976, 1979) himself was cautious about his proposed model, recommending that it be used in the course of preliminary investigation rather than as the sole model. Any final determination of which set of models provides the most credible parameter estimates requires careful consideration of underlying assumptions and the consequences of violating those assumptions.

Related Entries

- ▶ [Econometrics of Crime](#)
- ▶ [Parametric Sample Selection Models](#)
- ▶ [Sample Selection Models](#)

Recommended Reading and References

- Apel R, Bushway SD, Paternoster R, Brame R, Sweeten G (2008) Using state child labor laws to identify the causal effect of youth employment on deviant behavior and academic achievement. *J Quant Criminol* 24:337–362
- Berk RA (1983) An introduction to sample selection bias in sociological data. *Am Sociol Rev* 48:386–398
- Bushway S, Johnson BD, Slocum LA (2007) Is the magic still there? The use of the Heckman two-step correction for selection bias in criminology. *J Quant Criminol* 23:151–178
- Cragg JG (1971) Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica* 39:829–844
- Heckman JJ (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Ann Econ Soc Meas* 5:475–492
- Heckman JJ (1979) Sample selection bias as a specification error. *Econometrica* 47:153–161
- Smith DA, Brame R (2003) Tobit models in social science research: some limitations and a more general alternative. *Sociol Method Res* 31:364–388
- Stolzenberg RM, Relles DA (1990) Theory testing in a world of constrained research design: the significance of Heckman's censored sampling bias correction for nonexperimental research. *Sociol Method Res* 18:395–415
- Stolzenberg RM, Relles DA (1997) Tools for intuition about sample selection bias and its correction. *Am Sociol Rev* 62:494–507
- Winship C, Mare RD (1992) Models for sample selection bias. *Annu Rev Sociol* 18:327–350

Sanctioning

- ▶ [Sentencing Research](#)

Scared Straight Programs

Anthony Petrosino¹, Carolyn Petrosino², Meghan Hollis-Peel³ and Julia Lavenberg⁴

¹Learning Innovations, WestEd,

Woburn, MA, USA

²Bridgewater State University, Bridgewater, MA, USA

³Northeastern University and University of Massachusetts, Boston, MA, USA

⁴University of Pennsylvania Medical Center, Philadelphia, Pennsylvania, USA

Synonyms

[Juvenile awareness programs](#); [Prison aversion programs](#)

Overview

“Scared Straight” and other programs involve organized visits to prison by juvenile delinquents or children at risk for criminal behavior. Programs are designed to deter participants from future offending through firsthand observation of prison life and interaction with adult inmates.

Drawing on the findings from a recently updated systematic review, this entry describes these programs and examines the research evidence on the effects of these programs on delinquency. This entry shows that not only do these programs fail to deter crime but they actually increase it. Government officials permitting this program need to adopt rigorous evaluation to ensure that they are not causing more harm to the very citizens they pledge to protect.

Background

In the 1970s, inmates serving life sentences at a New Jersey (USA) prison began a program to “scare” or deter at-risk or delinquent children from a future life of crime. The program, known as “Scared Straight,” featured as its main component an aggressive presentation by inmates to juveniles visiting the prison facility. The presentation depicted life in adult prisons, and often included exaggerated stories of rape and murder (Finckenauer 1982). A television documentary on the program aired in 1979 provided evidence that 16 of the 17 delinquents remained law abiding for 3 months after attending “Scared Straight” – a 94 % success rate (Finckenauer 1982). Other data provided in the film indicated success rates that varied between 80 % and 90 % (Finckenauer 1982). The program received considerable and favorable media attention and was soon replicated in over 30 jurisdictions nationwide, resulting in special Congressional hearings on the program and the film by the United States House Committee on Education and Labor, Subcommittee on Human Resources (Petrosino et al. 2000).

Programs such as “Scared Straight” are based on deterrence theory. Advocates and developers of these programs believe that it is possible to use these realistic depictions of life in prison and/or presentations by inmates to deter juvenile offenders (or children at risk for becoming delinquent) from further (or initial) involvement with crime. Although the harsh and sometimes vulgar presentation in the earlier New Jersey version is the most famous, inmate presentations are now sometimes designed to be more educational than confrontational but with a similar crime prevention goal (Finckenauer and Gavin 1999). Programs that feature interactive discussions between youth and inmates as speakers who describe their life experiences and the current reality of prison life have a rather long history in the United States (Michigan D.O.C. 1967). It is not surprising such programs are popular: They fit with common notions of how to prevent or reduce crime (by “getting tough”); they are very inexpensive (a Maryland program was estimated

to cost less than \$1 US per participant); and they provide one way for incarcerated offenders to contribute productively to society by preventing youngsters from following down the same path (Finckenauer 1982).

A randomized controlled trial of the New Jersey program published in 1982, however, reported no effect on the criminal behavior of participants in comparison with a no-treatment control group (Finckenauer 1982). In fact, Finckenauer reported that participants in the experimental program were more likely to be arrested following the program. Other randomized trials reported in the United States also questioned the effectiveness of “Scared Straight”-type programs in reducing subsequent criminality (Greater Egypt Regional Planning and Development Commission 1979; Lewis 1983). Consistent with these findings, reviewers of research on the effects of crime prevention programs have not found deterrence-oriented programs like “Scared Straight” to be effective (Lipsey 1992).

Despite the convergence of evidence from studies and reviews, “Scared Straight”-type programs remain popular and continue to be used in the United States (Finckenauer and Gavin 1999). For example, a program in Carson City, Nevada (USA) brings juvenile delinquents on a tour of an adult Nevada State Prison (Scripps 1999). One youngster claimed that the part of the tour that made the most impact on him was “All the inmates calling us for sex and fighting for our belongings” (Scripps 1999). The United Community Action Network has its own program called “Wisetalk” in which at-risk youth are locked in a jail cell for over an hour with four to five parolees. They claim that only 10 of 300 youngsters exposed to this intervention have been rearrested (United Community Action Network 2001). In 2001, a group of guards – apparently without the knowledge of administrators – strip-searched Washington, DC students during their tours of a local jail under the guise of that they were using “a sound strategy to turn around the lives of wayward kids” – claiming the prior success of “Scared Straight” (Blum and Woodlee 2001).

“Scared Straight” and other “kids visit prison” programs are not unique to the United States. For example, it is called the “day in prison” or “day in gaol” in Australia (O’Malley et al. 1993), “day visits” in the United Kingdom (Lloyd 1995), and the “Ullersmo Project” in Norway (Storvoll and Hovland 1998). Hall (1999) reports positively on a program in Germany designed to scared straight young offenders with ties to Neo-Nazi and other organized hate groups. The program has been also tried in Canada (O’Malley et al. 1993). A recent television program (*Banged Up*) aired in the United Kingdom depicting their version of the program (Blunkett 2008; Wilson and Groombridge 2010).

In 1999, “Scared Straight: 20 Years Later” was shown on United States television and claimed similar results to the 1979 film (Muhammed 1999). In this version, the film reports that 10 of the 12 juveniles attending the program have remained offense free in the 3 months follow-up (Muhammed 1999). As in the 1979 television program, no data on a control or comparison group of young people were presented. Positive reports and descriptions of Scared Straight-type programs have also been reported elsewhere (e.g., in Germany [Hall 1999], and in Florida [Rasmussen 1996]), although it is sometimes embedded as one component in a multicomponent juvenile intervention package (Trusty 1995; Rasmussen 1996).

In 2000, Petrosino and his colleagues reported on a preliminary systematic review of nine randomized field trials, drawing on the raw percentage differences in each study. They found that programs such as “Scared Straight” generally increased crime between 1 % and 28 % in the experimental group when compared to a no-treatment control group. In 2002, a formal Campbell review was published (simultaneously with the Cochrane Collaboration) – updating the 2000 work and utilizing more sophisticated meta-analytic techniques. They reported similarly negative findings for Scared Straight and juvenile awareness programs. This entry provides an update of these earlier reports, extending searches to cover literature published through 2011.

Despite the results of this review and updates, Scared Straight and similar programs continue to be promoted as a crime prevention strategy. For example, Illinois’ then-Governor Rod Blagojevich signed a bill into law in 2003 that mandated the Chicago Public School system to set up a program called “Choices” (United Press International 2003). The program would identify students at risk for committing future crime and set up a program to give them “tours of state prison” to discourage any future criminal conduct (United Press International 2003). More recently, the Arts and Entertainment (A&E) station has been running a weekly series entitled, “Beyond Scared Straight.” Created by the producer of the original Scared Straight program (Arnold Shapiro), the program is now the highest rated in A&E’s history. The success of the television show has renewed interest in Scared Straight and similar programs as a crime prevention strategy (e.g., Dehnart 2011) but has also resulted in criticism that it ignores a long history of scientific evidence (e.g., Robinson and Slowikowski 2011). The Puerto Rico Department of Corrections and Rehabilitation also launched the “Loving Freedom” program, an intervention designed to imitate Scared Straight (Mullen 2012).

The question about whether Scared Straight and similar programs has a crime deterrent effect is best answered by examining the existing scientific evidence. Of course, prior research is no guarantee that interventions will work (or not work) in a future setting, but a reader might ask herself the following question upon reading the results of the aforementioned systematic review: Would I want a doctor to prescribe a treatment for my child that has the same track record of research results?

Methods

The findings reported here are based on the most recent update of the Campbell and Cochrane Collaboration reviews on the subject. In these reviews, studies were only included if they used random (or seemingly – quasi-random) procedures to assign participants to treatment and

control groups. The control must not have received another prevention program. Furthermore, only studies that involved juveniles (i.e., children 18 years of age or under) or overlapping samples of children and young adults (i.e., ages 13–21) were included. Studies could have included delinquents, pre-delinquents, or non-delinquent youths. Studies also were required to include a visit by program participants to a prison facility as a part of the intervention. Most of the included studies involved inmate presentations (Finckenauer 1982; Cook 1992), and they sometimes featured orientation sessions and prison tours. Finally, studies had to include at least one outcome of subsequent offending behavior. This outcome was measured by indices such as arrests, convictions, contacts with police, and self-reported offenses. Further information on the methodology and search methods is available in the original review.

Narrative Findings

Collectively, the nine included studies from the systematic review were conducted in eight different states of the United States, with Michigan the site for two studies (Yarborough 1979; Michigan D.O.C. 1967). No set of researchers conducted more than one experiment. The studies span the years 1967–1992. The first five studies located were unpublished and were disseminated in government documents or dissertations; the remaining four were found in academic journal or book publications. The average age of the juvenile participants in each study ranged from 15 to 17. Only the New Jersey study included girls (Finckenauer 1982). Racial composition across the nine experiments was diverse, ranging from 36 % to 84 % white. Nearly 1,000 (946) juveniles or young adults participated in the nine experimental studies. A narrative description of the included studies follows.

In an internal, unpublished government document, the Michigan Department of Corrections (1967) reported a trial testing a program that involved taking adjudicated juvenile boys on a tour of a state reformatory. Unfortunately, the report is remarkably brief. Sixty juvenile delinquent boys were randomly assigned to attend two

tours of a state reformatory or to a no-treatment control group. Tours included 15 juveniles at a time. No other part of the program is described. Recidivism was measured as a petition in juvenile court for either a new offense or a violation of existing probation order. The Michigan department of Corrections found that 43 % of the experimental group re-offended, compared to only 17 % of the control group. This large negative result curiously receives little attention in the original document.

Another program at the Menard Correctional Facility (GERP&DC 1979) started in 1978 and is described as a frank and realistic portrayal of adult prison life. The researchers randomly assigned 161 youths aged 13–18 to attend the program or a no-treatment control. The participants were a mix of delinquents or children at risk of becoming delinquent. Participants were compared on their subsequent contact with police, on two personality inventories (Piers-Berne and Jesness) and surveys of parents, teachers, inmates, and young people. The outcomes are also negative in direction but not statistically significant, with 17 % of the experimental participants being recontacted by police in contrast to 12 % of the controls (GERP&DC 1979). The authors concluded that “based on all available findings one would be ill advised to recommend continuation or expansion of the juvenile prison tours. All empirical findings indicate little positive outcome, indeed, they may actually indicate negative effects” (p. 19). Researchers report no effect for the program on two attitude tests (Jesness Inventory, Piers Harris Self-Concept Scale). In contrast, interview and mail surveys of participants and their parents and teachers indicated unanimous support for the program (p. 12). Researchers also note how positive and enthusiastic inmates were about their efforts.

In the Juvenile Offenders Learn Truth (JOLT) program (Yarborough 1979), juvenile delinquents in contact with one of four Michigan county courts participated. Each juvenile spent 5 total hours in the facility. Half of this time was spent in a confrontational “rap” session. This followed a tour of the facility, during which participants were escorted to a cell

and exposed to interaction with inmates (e.g., taunting). In the evaluation, 227 youngsters were randomly assigned to JOLT or to a no-treatment control. Participants were compared on a variety of crime outcomes collected from participating courts at 3- and 6-month follow-ups. This second Michigan study reported very little difference between the intervention and control group (Yarborough 1979). The average offense rate for program participants, however, was 69 % compared to 47 % for the control group. Yarborough (p. 14) concluded that, “. . .the inescapable conclusion was that youngsters who participated in the program, undergoing the JOLT experience, did no better than their control counterparts.”

The Face-to-Face program (Vreeland 1981) included a 13-h orientation session in which the juvenile lived as an inmate followed by counseling. Participants were 15–17 years of age and on probation from Dallas County Juvenile Court; most averaged two to three offenses before the study. A total of 160 boys were randomly assigned to four conditions: prison orientation and counseling, orientation only, counseling only, or a no-treatment control group. Vreeland examined official court records and self-reported delinquency at 6 months. This evaluation also reported little effect for the intervention (Vreeland 1981). Vreeland reported that the control participants outperformed the three treatment groups on official delinquency (28 % delinquent versus 39 % for the prison orientation plus counseling, 36 % for the prison only, and 39 % for the counseling only). This more robust measure contradicts data from the self-report measures used, which suggest that all three treatment groups did better than the no-treatment controls. None of these findings reached a level of statistical significance. Viewing all the data, Vreeland concluded that there was no evidence that Face-to-Face was an effective delinquency prevention program. He finds no effect for “Face-to-Face” on several attitudinal measures, including the “Attitudes Toward Obeying Law Scale.”

The New Jersey Lifers’ Program (Finckenauer 1982) began in 1975 and stressed confrontation

with groups of juveniles ages 11–18 who participated in a rap session. Finckenauer randomly assigned 82 juveniles, some of whom were not delinquents, to the program or to a no-treatment control group. He then followed them for 6 months in the community, using official court records to assess their behavior. Finckenauer reported that 41 % of the children and young people who attended the “Scared Straight” program in New Jersey committed new offenses, while only 11 % of the controls did, a difference that was statistically significant (Finckenauer 1982). He also reported that the program participants committed more serious offenses and that the program had no impact on nine attitude measures with the exception of a measure called “attitudes toward crime.” On this measure, experimental participants did much worse than controls. We deal with Finckenauer’s own concerns about randomization integrity in a sensitivity analysis, which is reported later.

The California SQUIRES Program (Lewis 1983) is supposedly the oldest such program in the United States, beginning in 1964. The SQUIRES program included male juvenile delinquents from two California counties between the ages of 14 and 18, most with multiple prior arrests. The intervention included confrontational rap sessions with rough language, guided tours of prison with personal interaction with prisoners, and a review of pictures depicting prison violence. The intervention took place 1 day per week over 3 weeks. The rap session was 3 h long, and normally included 20 youngsters at a time. In the study, 108 participants were randomly assigned to treatment or to a no-treatment control group. Lewis compared participants on seven crime outcomes at 12 months. Lewis reported that 81 % of the program participants were arrested compared to 67 % of the controls. He also found that the program did worse with seriously delinquent youths, leading him to conclude that such children and young people could not be “turned around by short-term programs such as SQUIRES. . .a pattern for higher risk youth suggested that the SQUIRES program may have been detrimental” (p. 222). The only deterrent effect for the program was the average length of

time it took to be rearrested: 4.1 months for experimental participants and 3.3 months for controls. Data were reported on eight attitudinal measures, and Lewis reported that the program favored the experimental group on all of them, again underscoring the difficulty of achieving behavioral change even when positively affecting the attitudes of juvenile delinquents.

The Kansas Juvenile Education Program (Locke et al. 1986) was designed to educate children about the law and the consequences of violating it. The program also tried to roughly match juveniles with inmates based on personality types. Fifty-two juvenile delinquents (ages 14–19) from three Kansas counties were randomly assigned while on probation to KEP or a no-treatment control. The investigators examined official (from police and court sources) and self-report crime outcomes at 6 months. Locke and his colleagues reported little effect of the Juvenile Education Program. Both groups improved from pretest to posttest, but the investigators concluded that there were no differences between experimental and control groups on any of the crime outcomes measured. Investigators also reported no effect for the program on the Jesness and Cerkovich attitude tests.

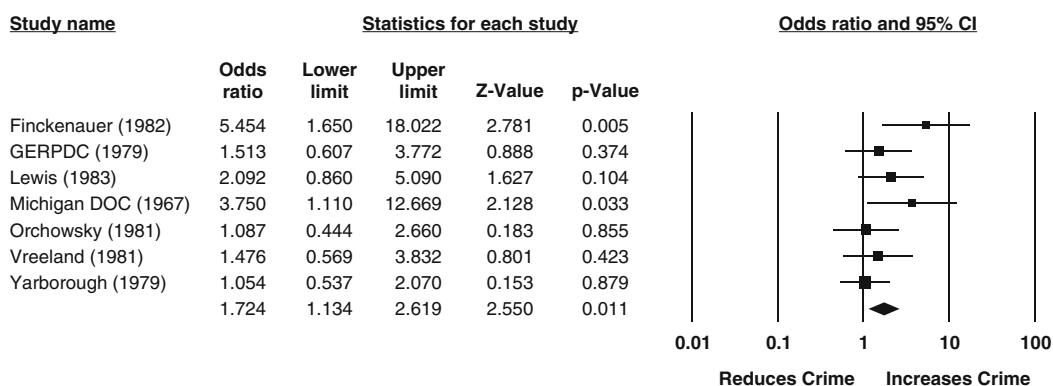
Project Aware (Cook and Spurrison 1992) was a nonconfrontational, educational program comprising one 5 h session run by prisoners. The intervention was delivered to juveniles in groups numbering from 6 to 30. In the study, 176 juveniles (ages 12–16) under the jurisdiction of the county youth court were randomly assigned to the program or to a no-treatment control. The experimental and control groups were compared on a variety of crime outcomes retrieved from court records at 12 and 24 months. Little difference was found between experimental and control participants in the study. For example, the mean offending rate for controls at 12 months was 1.25 for control cases versus 1.32 for Project Aware participants. Both groups improved from 12 to 24 months, but the control mean offending rate was still lower than the experimental group. The investigators concluded that, “attending the treatment program had no significant effect on the frequency or

severity of subsequent offenses” (p. 97). The investigators also reported on two educational measures: school attendance and dropout. Curiously, they report an effect for the program on school dropout data, but not that “. . . it is not clear how the program succeeded in reducing dropout rates. . .” (p. 97).

The only positive findings, though not statistically significant, were reported in Virginia (Orchowsky and Taylor 1981). The Insiders Program (Orchowsky and Taylor 1981) was described as an inmate-run, confrontational intervention with verbal intimidation and graphic descriptions of adult prison life. Juveniles were locked in a cell 15 at a time and told about the daily routine by a guard. They then participated in a 2-h confrontational rap session with inmates. Juvenile delinquents from three court service units in Virginia participated in the study. The investigators randomly assigned 80 juveniles ages 13–20 with two or more prior adjudications for delinquency to the Insiders program or a no-treatment control group. Orchowsky and Taylor report on a variety of crime outcome measures at 6-, 9-, and 12-month intervals. Although the difference at 6 months was not statistically significant (39 % of controls had new court intakes versus 41 % of experimental participants), they favor the experimental participants at 9 and 12 months. The investigators noted, however, that the attrition rates in their experiment were dramatic. At 9 months, 42 % of the original sample dropped out, and at 12 months, 55 % dropped out. The investigators conducted analyses that seemed to indicate that the constituted groups were still comparable on selected factors.

Most of the studies dealt with delinquent youths already in contact with the juvenile justice system. All of the experiments were simple two-group experiments except Vreeland’s evaluation of the Texas Face-to-Face program (Vreeland 1981). Only one study used quasi-random alternation techniques to assign participants (Cook 1992); the remaining studies claimed to use randomization although not all were explicit about how such assignment was conducted. Only the Texas study (Vreeland 1981) included data from

Effects of Scared Straight and other similar programs: Meta-analysis of first effect crime outcomes (Random Effects Analysis)



Scared Straight Programs, Fig. 1 Random effects model results (Petrosino et al. 2004)

self-report measures. In two studies (Cook 1992; Locke et al. 1986), no post-intervention offending rates were reported. Also, the follow-up periods were diverse and included measurements at 3, 6, 9, 12, and 24 months.

The results of the meta-analysis presented in the systematic review are further evidence of the harmfulness of these programs. The review reported the crime outcomes for official measures at the first-effect or first follow-up interval (and usually the only) period reported. Each analysis focused on proportion data (i.e., the proportion of each group re-offending), as the outcomes reporting means or averages is sparse and often does not include the standard deviations. Thus, because the data rely on dichotomous outcomes, analyses used the odds ratios (OR) as the measure of program effect, contrasting the odds of crime in the treatment program relative to the control.

The analysis of the data in comparison Fig. 1 from the seven studies reporting reoffending rates shows that intervention increases the crime or delinquency outcomes at the first follow-up period. The mean odds ratio across studies assuming a random effects model indicates an overall harmful effect of these programs (mean odds ratio = 1.72, 95 % confidence interval of 1.13–2.62). Thus, the intervention increases the odds of offending (Fig. 1).

These randomized trials, conducted over a 25-year period in eight different jurisdictions, provide evidence that “Scared Straight” and other

“juvenile awareness” programs are not effective as a stand-alone crime prevention strategy. More importantly, they provide empirical evidence that these programs likely increase the odds that children exposed to them will commit offenses in future. Despite the variability in the type of intervention used, ranging from harsh, confrontational interactions to tours of the facility converge on the same result: an increase in criminality in the experimental group when compared to a no-treatment control. Doing nothing would have been better than exposing juveniles to the program.

Converging Evidence

Other reviews also examined the efficacy of “Scared Straight” and similar programs. A meta-analysis of juvenile prevention and treatment programs by Lipsey (1992) found a small negative effect for 11 “shock incarceration” and “Scared Straight” programs, with the experimental groups having a 7 % higher recidivism rate than controls relative to a 50 % recidivism baseline. Gendreau and his colleagues (1997) also reported a meta-analysis of “get tough” or “get smart” sanctions. These included interventions designed to deter future crime like “Scared Straight” as well as interventions designed to punish or control offenders at less cost such as intensive supervision while on probation or parole. The reviewers computed correlations of program participation and recidivism outcomes. Examining 15 experimental or

quasi-experimental evaluations of Scared Straight-type programs, they also found a small negative (harmful) effect (average correlation of 0.07). Simply put, participating in the program was *associated with an increase in crime*.

Possible Controversies in the Literature

This evidence places the onus on every jurisdiction to show how their current or proposed program is different than the ones studied here. Given that, they should then put in place rigorous evaluation to ensure that no harm is caused by the intervention. Some literature indicates the program can have a positive effect on the inmate providers and that argument is sometimes used to legitimize use of the program. These arguments are undoubtedly used under the assumption that the program does no harm. In light of these findings, assertions that “Scared Straight” and similar programs ought to be used because it achieves other things raises ethical questions about potentially harming children (and others in the community who may be victimized) in order to accomplish other important, but latent, goals.

Petrosino and colleagues have received communications from different prison facilities that are using a juvenile awareness program. One argument these programs make to sustain using such programs is that the research reported here does not apply to their particular program. One recommendation here is that correctional research units, either at the facility or at a regional or national government level, collaborate with program staff to conduct a rigorous evaluation. If such units do not exist or cannot conduct their own study, they should collaborate with a local university, college, or research firm that could undertake this work to ensure that the program is working as planned and not unintentionally causing more harm than good. The findings, however, across a diversity of programs, jurisdictions, and samples (in conjunction with the converging evidence from other reviews) suggests that the fundamental concept behind Scared Straight and other juvenile awareness programs may be flawed.

Open Questions

One question that continues to arise about these findings is why “Scared Straight” and similar programs seem to lead to more crime rather than less in its participants. What is the critical mechanism? Understanding why something works or fails is of great interest to evaluators, program designers, and criminological theorists. Holley and Brewster (1996), evaluators for the Oklahoma “Speak Outs” program, wondered about the criminogenic effect of these programs when they asked:

If one argued that a two hour visit cannot perform the miracle of deterring socially unacceptable behavior . . . , it can also be argued that it was extremely simplistic to assert that a two hour visit can perform the miracle of *causing* socially unacceptable behavior. (p. 130)

Although there were many good post hoc theories about why these programs had negative effects, the evaluations were not structured to provide the kind of mediating variables or “causal models” necessary for an empirical response to this question in a systematic review.

Another key question concerns why these programs continue to be used. The research evidence to date indicates that these programs simply do not work. Despite this evidence, these programs continue to be used. This concern is particularly problematic given the recent paucity of high-quality research studies evaluating these programs. Why do policymakers continue to implement programs that are found to be harmful?

Arnold Shapiro (cited in Dehnart 2011) criticized the studies reviewed here because none of them were reported after 1992. “Scared Straight” has evolved and is now a very different program, and two decades have passed since that last study was published. This further reinforces the need for jurisdictions using this strategy to conduct rigorous evaluation, but it is difficult to obtain funding from agencies because Scared Straight is viewed as a failed strategy for youth.

Conclusions

The research on Scared Straight–type programs cannot predict with certainty that every such program will fail or – worse yet – lead to harmful effects on juvenile participants. But, the prior evidence indicates that there is a greater probability than not that it will be harmful. Would you permit a doctor to use a medical treatment on your child with a similar track record of results?

Despite the gloomy findings reported here and elsewhere, “Scared Straight” and its derivatives continue in use, although a randomized trial has not been reported since 1992. As Finckenauer and Gavin (Finckenauer 1999) noted, when the negative results from the California SQUIRES study came out, the response was to end the evaluation – not the program. Today, the SQUIRES program continues, evaluated by the testimonials of prisoners and participants alike. Despite evidence, belief in the program’s efficacy continue. Middleton and his colleagues report on the extension of this strategy in one UK town to scare ordinary schoolchildren by using former correctional officers to set up a prison-type atmosphere in the public school system (Middleton et al. 2001). In 1982, Finckenauer called this the “Panacea Phenomenon,” describing how policymakers, practitioners, media reporters, and others sometimes latch onto quick, short-term, and inexpensive cures to solve difficult social problems (Finckenauer 1982). Others claim that the program by itself is of little value but could be instrumental if embedded in an overall multicomponent package of interventions delivered to youths. More recently, the success of the A&E program, Beyond Scared Straight, has increased enthusiasm for this program as a crime prevention strategy.

It may be true that Scared Straight and like programs do not work because they only convey a threat that juveniles do not think will be carried out. What about the evidence for deterrence when it is not an inmate providing a third-party threat but the juvenile system officially processes the youth? There have been a wide range of randomized trials that test for the effects of official processing in juvenile courts with some other

intervention (such as diverting the kid from such processing). Is there evidence that the delivery of a threat – official system processing – deters future criminal behavior? In 2010, Petrosino, et al. examined 29 randomized trials that evaluated the effects of some diversionary alternative (services or outright release) and compared it to official processing or progression deeper into the juvenile justice system. That review, published by the Campbell Collaboration, also indicated that formal system processing or progression had no crime deterrent effect, and in some instances increased crime in contrast to diversionary alternatives. In addition, formal processing is a more expensive approach than most diversionary programs, and coupled with the crime reduction effect, could result in some savings for jurisdictions (Petrosino et al. 2010). This review indicates that the delivery of a threat (official processing) did not deter future juvenile offending, compared to doing nothing, and actually reported worse outcomes than if the youth was assigned to a diversionary program with services.

Related Entries

- ▶ [Deterrence](#)
- ▶ [Examining the Effectiveness of Correctional Interventions](#)
- ▶ [Juvenile Diversion](#)
- ▶ [Juvenile Justice in the Get Tough Era](#)

Recommended Reading and References

- Blum J, Woodlee Y (2001) Trying to give kids a good scare. *Washington Post*, 3 June 2001, p C01
- Blunkett D (2008) Scared straight? *Public Pol Res* 15(2):77–79
- Cook DD (1990) Effects of a non-confrontational prisoner-run juvenile delinquency deterrence program. Unpublished Masters thesis, Department of Psychology, Mississippi State University
- Cook DD, Spirison CL (1992) Effects of a prisoner-operated delinquency deterrence program: Mississippi’s project aware. *J Offender Rehab* 17:89–99
- Dehnart A (2011) Beyond scared straight’s real-life controversy. *The Daily Beast*, 23 Feb 2011.

- <http://www.thedailybeast.com/articles/2011/02/23/beyond-scared-straight-real-life-controversy.html>
- Finckenauer JO (1982) Scared straight and the panacea phenomenon. Prentice-Hall, Englewood Cliffs
- Finckenauer JO, Gavin PW (1999) Scared straight: the panacea phenomenon revisited. Waveland Press, Prospect Heights
- Gendreau P, Goggin C, Cullen F, Andrews D, Bonta J (1997) The effectiveness of “get tough” strategies on offender recidivism. Unpublished manuscript, University of New Brunswick, Center for Criminal Justice Studies, Saint John
- Greater Egypt Regional Planning & Development Commission (1979) Menard correctional center: juvenile tours impact study. Greater Egypt Regional Planning and Development Commission, Carbondale
- Hall A (1999) Jailhouse shock aims to scare youths straight. *The Scotsman*, 26 Oct 1999, p 12
- Holley PD, Brewster D (1996) A description of speak outs in Oklahoma prisons. *Okla Crim Just Res Consort J* 3(Aug):123–133
- Lewis RV (1983) Scared straight—California style: evaluation of the San Quentin SQUIRES program. *Crim Just Behav* 10(2):209–226
- Lipsey MW (1992) Juvenile delinquency treatment: a meta-analytic inquiry into the variability of effects. In: Cook TC, Cooper H, Cordray DS, Hartmann H, Hedges LV, Light RL, Louis TA, Mosteller FM (eds) *Meta-analysis for explanation*. Russell Sage, New York, pp 83–127
- Lloyd C (1995) To scare straight or educate? The British experience of day visits to prison for young people. Home Office Research Study No. 149. Home Office, London
- Locke TP, Johnson GM, Kirigin-Ram K, Atwater JD, Gerrard M (1986) An evaluation of a juvenile education program in a state penitentiary. *Eval Rev* 10:281–298
- Michigan Department of Corrections (1967) A six month follow-up of juvenile delinquents visiting the Ionia Reformatory. Research report no. 4. Michigan Department of Corrections, Lansing
- Middleton J, Reeves E, Lilford R, Howie F, Hyde C, Elbourne D, Oakley A, Gough D (2001) Collaboration with the Campbell Collaboration. *Brit Med J* 323:1252
- Muhammed L (1999) Kids and crooks revisited: some were “Scared Straight!” *USA Today*, 12 Apr 1999, p 4D
- Mullen C (2012) Bar Association: stop putting innocent kids in jail. *San Juan Daily Star* (Letter to the Editor), 14 Feb 2012, p 4
- O’Malley P, Coventry G, Walters R (1993) Victoria’s day in prison program: an evaluation and critique. *Aust N Z J Criminol* 26(2):171–183
- Orchowsky S, Taylor K (1981) The insiders juvenile crime prevention program. Virginia Department of Corrections, Richmond
- Petrosino A, Turpin-Petrosino C, Finckenauer JO (2000) Well meaning programs can have harmful effects!: Lessons from experiments in Scared Straight and other like programs. *Crime Delinq* 46:354–379
- Petrosino A, Turpin-Petrosino C, Buehler J (2004) “Scared Straight” and other juvenile awareness programs for preventing juvenile delinquency. *Campbell Syst Rev* 2004:2
- Petrosino A, Turpin-Petrosino C, Guckenburg S (2010) Formal system processing of juveniles: effects on delinquency. *Campbell Syst Rev* 2010:1
- Rasmussen DW, Yu Y (1996) An evaluation of juvenile justice innovations in Duval County, Florida. Florida State University, Tallahassee
- Robinson LO, Slowikowski J (2011) Scary – and ineffective. Traumatizing at-risk kids is not the way to lead them away from crime and drugs. *Baltimore Sun*, 31 Jan 2011. http://articles.baltimoresun.com/2011-01-31/news/bs-ed-scared-straight-20110131_1_straight-type-programs-straight-program-youths
- Scripps J (1999) Prison tour serves as a wake-up call. *The Forum*, 27 Oct 1999, p 1
- Shapiro A (1978) Scared straight [film]. Pyramid Films, Santa Monica
- Storvall AE, Hovland A (1998) The Ullersmo project: scared straight in Norway 1992–1996 (Ullersmoprojektet: “Scared Straight” I Norge 1992–1996). *Nordisk Tidsskrift for Kriminalvidenskab* 85(2):122–135
- Trusty G (1995) Sheriff’s office, juvenile court resources used for restarts. *Louisiana Youth Care Magazine* 18(5):3–5
- United Community Action Network (2001) Services: Wisetalk (scared straight). www.ucan.av.org/services.htm. Accessed 26 Apr 2001
- United Press International (2003) Scared straight, 21 Aug, p 1008233w3635
- Vreeland AD (1981) Evaluation of face-to-face: a juvenile aversion program. Unpublished doctoral dissertation, University of Texas, Dallas
- Wilson D, Groombridge N (2010) I’m making a TV programme here: reality TV’s *Banged up* and public criminology. *Howard J Criminol* 49(1):1–17
- Yarborough JC (1979) Evaluation of JOLT as a deterrence program. Michigan Department of Corrections, Lansing

Scenario Designs

M. Lyn Exum and Megan Sims
Department of Criminal Justice and
Criminology, University of North Carolina at
Charlotte, Charlotte, NC, USA

Overview

Suppose you drove by yourself one evening to meet some friends at a bar that is about 10 miles from

your house. You have been drinking throughout the evening and by the time you're ready to leave, you suspect your blood alcohol level might exceed the legal limit. Suppose you have to be at work early the next morning. You can either drive home or find some other way home, but if you leave your car, you will have to return early the next morning to pick it up. (Pogarsky 2004, p. 119)

Hypothetical scenarios like the one above are commonly used to study criminal decision making. In the typical study, participants are asked to envision themselves experiencing the situation described in the scenario and then to self-report how likely they would be to respond in an illegal manner (e.g., drive while drunk). They also rate various costs and benefits that might result from engaging in the offense. Using these data, researchers attempt to reconstruct the decision to offend by modeling participants' self-reported criminal intentions as a function of the perceived costs and benefits, along with other individual and situational factors.

This entry discusses the contributions of the hypothetical scenario method (HSM) to the study of criminal decision making. The rise of the HSM within the deterrence/rational choice literature is first presented, followed by a summary of the key findings from this literature. Finally, emerging areas of research and future challenges for HSM scholars are highlighted.

The Rise of the HSM

The deterrence and rational choice perspectives are the dominant theoretical frameworks in which to study criminal decision making. Both theories assert that human beings freely chose to engage in (or abstain from) crime after assessing the consequences of the act. Both theories also contend that differences in the qualities of these consequences – such as their certainty, swiftness, and severity – will lead to different decisional outcomes. However, the two theories diverge with respect to the *types* of consequences that influence criminal decision making. In its purest form, deterrence theory focuses on the crime-inhibiting effects of formal legal sanctions, such as arrest and incarceration (Nagin 1978).

In contrast, the rational choice perspective recognizes that informal sanctions as well as the benefits of crime operate in conjunction with formal sanctions to shape the decision to offend (Grasmick and Bursik 1990; Piliavin et al. 1986).

Although the philosophical underpinnings of deterrence and rational choice can be traced back to the eighteenth century, empirical research on these theories did not appear until the 1960s (Paternoster 2010). These initial tests typically compared objective measures of punishment (such as incarceration rates) to aggregate rates of crime, and in general, the results showed weak/mixed support for the notion of deterrence. Scholars subsequently argued that instead of relying on objective measures of punishment, test of criminal decision making should focus on individuals' *perceptions* of punishment, regardless of how accurate those perceptions may be. From a methodological standpoint, this shift from objective to perceptual measures meant that researchers would now need to survey individuals and have them self-report their involvement in crime as well as their perceptions of the legal consequences of their behavior. Initially, this vein of research relied heavily on cross-sectional surveys; however, given the inherent temporal ordering problems in cross-sectional data, many scholars turned to longitudinal studies to study criminal decision making.

Notwithstanding their methodological advantages, longitudinal studies of deterrence and rational choice theory are unable to capture participants' perceptions of consequences *at the exact moment* the decision to offend (or abstain) is made. Instead, these studies capture self-reported perceived consequences at Time 1 and then use these perceptions to predict participants' subsequent involvement in crime occurring prior to the Time 2 testing period (which may be as much as 1–2 years in the future). With such a large time interval, the opportunity for perceptions to change is ever present. Although evidence suggests there is some similarity in the perceived consequences at Time 1 and Time 2, one cannot be fully certain that these perceptions remained stable throughout the duration of the study interval – including the moment when

participants came face to face with a real-world criminal opportunity (see Piliavin et al. 1986, pp. 115–116).

In order to better measure the perceived consequences of crime at the moment the decision to offend/abstain occurs, some researchers began focusing on whether participants *believed* they would engage in a particular criminal act, as opposed to focusing on whether participants had actually committed the act (e.g., Grasmick and Bursik 1990). As participants then contemplated their penchant toward the offense, researchers could query them on the perceived consequences of offending. While this technique seemingly allows researchers to study criminal decision making at the moment it unfolds, it does not provide participants with adequate contextual information about the potential criminal opportunity – information that participants may need in order to accurately assess the consequences of the offense. As a result, participants' perceived consequences in these studies may be unreliable (Klepper and Nagin 1989).

In response, researchers turned to the hypothetical scenario method. Through the use of structured vignettes, the HSM allows researchers to present participants with an opportunity to offend – albeit an imaginary one. The vignettes, which are usually no more than a paragraph in length, contain important contextual information about the criminal opportunity (e.g., you are drunk at a bar 10 miles from your house, it's late, you need your car early the next day). Through the use of these vignettes, researchers can present participants with virtually any type of criminal/imprudent opportunity, such as sexual assault (Bouffard 2002b), physical aggression (Exum 2002), drunk driving (Piquero and Pogarsky 2002), corporate price fixing (Piquero et al. 2005), academic cheating (Tibbetts 1999), sports doping (Strelan and Boeckmann 2006), and even police misconduct (Pogarsky and Piquero 2004).

Unlike the typical cross-sectional or longitudinal studies of decision making, the measure of offending in HSM studies is not the participants' actual involvement in crime; instead, it is the participants' self-reported *intentions* to offend.

Researchers acknowledge that such self-reported intentions are not measures of real-world behavior (Bachman et al. 1992; Paternoster and Simpson 1996) nor should they be interpreted as “synonymous with actual performance” (Tibbetts and Herz 1996, p. 203). At the same time, researchers also contend that measures of criminal intentions should be strongly correlated with actual offending (Tibbetts 1999; Piquero and Tibbetts 1996) and can be interpreted as indicators of participants' predisposition or proneness to crime (Paternoster and Simpson 1996; Pogarsky 2004).

In sum, the HSM is wholly distinct in its approach to studying the decision-making processes proposed by deterrence and rational choice theories. Unlike research using objective measures of punishment, the HSM captures individuals' subjective perceptions of the consequences of crime. Unlike traditional cross-sectional studies, the HSM does not have inherent temporal ordering problems. Unlike traditional longitudinal studies, the HSM allows researchers to measure participants' perceived consequences at the moment they contemplate a criminal opportunity. Unlike studies that simply ask about participants' intentions to commit some kind of generically defined crime, the HSM provides participants with a specific narrative for the criminal opportunity that, in turn, should make participants' responses more reliable. As a result of all these features, the HSM stands as one of the most versatile and methodologically sophisticated techniques in which to study the decision to offend.

Key Findings from the HSM Literature

The Role of Perceived Punishments

Much has been written about the empirical status of deterrence theory (e.g., Nagin 1978; Paternoster 2010; Pratt et al. 2008), especially as it pertains to the certainty and severity of formal legal sanctions. In general, these reviews find that the deterrent effect of sanction certainty – though weak – is greater than the effect for sanction severity. However, these reviews generally

provide a summary of the deterrence literature as a whole and therefore cannot isolate the findings from only those studies using the HSM.

Pratt et al.'s (2008) meta-analysis is the exception. It disaggregates effect size estimates for punishment certainty and severity across studies of differing research methodologies, including the HSM. Interestingly, the findings from these HSM studies largely mirror those of other, more global reviews of the deterrence literature. The certainty effect for formal legal sanctions was found to be weak (mean effect size = -0.14) but nevertheless stronger than the effect for sanction severity (mean effect size = -0.05). The threat of nonlegal sanctions, such as the loss of respect or employment, was found to exert a deterrent effect that was identical to that of the certainty of formal sanctions (mean effect size = -0.14). (Note that the severity effect for these nonlegal sanctions was not estimated in the study.) Collectively, these meta-analytic findings suggest that punishment certainty and severity play a minor role in the decision to offend.

Far less attention has been devoted to studying the impact of the swiftness of punishment. In one of the few HSM studies to examine celerity, Nagin and Pogarsky (2001) used a factorial survey to manipulate the length of time in which the punishment for drunk driving was expected to occur. Across a variety of multivariate models and after controlling for a host of factors such as the certainty and severity of punishment, celerity did not exert a significant deterrent effect on self-reported intentions to drive drunk. In a subsequent HSM study examining police misconduct, Pogarsky and Piquero (2004) manipulated the swiftness of punishment an officer might experience from failing to charge an off-duty officer with drunk driving and from performing an unauthorized background check on a new neighbor. This time, the results for celerity were found to vary by the type of misconduct. The swiftness of punishment had a significant deterrent effect on failing to report a fellow officer for drunk driving but had no significant effect on performing an unauthorized background check. Pogarsky and Piquero speculate that these differences may be partly

attributed to differences in the perceived gravity of the offenses, suggesting severity and swiftness may interact to produce an overall deterrent effect. Be that as it may, the findings from these two HSM studies tentatively suggest that the swiftness of punishment – at least, by itself – has a minimal impact on criminal decision making.

The Role of Perceived Benefits

Many HSM studies, especially those with a rational choice orientation, have examined the influence of the perceived benefits of crime – with “benefits” defined to include some form of physical and/or psychological reward. Across a host of HSM studies examining a variety of offenses, this research generally shows that the potential benefits of crime are a critical component in the decision to offend. For example, HSM research finds that the benefits from crime exert a positive and significant effect on self-reported intentions to engage in computer software piracy (Higgins 2007), drunk driving (Bouffard 2007; Nagin and Paternoster 1994), physical assault (Carmichael and Piquero 2004; Exum 2002), sexual assault (Loewenstein et al. 1997; Nagin and Paternoster 1994), theft/shoplifting (Piquero and Tibbetts 1996; Tibbetts and Herz 1996), corporate offending (Paternoster and Simpson 1996; Piquero et al. 2005), plagiarism (Ogilvie and Stewart 2010), and test cheating (Sitren and Applegate 2007; Tibbetts 1999). Furthermore, the benefits of crime have been shown to be even more influential to the decision to offend than are formal and informal criminal sanctions (Nagin and Paternoster 1993).

This is not to say that perceived benefits are universally predictive of offending. In fact, some HSM studies find certain benefits of crime (and under certain conditions) to be unrelated to offending intentions and/or to contribute little predictive power in multivariate models of criminal decision making (e.g., Bouffard 2002b; Elis and Simpson 1995; Ogilvie and Stewart 2010; Paternoster and Simpson 1996; Piquero et al. 2005; Simpson and Piquero 2002). Nevertheless, and as a whole, the empirical evidence from the HSM literature overwhelmingly indicates that the

perceived payoffs from crime promote the decision to offend.

The Role of Consequence Salience

Bouffard (2002b) has argued that models of criminal decision making should not only consider the certainty and severity of perceived consequences but also the salience of those consequences. Conceptually, salience is distinct from certainty, severity, and swiftness in that it captures “how much weight a given factor carried in the actual decision-making process” (p. 123; see also Bouffard 2007). In other words, salience reflects how *important* the consequence is to the potential offender when contemplating the criminal opportunity. Additionally, salience is purported to be a dynamic quality that is influenced by state-level factors such as emotional arousal. As a result, individuals in emotionally charged states may identify consequences that are highly certain and highly severe, but in the “heat of the moment,” individuals simply do not view these consequences as terribly important. Similarly, under certain emotional states, the benefits of a criminal act – no matter how unlikely or how inconsequential – may nevertheless carry a lot of weight in the decision to offend.

To examine whether salience is a unique concept, Bouffard (2007) presented participants with a hypothetical drunk driving scenario and asked them to report the certainty, severity, and importance of the consequences associated with the offense. Factor analysis revealed mixed results, with the salience measures for perceived costs indeed loading on a unique factor, but the salience measures for perceived benefits loading on a factor with perceived certainty. To determine whether salience is a dynamic element of decision making, Bouffard (2002b) randomly assigned male participants to view either sexually explicit or neutral images prior to reading a sexual assault scenario. Participants then rated the certainty, severity, and salience of their perceived consequences. Here again, the findings were mixed. Sexual arousal appeared to minimize the salience of the negative consequences of sexual aggression, but it did not enhance the importance of the perceived benefits.

Arguably, the most important question regarding consequence salience is whether it influences an individual’s decision to offend. The HSM research examining this question is scarce and somewhat inconsistent. For example, Bouffard (2002a) reports a series of bivariate correlations between participants’ intentions to drive drunk and the salience of their perceived consequences of the act. These correlations were found to be in the expected direction, statistically significant, and weak to moderate in magnitude. However, when salience measures were correlated with intentions to engage in sexually coercion, the correlations were generally weaker with only a few statistically distinguishable from zero. Furthermore, in a series of multivariate analyses of the sexual coercion data, Bouffard (2002b) found the salience of perceived benefits – but not of the perceived costs – to impact self-reported intention scores. As a whole, the extant research on salience suggests that it is a dimension of participants’ perceived consequences that appears to have some level of influence on criminal decision making. More research in this area is needed, however.

The Role of Morality

Many HSM studies have examined the role of morality on criminal decision making. In this research, the perceived immorality of the act is treated as a type of internal sanction, with high immorality ratings predicted to be inversely related with offending intentions. Note that while immorality is similar in nature to the informal sanctions of guilt and shame, it is typically regarded as a distinct construct that is generally rated in terms of its severity (i.e., “how morally wrong would it be to...”), but not its certainty.

Bachman and colleagues (1992) were among the first to use the HSM to examine the influence of morality on criminal decision making. After controlling for perceptions of formal and informal sanctions, participants’ ratings of immorality exerted a significant inhibitory effect on their self-reported intentions to engage in sexual assault. This inhibitory effect of morality has been found in other HSM studies as well,

including those examining physical assault, drunk driving, corporate offending, shoplifting/theft, sports doping, and academic cheating (Bachman et al. 1992; Carmichael and Piquero 2004; Elis and Simpson 1995; Higgins 2007; Loewenstein et al. 1997; Nagin and Paternoster 1994; Paternoster and Simpson 1996; Piquero et al. 2005; Piquero and Tibbetts 1996; Simpson and Piquero 2002; Strelan and Boeckmann 2006; Tibbetts and Herz 1996).

Morality is sometimes viewed as the gatekeeper of the hedonic calculus. That is, before the perceived costs or benefits of a criminal act can be incorporated into the decision to offend, the act itself must first be a part of the person's repertoire of morally acceptable behavior. If it is not, then the individual is fully constrained by his morality and – as a result – any threats of punishment or enticements from benefits are irrelevant. If such a gatekeeper conceptualization is accurate, then the perceived consequences of crime should have their greatest effects among those with weak moral inhibitions and have their weakest effects among those with strong moral inhibitions (Bachman et al. 1992). Although only a few HSM studies have specifically sought to examine this interaction between morality and the role of perceived consequences, there is evidence to support the gatekeeper function of morality.

For example, Bachman et al. (1992) found that among participants who rated a sexual assault scenario to be more morally offensive, perceptions of formal punishments did not significantly alter the intention to offend. Yet, among participants who found the sexual assault to be less morally offensive, formal sanctions exerted a significant deterrent effect. Similar results have also been found among business executives and MBA students who read corporate crime scenarios (Paternoster and Simpson 1996). Although more research is needed in this area, the HSM studies that are available indicate that crime is a nonmarket area among individuals with strong moral fortitude. In other words, among the righteous, the decision to act is less utilitarian and more deontological.

The Role of Self-Control

Conceived as time-stable personality trait, self-control is purported to exert a *direct* effect on criminal intentions (Piquero and Tibbetts 1996). That is, because individuals with low self-control are impulsive risk takers, they should be less able to abstain from a criminal opportunity when it presents itself. However, self-control is also thought to exert an *indirect* effect on offending intentions by altering the individuals' perceptions of consequences of crime. That is, given that individuals with low self-control are present oriented, they should be less able to perceive the long-term negative consequences of their criminal actions, thereby resulting in greater criminal proclivities. There is some HSM research to support both of these claims.

For example, Nagin and Paternoster (1993) found evidence for a direct effect of self-control on offending intentions. After controlling for such factors as prior criminal behavior, the characteristics of the criminal opportunity, and the perceived costs/benefits of the act, a one standard deviation change in low self-control scores resulted in a 39 % increase in theft intentions, a 17 % increase in drunk driving intentions, and an 83 % increase in sexual assault intentions (see also Nagin and Paternoster 1994). In support of an indirect effect, Piquero and Tibbetts (1996) found that low self-control accentuated the perceived benefits and minimized the perceived costs associated with shoplifting and drunk driving – effects that, in turn, significantly influenced participants' intentions to offend (see also Carmichael and Piquero 2004; Higgins 2007).

However, most HSM studies that have included a measure of self-control have found either null effects or effects that are not stable across different models of decision making. For example, neither Tibbetts and Myers (1999) nor Sitren and Applegate (2007) found evidence that self-control levels directly influence college students' intentions to cheat on an exam. Similarly, Simpson and Piquero (2002) found no evidence to suggest that self-control levels predict corporate offending among business executives and MBA students. Furthermore, others have shown that the effects of self-control operate

inconsistently across the sex of the participant (Tibbetts and Herz 1996), the operational definition of self-control (Piquero and Bouffard 2007), the dependent variable (Loewenstein et al. 1997; Pogarsky and Piquero 2004), and the participants' anticipated mood when completing the hypothetical act (Carmichael and Piquero 2004). Thus, while there is some support for the notion that self-control levels directly and indirectly impact criminal decision making, the overarching findings from the HSM literature raise doubt as to whether these effects are, in fact, noteworthy and robust.

The Role of Altered States of Mind

Tests of deterrence and rational choice theory have generally ignored the offender's emotional, physiological, and psychopharmacological state during the criminal decision-making process. This omission is striking given the proposed theoretical relationships between altered states of mind and cognition (see Exum 2002; Loewenstein et al. 1997). Fortunately, the HSM is easily capable of incorporating the study of altered states on decision making, and a few such studies have taken advantage of this.

For example, in their study on the effects of sexual arousal on sexually aggressive decision making, Loewenstein et al. (1997) randomly assigned male participants to view photographs of nude women or of fully clothed fashion models prior to reading a hypothetical date rape scenario. Compared to the non-aroused participants, those who were aroused endorsed greater sexually aggressive intentions. However, few differences were found across arousal conditions with respect to participants' perceived consequences of sexual aggression. In a similar experiment, Bouffard (2002b) found that self-reported arousal levels were positively related to offending intentions but inconsistently related to the perceived consequences. In a subsequent and more detailed analysis, Bouffard (2011) concluded there are in fact differential deterrent effects across arousal states, with the perceived costs of crime exerting their greatest deterrent effects among those who reported lower arousal levels. Collectively, these studies suggest that arousal increases

sexually aggressive tendencies – not by altering participants' perceptions of the consequences – but by somehow moderating the underlying relationship between perceived consequences and the intention to offend.

Also using an experimental design, Exum (2002) examined the effect of alcohol intoxication and anger on participants' responses to a physical assault scenario. Participants were randomly assigned to drink either a nonalcoholic or alcoholic beverage, the latter designed to elevate blood alcohol levels to approximately 0.08 %. Half of the participants in each group were then confronted by the researcher in such a way to induce anger, whereas the remaining half was not. Participants then read a physical assault scenario and responded to a battery of rational choice questions. Results indicated that alcohol and anger interacted to increase one of two measures of aggressive intentions; however, consistent with other research (Loewenstein et al. 1997; Bouffard 2011), there was no support for the idea that the altered states of mind affected participants' perceptions of costs and benefits. At the same time, exploratory analyses revealed that the collection of consequences used to model decision making were not equally predictive of intentions across the experimental conditions. This suggests that – while the perceptions of costs and benefits were largely the same across all participants – those in the intoxication and anger conditions used these perceptions differently in their decision-making processes. These findings further suggest that altered states of mind moderate the decision to offend.

Carmichael and Piquero (2004) also examined the effect of perceived anger on physical assault. Anger was not experimentally manipulated in this study; instead, participants were asked to read a bar fight scenario and then self-report how angry they thought they would feel in that particular situation. Controlling for the perceived costs and benefits of the assault, anger scores were positively and significantly related to aggressive intentions. Furthermore, the perceived negative consequences of assault exerted significant deterrent effects among the low anger group but not in the high anger group, again suggesting

that anger moderates the decision-making process. Collectively, this small body of HSM research serves as a reminder that decision making does not occur in a vacuum and that emotional, visceral, and pharmacological forces shape the decision to offend.

Conclusions from the HSM Literature

The HSM research on criminal decision making finds that the certainty of punishment (both formal and informal) has a weak deterrent effect and the severity of punishment has an even weaker effect. Although rarely examined, the swiftness of punishment also seems to carry little weight in the decision to offend. In stark contrast, the perceived benefits from crime appear to be a key element in criminal decision making. However, some criminal acts appear to be so morally objectionable to participants that they will simply refrain from the act regardless of the benefits. Similarly, dynamic factors such as one's state of mind may alter the salience of these perceived consequences as well as moderate the rational process thought to underlie criminal decision making. Trait-level characteristics such as low self-control may not only make it more difficult for individuals to self-regulate their behavior in the presence of a criminal opportunity, but they may also change the way individuals assess the risks and benefits associated with that opportunity (but note, however, that the self-control effects may not be very robust). As a result of these various state- and trait-level factors, criminal decisions that appear to be completely "irrational" to the outside observer be seen as wholly rational from within the perspective of the actor.

Emerging Areas of Research and Challenges

Given that the HSM is a relatively new technique within the study of criminal decision making, one could consider most any line of HSM research to be an "emerging" area. With that being said, below are four areas of criminological HSM research that deserve special attention. These four issues have great potential to advance (or

challenge) much of what criminologists claim to know about the decision to offend.

Studying Deterrence Among Deterrables

Although there is reason to believe that individuals vary in their responsivity to the perceived consequences of crime (see Bachman et al. 1992; Carmichael and Piquero 2004; Paternoster and Simpson 1996), tests of deterrence theory typically ignore these individual susceptibilities and instead examine the properties of punishment across an aggregate sample of participants. Pogarsky (2002) challenged this approach, arguing that deterrence theory should be tested only among those who are truly deterrable, as opposed to those who are either so bound to conformity or so wedded to deviance that extralegal sanctions have no impact on their behavior. Using the HSM and participants' self-reported intentions to drive drunk, Pogarsky (2002) developed a method of identifying the subsample of "deterrables" within a group of study participants. He then examined the impact of formal sanctions among the deterrables and, for comparison, among the full sample.

Results from the full sample largely confirm the findings from past studies of deterrence theory – namely, that certainty has a greater effect than severity. However, in the analysis of only the deterrables, severity had a greater effect than certainty. Thus, the conventional wisdom that "certainty is more important than severity" appears to be challenged when studying only those people who are, in fact, swayed by sanction threats. Future HSM research should build on Pogarsky's method of disaggregating samples into deterrable and non-deterrable groups and continue to examine the role of formal and informal sanctions among those most responsive to punishments. The findings have important implications not only for criminological theory but for public policy.

The Positive Punishment Effect

Contrary to the notion that sanctions deter future behavior, some studies find a *positive* relationship between punishment and subsequent offending (e.g., Piquero and Pogarsky 2002).

Pogarsky and Piquero (2003) outlined two possible explanations for this positive punishment effect: selection and resetting. The selection explanation contends that offenders with stronger criminal propensities are more likely to come to the attention of the police (i.e., get punished) and also be more likely to recidivate. In contrast, the resetting explanation argues that offenders who experience a stroke of bad luck (e.g., get punished) are more apt to take elevated risks in the future (e.g., recidivating) because they believe their bad luck is not likely to return anytime soon. In other words, shortly after experiencing punishment, the offender's perceived certainty of sanctioning is "reset" to a nominal value.

To examine the possibility of resetting, Pogarsky and Piquero (2003) administered a hypothetical drunk driving scenario to a sample of undergraduates and asked them a series of questions about their perceptions of punishment and prior punishment experiences. Support for the resetting explanation was mixed and varied by the participant's level of risk for offending. That is, consistent with the notion of resetting, low-risk offenders who had previously been punished reported lower certainty estimates than those low-risk offenders who had not been previously punished. However, no evidence of resetting was found among high-risk offenders, with all high-risk offenders perceiving comparable sanctioning risks regardless of past punishment experiences. While more research is needed in this area to better understand the nature of positive punishment effects, future studies should also seek to integrate research on "deterables" with that on resetting. Although it is too early to tell, it may be the case that those who are most responsive to sanction threats (the deterables) are also the very same "low-risk" offenders who tend to reset their perceived threat levels following punishment.

Researcher-Generated Versus Self-Generated Consequences

When potential offenders encounter a criminal opportunity in the real world, they freely deduce their own set of costs and benefits to contemplate.

However, in the typical HSM study, participants are asked to contemplate a predetermined, uniform set of consequences generated by researchers. Such a reliance on researcher-generated consequences inevitably forces an artificial structure on the otherwise organic decision to offend. Bouffard, Exum, and Collins (2010) examined the impact of this artificial structure by presenting participants with a hypothetical shoplifting scenario and then randomly assigning them to either (1) assess a standard battery of researcher-generated consequences or (2) assess a set of costs and benefits that the participants themselves self-generated. By comparing and contrasting the two sets of consequence ratings, Bouffard and colleagues could begin to examine the extent to which individuals' decision-making processes were being influenced by the study's methodology.

The findings show that when participants are asked to assess the benefits of crime that were identified by the researcher, they are more likely to regard the consequences as possible outcomes (Bouffard et al., 2010). Illustratively, when participants were specifically asked to consider the "thrill" derived from the hypothetical theft, 64 % reported the act would afford some degree of fun/excitement. In contrast, when participants were asked to self-identify all the "good things" that might happen to them if they committed the hypothetical act, only 10 % indicated it would be fun/exciting. Similarly, participants were more likely to see the costs of shoplifting as potential outcomes when these costs were identified by the researcher versus when they were self-generated. A noteworthy example is the cost of deviating from one's morality. Whereas 99 % of participants who received researcher-generated consequences reported that shoplifting felt immoral, only 3 % of participants who self-generated consequences reported feelings of immorality.

Recall that past research using the HSM finds moral inhibitions to be influential considerations in the decision to offend (e.g., Bachman et al. 1992; Loewenstein et al. 1997; Paternoster and Simpson 1996). The findings from Bouffard et al.'s (2010) study suggest that the morality

effect reported in this literature may, in part, be a methodological artifact and largely attributed to the use of researcher-generated consequences. However, future research is needed to better understand how the hypothetical scenario methodology (in general) and the manner in which costs and benefits are presented (in particular) impact the findings from HSM studies.

The Accuracy of Self-Reported Intentions to Offend

Recall that the typical dependent variable in HSM studies is not participants' actual offending behavior but is instead their self-reported *intentions* to offend. According to the theory of reasoned action, such intentions can be viewed as valid estimates of real-world behavior because "...barring unforeseen events, a person will usually act in accordance with his or her intention" (Ajzen and Fishbein 1980, p. 5). Consistent with this notion, research generally finds the correlation between participants' intentions to act and their corresponding behavior to be positive and significant (Ajzen and Fishbein 1980; Armitage and Conner 2001). However, this correlation is largely based on the study of conventional behaviors such as voting, watching television, or going to church. Only a handful of studies have examined the intentions/behavior relationship for deviant behaviors, and while the results from these studies are promising on the surface, a closer examination reveals potential methodological and/or measurement problems that may weaken the interpretation of these intentions/behavior correlations (Exum et al. 2011; cf. Pogarsky 2004).

To further examine the predictive accuracy of self-reported intentions to offend, Exum et al. (2011) gave undergraduate students a copy of a fictitious newspaper article describing an ongoing music piracy operation on campus. Students were led to believe the article was real and had recently appeared in the local paper. The article described a graduate student who was reportedly emailing illegal copies of digital music files for free to anyone with a university email account. Also included in the article were the graduate student's name and email address. After reading the article, study participants were asked to self-

report their likelihood of contacting the graduate student and requesting illegal files from him. In the following weeks, the graduate student – who was a research confederate – monitored his email account for any requests from participants. In this way, participants' self-reported intentions to offend could be validated against their own requests for illegal music files.

Despite the variability in offending intentions, including a few participants who reported a 100 % chance of emailing the graduate student, no one in the study actually requested music files from the confederate. This suggests that self-reported criminal intentions (at least, as they apply to this form of music piracy) have a low false-negative rate but a high false-positive rate. Stated differently, all participants who reported little-to-no intentions of contacting the confederate abstained from the behavior, resulting in a false-negative rate of 0 %. However, all participants who reported strong-to-definitive intentions of contacting the confederate abstained as well, resulting in a false-positive rate of 100 % (Exum et al. 2012). This suggests that when presented with a hypothetical offending scenario, participants may fall prey to a false sense of bravado and self-report strong intentions to offend, only later to cower at the real-world opportunity. Given that self-reported intentions to offend are an essential element of HSM studies of deterrence/rational choice, additional research is needed to examine the accuracy of these offending intentions (see also Pogarsky 2004). The implications of this line research have the potential to challenge most everything we know about the nature of criminal decision making.

Conclusion

More than 200 years ago, Jeremy Bentham (1970/1789) described the deliberative process in which all individuals are thought to engage as they contemplate a criminal act:

Sum up all the values of all the *pleasures* on one side, and those of all the *pains* on the other. The balance, if it be on the side of pleasure, will give the *good* tendency of the act upon the whole with the respect to the interest of that *individual* person. ... (p. 40, emphasis in the original)

Two centuries later, criminologists continue to struggle with the best way to test these ideas. The biggest impediment appears to be obtaining a measure of participants' perceived "pleasures" and "pains" at the moment participants decide to engage in or abstain from crime. The hypothetical scenario method is arguably the best technique to date for doing this, thereby making it uniquely situated to study criminal decision making at the moment it unfolds. Today, much of what we now know about the decision to offend – e.g., the negligible effect of punishment, the lure of perceived benefits, and the different calculative processes across individual- and state-level factors – has been either discovered through or confirmed by HSM research.

As scholars seek to build upon our current understanding of criminal decision making, they will inevitably rely on the HSM to some degree. The technique's versatility, its ease of administration, and its ability to accommodate experimental manipulations make it an ideal platform for designing and implementing research studies. At the same time, the HSM's greatest strength – i.e., its ability to get inside the "black box" of the mind and study decision making during the contemplation of a criminal opportunity – comes with a great weakness. The HSM can only allow researchers to model the decisions that underlie criminal *intentions* and not actual criminal *behavior*. As a result, the HSM can inform our understanding of real-world criminal decision making only to the extent that individuals' self-reported intentions are predictive of their actual criminal conduct. Until the predictive accuracy of criminal intentions is more fully understood, the study of criminal decision making will benefit most from triangulating the findings from a host of studies that use a variety of research methodologies, including the HSM.

Related Entries

- ▶ [Bayesian Updating and Crime](#)
- ▶ [Deterrence](#)
- ▶ [Deterrence: Actual Versus Perceived Risk of Punishment](#)

- ▶ [Rational Choice and Prospect Theory](#)
- ▶ [Rational Choice Theory](#)
- ▶ [Rational Choice, Deterrence, and Crime: Sociological Contributions](#)

Recommended Reading and References

- Ajzen I, Fishbein M (1980) Understanding attitudes and predicting social behavior. Prentice-Hall, Englewood Cliffs
- Armitage CJ, Conner M (2001) Efficacy of the theory of planned behavior: a meta-analytic review. *Br J Soc Psychol* 40:471–499
- Bachman R, Paternoster R, Ward S (1992) The rationality of sexual offending: testing a deterrence/rational choice conception of sexual assault. *Law Soc Rev* 26:343–372
- Bentham J (1970) An introduction to the principles of morals and legislation. Oxford University Press, New York (Original work published 1789)
- Bouffard JA (2002a) Methodological and theoretical implications of using subject-generated consequences in tests of rational choice theory. *Justice Q* 19:747–769
- Bouffard JA (2002b) The influence of emotion on rational decision making in sexual aggression. *J Crim Justice* 30:121–134
- Bouffard JA (2007) Rational choice theory revisited: a preliminary examination of the perceived importance of consequences in offender decision-making. *Int J Crime Crim Justice Law* 2:49–67
- Bouffard J (2011) 'In the heat of the moment': mediating versus moderating relationships between sexual arousal and perceived sanctions. *J Crime Justice* 34:24–44
- Bouffard J, Exum ML, Collins P (2010) Methodological artifacts in tests of rational choice theory. *J Crim Justice* 38:400–409
- Carmichael S, Piquero AR (2004) Sanctions, perceived anger, and criminal offending. *J Quant Criminol* 20:371–393
- Elis LA, Simpson SS (1995) Informal sanction threats and corporate crime: additive versus multiplicative models. *J Res Crime Delinquency* 32:399–424
- Exum ML (2002) The application and robustness of the rational choice perspective in the study of intoxicated and angry intentions to aggress. *Criminol* 40:933–966
- Exum ML, Turner MG, Hartman JS (2012) Self-reported intentions to offend: all talk and no action? *Am J Crim Justice*. 4:523–543
- Grasmick HG, Bursik RJ (1990) Conscience, significant others, and rational choice: extending the deterrence model. *Law Soc Rev* 24:837–861
- Higgins GE (2007) Digital piracy, self-control theory, and rational choice: an examination of the role of value. *Int J Cyber Criminol* 1:33–55

- Klepper S, Nagin D (1989) The deterrent effect of perceived certainty and severity of punishment revisited. *Criminol* 27:721–746
- Loewenstein G, Nagin D, Paternoster R (1997) The effect of sexual arousal on expectations of sexual forcefulness. *J Res Crime Delinq* 34:443–473
- Nagin DS (1978) General deterrence: a review of the empirical evidence. In: Blumstein A, Cohen J, Nagin D (eds) *Deterrence and incapacitation: estimating the effects of criminal sanction on crime rates*. National Academy Press, Washington, DC, pp 95–139
- Nagin D, Paternoster R (1993) Enduring individual differences and rational choice theories of crime. *Law Soc Rev* 27:467–496
- Nagin DS, Paternoster R (1994) Personal capital and social control: the deterrence implications of a theory of individual differences in criminal offending. *Criminology* 32:581–606
- Nagin DS, Pogarsky G (2001) Integrating celerity, impulsivity, and extralegal sanction threats into a model of general deterrence: theory and evidence. *Criminology* 39:865–892
- Ogilvie J, Stewart A (2010) The integration of rational choice and self-efficacy theories: a situational analysis of student misconduct. *Aust New Zeal J Criminol* 43:130–155
- Paternoster R (2010) How much do we really know about criminal deterrence? *J Crim Law Criminol* 100:765–824
- Paternoster R, Simpson S (1996) Sanction threats and appeals to morality: testing a rational choice model of corporate crime. *Law Soc Rev* 30:549–583
- Piliavin I, Gartner R, Thornton C, Matsueda RL (1986) Crime, deterrence, and rational choice. *Am Sociol Rev* 51:101–119
- Piquero AR, Bouffard JA (2007) Something old, something new: a preliminary investigation of Hirschi's redefined self-control. *Justice Q* 24:1–27
- Piquero AR, Pogarsky G (2002) Beyond Stafford and War's reconceptualization of deterrence: personal and vicarious experiences, impulsivity, and offending behavior. *J Res Crime Delinq* 39:153–186
- Piquero A, Tibbetts S (1996) Specifying the direct and indirect effects of low self-control and situational factors in offenders' decision-making: toward a more complete model of rational offending. *Justice Q* 13:481–510
- Piquero NL, Exum ML, Simpson SS (2005) Integrating the desire-for-control and rational choice in a corporate crime context. *Justice Q* 22:252–280
- Pogarsky G (2002) Identifying “deterable” offenders: implications for research on deterrence. *Justice Q* 19:431–452
- Pogarsky G (2004) Projected offending and contemporaneous rule-violation: implications for heterotypic continuity. *Criminology* 42:111–136
- Pogarsky G, Piquero AR (2003) Can punishment encourage offending? Investigating the “resetting” effect. *J Res Crime Delinq* 40:95–120
- Pogarsky G, Piquero AR (2004) Studying the reach of deterrence: can deterrence theory help explain police misconduct? *J Crim Justice* 32:371–386
- Pratt TC, Cullen FT, Blevins KR, Daigle LE, Madensen TD (2008) The empirical status of deterrence theory: a meta-analysis. In: Cullen F, Wright J, Blevins K (eds) *Taking stock: the status of criminological theory*. Transaction, New Brunswick, pp 367–395
- Simpson SS, Piquero NL (2002) Low self-control, organizational theory, and corporate crime. *Law Soc Rev* 36:509–548
- Sitren AH, Applegate BK (2007) Testing the deterrent effects of personal and vicarious experience with punishment and punishment avoidance. *Deviant Behav* 28:29–55
- Strelan P, Boeckmann RJ (2006) Why drug testing in elite sports does not work: perceptual deterrence theory and the role of personal moral beliefs. *J Appl Soc Psychol* 36:2909–2934
- Tibbetts SG (1999) Differences between women and men regarding decisions to commit test cheating. *Res High Educ* 40:323–342
- Tibbetts SG, Herz DC (1996) Gender differences in factors of social control and rational choice. *Deviant Behav Interdiscip J* 17:183–208

School Aggression

► [Bullying Prevention: Assessing Existing Meta-Evaluations](#)

School Bullying as a Risk Factor for Later Criminal Offending

Maria M. Ttofi¹, David P. Farrington¹ and Friedrich Lösel^{1,2}

¹Institute of Criminology, University of Cambridge, Cambridge, UK

²Institute of Psychology, University of Erlangen–Nuremberg, Erlangen, Germany

Overview

School bullying has received the attention of researchers and program planners in both developed and developing countries. It is a special category of aggressive behavior that

has been addressed through numerous anti-bullying programs and, in some cases, through wider multiple component programs. Various anti-bullying agencies have highlighted the importance of intervention research for the development of safer school communities, where students can develop their full potential without being exposed to bullying and its detrimental effects. A vast number of cross-sectional studies have provided evidence of the negative impact of bullying on children's concurrent health.

This entry reports on an updated systematic review and meta-analysis that was undertaken under the aegis of the Swedish National Council for Crime Prevention and further supported by the British Academy and conducted by the current authors. Only longitudinal prospective studies were included in the review, which aimed to examine to what extent school bullying predicts later offending and violence. Significant effect sizes were found even after controlling for other major childhood risk factors. Being a bully increased the likelihood of being an offender by more than half and increased the likelihood of being violent by two thirds. These results either reflect the persistence of an underlying aggressive or antisocial tendency or a facilitating effect of school bullying on later offending and violence (or both).

The implication is that high quality bullying prevention programs (and possibly multiple component programs which also target aggression) should be promoted. They could be viewed as an early form of crime prevention. They can potentially have long-term effects by improving the future psychosocial adjustment of school bullies and reducing the associated health, welfare, education, and other costs.

Introduction

School bullying has recently become a topic of major public concern and has attracted a lot of media attention, with articles in major newspapers and magazines reporting cases of children who committed (or attempted) suicide because

of their victimization at school and parents suing school authorities for their failure to protect their offspring from continued bullying victimization (e.g., Ttofi and Farrington 2012). There is, nevertheless, a number of "skeptics" who still perceive school bullying as being part of a normal developmental process, or as one of those school experiences that prepare children for the grown-up world. Scientific evidence regarding possible detrimental effects of school bullying on children's mental health and future psychosocial adjustment can only be provided through a systematic review and meta-analysis, providing an unbiased standardized effect size and defining the magnitude of the effect.

Background Research

School bullying is a special category of aggressive behavior involving repeated unprovoked acts against less powerful (emotionally or physically) individuals (Farrington 1993; Olweus 1993). Of course, schools, like other institutions, will always be a place in which the basic human motive of aggression will be demonstrated. However, school bullying should not be confused with more or less normal aggressive interactions such as rough and tumble play.

Scientific interest in the problem of bullying and its negative short-term and long-term effects emerged after the well-publicized suicides of three Norwegian boys in 1982, which were attributed to severe peer bullying (Olweus 1993). School bullying has gradually become a topic of major public concern via "bullying awareness days," national initiatives in various (European) countries (Smith and Brain 2000), and anti-bullying research networks across the world (e.g., Anti-Bullying Alliance; BRNET; International Observatory for Violence in Schools; PREVNet).

Any suggestion regarding the short-term negative impact of peer aggression and victimization seems reasonable even to the lay mind. Establishing, on the other hand, the long-term effects of school bullying and arguing that children involved in peer aggression are more

likely to follow an antisocial path (compared with noninvolved students) is more challenging. Some early longitudinal studies did provide evidence of the long-term impact of school bullying and, notably, established the intergenerational transmission of school bullying. In the Cambridge Study in Delinquent Development, for example, boys who were bullies at age fourteen tended, at age thirty-two, to have children who were bullies (Farrington 1993). As another example, in his follow-up study of over 700 Stockholm boys, Olweus (1993) reported that 36 % of bullies at ages thirteen to sixteen were convicted three or more times between ages sixteen and twenty-four, compared with 10 % of the remainder.

There have been surprisingly few recently published longitudinal studies on the developmental pathways of children involved in school bullying since the seminal work of Olweus in Scandinavia and some other European examples. Two special issues in peer-reviewed journals have recently been published in an attempt to address this gap in research literature (Farrington et al. 2011; Ttofi et al. 2011a). Both issues presented new findings on the long-term negative consequences of school bullying based on major prospective longitudinal studies from around the world. Longitudinal investigators of twenty-nine studies conducted analyses for a more comprehensive British Academy funded project, which examined the long-term association of school bullying with both internalizing (such as anxiety, self-esteem, and stress) and externalizing (such as aggression, alcohol, and drug use) problems (see Farrington et al. 2012, Table 4, for a list of all contributors).

The special issue of *Criminal Behaviour and Mental Health* focused on the association between bullying perpetration at school and offending later in life. A systematic review and meta-analysis on the topic was carried out (Ttofi et al. 2011c). The special issue of the *Journal of Aggression, Conflict and Peace Research* focused on the association between bullying victimization (i.e., being bullied) and internalizing problems later in life, such as anxiety and depression. A systematic review and

meta-analysis was again carried out examining the extent to which bullying victimization at school predicted depression (Ttofi et al. 2011b), showing that the probability of being depressed up to seven years later in life ($M = 7.13$ years; $SD = 8.79$) was significantly higher for victims of school bullying than for control students, i.e., children not involved in school bullying.

Building upon the above-mentioned research activities, an effort was made to update the relevant systematic reviews (Farrington et al. 2012) and to study further outcomes, such as violence (Ttofi et al. 2012). This entry presents results from the updated systematic review on the association of school bullying with offending later in life. Additional analyses are presented on the long-term link of bullying with violence.

Methods

The main objectives of the systematic review were two-fold. Firstly, to assess whether bullying at school (perpetration and victimization) was a significant risk factor predicting offending and violence later in life (unadjusted effect sizes). Secondly, to assess whether these associations were still significant after controlling for other major childhood risk factors, measured at the baseline period (adjusted effect sizes). Results on offending and violence were carefully treated in separate analyses and the outcome measures under each category generally did not overlap. However, it is possible that in some studies, outcome measures such as “police arrests” would include violence. “Offending” included outcome measures such as police or court contact, property offending, criminal convictions, property theft, vandalism, shoplifting, vehicle theft, etc. “Violence” included outcome measures such as forced sexual contact, criminal violence, physical fights, violent convictions, violent offending, weapon carrying, assault, etc.

Further analyses were conducted to investigate moderators that might explain variability in effect sizes, such as the age at which bullying was measured (Time 1), the age at which the outcome measures were taken (Time 2), the

number of covariates controlled for in the adjusted effect sizes, the length of the follow-up period (measured in years), and the way in which the outcomes were measured (i.e., official data versus self-reports).

Stringent inclusion and exclusion criteria were set in advance. For example, reports were included only if they were based on prospective longitudinal data. The predictor must have been a measure of school bullying (and not other more general forms of peer aggression/victimization) and must have preceded the outcome (i.e., offending or violence). A clear measurement of offending and/or violence must have been included in the report as an outcome measure. Studies were included if participants were school-aged children in the community and exposure to bullying (perpetration and victimization) specified the school years. Published and unpublished reports of the literature were included in order to minimize the possibility of publication bias in the results.

Reports were excluded if the character of the data was qualitative in nature (e.g., qualitative data based on interviews) and did not allow calculation of an effect size. This did not apply if a qualitative method (e.g., interviews or observation studies) was used to obtain a quantitative measure. If the outcome measure (offending or violence) was part of a wider theoretical construct (e.g., externalizing or antisocial behavior), then the relevant report was again excluded. Reports based on clinic samples or incarcerated youth were also excluded.

Extensive searches were carried out and a detailed description of them can be found in the Swedish report (Farrington et al. 2012) and the most recent work focusing on violent outcomes (Ttofi et al. 2012). In total, the same searching strategies were repeated in 19 electronic databases, and the full volumes of 63 journals were searched either online or in print. In the Swedish report, readers can also find detailed tables of the key features of each report, such as the sample size, the country where the study took place, the exact confounds controlled for at the baseline period, etc.

Results

In total, 661 reports concerned with the association of school bullying with internalizing (e.g., anxiety, depression, self-esteem, etc.) and externalizing (e.g., aggressive behavior, conduct problems, offending, etc.) problems were located. All reports were screened in line with the inclusion and exclusion criteria and classified in five different categories (see Farrington et al. 2012, Table 5). Further to a detailed screening of all manuscripts, a total number of 48 reports, corresponding to 29 longitudinal studies, presented data on the long-term association of school bullying (perpetration and victimization) with offending in adolescence or young adulthood (see Farrington et al. 2012, Table 6). A total number of 51 reports from 28 longitudinal studies were included in the systematic review on the association of school bullying (perpetration and victimization) with violence in adolescence or young adulthood (see Ttofi et al. 2012, Table 1).

When different manuscripts relating to the same longitudinal study reported different effect sizes (because of differences, e.g., in the sample size or in the follow-up period that the authors have used), the combination of effect sizes across reports is not straightforward as these effect sizes are based on dependent samples. These dependencies were taken into account, as ignoring them would result in standard errors that were too small, often by a large degree. Advice from leading experts in the field was sought on this matter (Wilson 2010). Clear rules were set in advance for computing effect sizes across reports from the same longitudinal study (see Farrington et al. 2012; Ttofi et al. 2012).

Bullying Perpetration at School and Offending Later in Life

Eighteen studies provided an effect size for bullying perpetration versus offending. The summary unadjusted effect size across the 18 studies was $OR = 2.64$ (95 % CI: 2.17–3.20; $z = 9.83$) for the random-effects model. The random-effects model was used since the

heterogeneity test, Q , of 84.89 was highly significant at $p = 0.0001$. When the three studies with only unadjusted effect sizes were excluded, the summary effect size for the remaining 15 studies – for the random-effects model – was $OR = 2.54$ (95 % CI: 2.05–3.14, $z = 8.52$). Again, there was significant variability in effect sizes across these studies ($Q = 76.03$, $p = 0.0001$).

After controlling for covariates, the adjusted summary effect size was reduced to $OR = 1.89$ (95 % CI: 1.60–2.23; $z = 7.49$) but this was still highly significant (see Farrington et al. 2012; Figs. 3 and 4). This OR indicates quite a strong relationship between bullying perpetration and later offending. For example, if a quarter of children were bullies and a quarter were offenders, this value of the OR would correspond to 34.5 % of bullies becoming offenders, compared with 21.8 % of non-bullies. Thus, being a bully increases the risk of being an offender (even after controlling for other childhood risk factors) by more than half.

For the adjusted summary effect size, various moderators were investigated to explain the heterogeneity in effect sizes across studies, which was significant ($Q = 36.82$, $p = 0.001$). These included the number of covariates controlled for at baseline (range: 1–20; $M = 7.00$; $SD = 5.22$), the age at which school bullying was measured (range: 6.23–15.54; $M = 11.26$; $SD = 2.68$), the age of participants when outcome measures were taken (range: 10.00–24.64; $M = 17.10$; $SD = 4.91$), and the length of the follow-up period, measured in years (range: 0.42–16.50; $M = 5.84$; $SD = 4.56$).

The age at which bullying was measured was positively associated with the effect size, but the regression coefficient was not statistically significant ($B = 0.019$, $SE = 0.024$, $p = 0.428$). The length of the follow-up period was significantly negatively associated with the effect size ($B = -0.027$, $SE = 0.012$, $p = 0.018$). As expected, the age of the study participants when outcome measures were taken was significantly negatively related to the effect size ($B = -0.025$, $SE = 0.012$, $p = 0.039$). The above two negative relationships suggest that bullying perpetration

has a stronger effect in the short term. The relationship between the number of covariates controlled for and the effect size was in the expected negative direction and also significant ($B = -0.027$, $SE = 0.013$, $p = 0.037$). Therefore, the adjusted effect size decreased as the number of covariates controlled for increased.

Other moderators that may explain variability in effect sizes include the type of longitudinal studies (i.e., prospective versus retrospective) and the way in which the outcomes were measured (i.e., official data versus self-reports). In the Farrington et al. (2012) report, the reader can obtain information about these moderators (see their Table 6). Only three studies out of fifteen presented outcome measures based on official records for offending, making a moderator analysis inappropriate (due to uneven study numbers). Finally, only one study presented results based on a retrospective measure of bullying victimization, so any analyses on this matter would be meaningless.

If the studies included in a meta-analysis are a biased sample of all relevant studies, then the mean effect computed will reflect this bias (Borenstein et al. 2009, p. 277). It is clear from our thorough searching strategies that every precaution was taken to ensure that all eligible studies would be represented in the meta-analysis. In order to further increase the validity of the meta-analysis findings, a number of publication bias analyses were carried out.

Firstly, the Duval and Tweedie's Trim-and-Fill procedure was used. This technique displays the differences in effect sizes that could be attributable to bias by imputing effect sizes until the error distribution more closely approximates normality, offering the best *estimate* of the unbiased effect size (Borenstein et al. 2009, p. 286). No imputed effect sizes appeared on the relevant funnel plot (they would have been presented as solid black dots; see Farrington et al. 2012, Fig. 6), indicating no publication bias. The imputed summary effect size (represented by a solid black diamond) had not shifted at all.

Indeed, under the fixed effect model, the point estimate and 95 % confidence interval for the

combined studies was 1.86 (95 % CI: 1.71–2.03). Using Trim-and-Fill procedure, these values remained unchanged. Under the random-effects model, the point estimate and 95 % confidence interval for the combined studies was 1.89 (95 % CI: 1.60–2.23). Using Trim-and-Fill procedure, these values were again unchanged.

Furthermore, Rosenthal's Fail-Safe N test (Rosenthal 1979) was conducted. One concern of publication bias is that some nonsignificant studies are missing from a given analysis and that these studies, if included, would nullify the observed effect. Rosenthal suggested that, rather than simply speculate about the impact of the missing studies, we compute the number of nonsignificant studies that would be required to nullify the effect. If this number is small, then there is reason for concern because some nonsignificant studies may have been never communicated to the scientific community (e.g., due to "publication bias"). However, if this number is large, one can be confident that the treatment effect, while possibly inflated by the exclusion of some studies, is nevertheless not zero.

Bullying Perpetration at School and Violence Later in Life

A total number of 15 studies were concerned with the association of bullying perpetration with aggression and violence later in life. The unadjusted summary effect size across these studies was OR = 3.09 (95 % CI: 2.35–4.07; $z = 8.10$). For one study, only an unadjusted effect size was available. The unadjusted effect size for the remaining 14 studies was OR = 2.97 (95 % CI: 2.25–3.92; $z = 7.71$; $Q = 151.81$, $p = 0.0001$; $I^2 = 91.44$). All individual studies yielded a significant effect size (see Ttofi et al. 2012, Fig. 1). After controlling for covariates, the adjusted summary effect size was reduced to OR = 2.04 (95 % CI: 1.69–2.45; $z = 7.53$) but this was still highly significant (see Ttofi et al. 2012, Fig. 2). This OR indicates quite a strong relationship between bullying perpetration and later violence. For example, if a quarter of children were bullies and a quarter were violent, this value of OR would correspond to 35.8 % of bullies becoming

violent, compared with 21.4 % of non-bullies. Thus, being a bully increases the risk of being violent (even after controlling for other childhood risk factors) by two thirds.

Although all individual studies yielded an effect size supporting the link between school bullying and aggression/violence later in life, the magnitude and the significance of the effect varied across these studies. Various moderator analyses were conducted in order to explain this variability ($Q = 75.801$, $p = 0.0001$, $I^2 = 82.85$). These included the number of covariates controlled for at baseline (range: 2–20; $M = 6.93$; $SD = 5.25$), the age at which school bullying was measured (range: 8.00–15.54; $M = 12.04$; $SD = 2.35$), the age of participants when outcome measures were taken (range: 10.00–24.64; $M = 17.65$; $SD = 4.83$), and the length of the follow-up period, measured in years (range: 0.42–16.50; $M = 5.61$; $SD = 4.88$).

The age of participants when bullying was measured was significantly negatively correlated with the effect size ($B = -0.065$; $SE = 0.021$; $p = 0.002$), suggesting that the younger the children were when they exhibited this form of problem behavior, the more likely it was that they would be violent later in life. The age of participants when outcome measures were taken was also significantly negatively related to the effect size ($B = -0.033$; $SE = 0.009$; $p = 0.0005$). In other words, the lower the age of the participants when aggression or violence was measured, the larger the effect, possibly because this was associated with a shorter follow-up period. This is consistent with the significant negative association between the length of follow-up period and the magnitude of the effect size ($B = -0.017$; $SE = 0.009$; $p = 0.051$). As expected, the magnitude of the effect size decreased as the number of confounds controlled for increased ($B = -0.013$; $Intercept = 0.668$; $SE = 0.010$; $p = 0.185$), but the relevant regression coefficient was not significant.

As with the previous meta-analysis, a number of sensitivity analyses were conducted. Firstly, the Duval and Tweedie's Trim-and-Fill procedure was performed. Three imputed effect

sizes appeared on the relevant funnel plot (see Ttofi et al. 2012, Fig. 3) and the imputed summary effect size (represented by a solid black diamond) had shifted slightly, suggesting a trivial overestimation of the summary effect size.

As already mentioned, the difference was very small. Under the fixed effect model, the point estimate and 95 % confidence interval for the combined studies was 1.83 (95 % CI: 1.71–1.95). Using Trim-and-Fill procedure, the imputed point estimate was 1.76 (95 % CI: 1.65–1.88). Under the random-effects model, the point estimate and 95 % confidence interval for the combined studies was 2.04 (95 % CI: 1.69–2.45). Using Trim-and-Fill procedure, the imputed point estimate was 1.77 (95 % CI: 1.45–2.16).

Finally, the Rosenthal's Fail-Safe N test was performed. This meta-analysis incorporated data from 14 studies, which yielded a z -value of 17.12216 and corresponding 2-tailed p -value of 0.000001. The fail-safe N is 1055. This means that one would need to locate and include 1055 "null" studies in order for the combined 2-tailed p -value to exceed 0.050. Put another way, 75.4 missing studies would be needed for every observed study for the effect to be nullified. It is impossible that such a large number of studies were conducted but not published or not included in our analysis.

Further Findings

Further analyses were performed to examine the association of bullying victimization with later offending before (Unadjusted OR = 1.32; 95 % CI: 1.13–1.55, $z = 3.40$) and after controlling for other major childhood risk factors (Adjusted OR = 1.14; 95 % CI: 0.997–1.310, $z = 1.91$) and relevant forest plots are shown in the Farrington et al. (2012) report (see Figs. 11 and 12). This was a very weak relationship. Moderator analyses and publication bias analyses similar to those presented in the current entry were also presented in that report.

Finally, analyses were performed to examine the association of bullying victimization with later violence before (Unadjusted

OR = 1.65; 95 % CI: 1.42–1.92; $z = 6.48$) and after (Adjusted OR = 1.42; 95 % CI: 1.248–1.6172; $z = 5.3117$) controlling for covariates (see Ttofi et al. 2012, Figs. 4 and 5). Again, moderator analyses and publication bias analyses similar to those presented in the current entry are also presented in that report.

Sensitivity analyses were performed for these two sets of meta-analyses and the results showed in general no evidence of publication bias (see Farrington et al. 2012; Ttofi et al. 2012).

Possible Controversies in the Literature

The results of these systematic reviews and meta-analyses suggest that there are long-term detrimental effects of school bullying on later offending and violence. This was even the case when confounded variables that are risk factors for bullying or victimization as well as the undesirable outcomes were controlled for. Therefore, one can conclude that school bullying is an independent predictor of the later psychosocial development of perpetrators as well as of victims. It is the first time that this conclusion is not only based on a few selected primary studies and narrative reviews, but, instead, on comprehensive meta-analyses of prospective longitudinal studies that included new data from a substantial body of yet unpublished research. The findings remained robust in sensitivity analyses testing potential publication biases, of which there was no sign.

The relation of bullying perpetration with later offending and violence might reflect the persistence of an underlying disposition for antisocial behavior that has different manifestations over time (Farrington 1993; Lösel and Bliesener 2003). However, as the relation remained after controlling for other childhood risk factors, bullying perpetration may also increase the likelihood of later offending and violence.

Of course, one should acknowledge that any direct mention of causality should be carefully treated. Although most studies use bullying as the predictor of later outcomes, implying in

this way a specific temporal sequence, alternative models have been suggested. Very few bullying studies have examined alternative models on whether bullying is a cause or a consequence of psychopathic behavior (e.g., Boulton et al. 2010; Kim et al. 2006). This is not a trivial matter and it would shed more light on the temporal sequence and the causal ordering between bullying and other internalizing or externalizing behaviors. The substantial adjusted effect size for victimization versus later depression found in a previous meta-analysis (Farrington et al. 2012; Ttofi et al. 2011b), for example, suggests in a way that the frequent internalizing symptoms of victims are not only a trigger for being bullied, but a psychological consequence.

Systematic reviews on risk factors are important as they can advance theory and also help to develop effective prevention programs (Murray et al. 2009). For example, it would be interesting to examine whether victims of bullying suffer from low self-esteem or whether school bullies lack cognitive or affective empathy. Such findings, based on relevant systematic reviews, could guide future intervention initiatives, while also refining theory about the causes of bullying perpetration and victimization.

Open Questions and Future Research Directions

In the current meta-analysis, studies were included and analyzed based on “level analyses.” Levels of bullying perpetration were compared with later levels of offending and violence. It would also have been interesting to complete a systematic review on “change analyses,” examining whether changes in bullying from Time 1 to Time 2 are followed by changes in an outcome from Time 2 to Time 3. However, there are hardly any studies on this matter, since such analyses would require relevant data from multiple waves. Such analyses would allow, to an extent, making safer inferences about causality, although change data are subject to greater variability than level data. Systematic reviews

of longitudinal studies which control for confounded variables can give some hints on whether variables are simple correlational risk factors, risk markers, or causal risk factors (Kraemer et al. 2005).

Future research should also examine possible gender-specific and ethnic-specific effects of bullying on later violent behavior and offending. Such information was hardly ever available in the current literature. To investigate and disentangle the impact of these and other variables on the relation between bullying and later outcomes, more longitudinal studies with a sound control for childhood risk factors are needed. The results of meta-regression analyses were not always as expected in the meta-analyses for the British Academy project on “Health and Criminal Outcomes of School Bullying” because of the large differences in the type of covariates researchers controlled for. However, one should note that the lack of a sufficient number of studies with consistent patterns of characteristics is a typical problem in meta-analyses (Lipsey 2003).

Future research should also examine mediators or possible causal mechanisms between school bullying and the various outcomes. The underlying mechanisms, for example, may be the reinforcement obtained by dominating others and the development of an identity as a “bully” that goes beyond the school context.

Conclusions

This is the first time that research has provided an *unbiased standardized effect size* regarding the predictive efficiency of school bullying in relation to violence and offending later in life. The significant summary effect sizes have important implications for policy and practice as they give a stronger voice to anti-bullying agencies and reestablish the moral imperative of school communities to create an appropriate violence-free school climate.

High quality bullying prevention programs should be promoted (Farrington and Ttofi 2009; Ttofi and Farrington 2011). They could be

viewed as an early form of crime prevention. These programs can potentially have long-term effects by improving the future psychosocial adjustment of school bullies and reducing the associated health, welfare, education, and other costs. The effectiveness of other school-based programs for the prevention of problem behaviors has been examined through thorough systematic reviews (e.g., Wilson et al. 2001) and it is possible that such programs, or other general multicomponent programs, might have positive effects in reducing aggression and bullying behavior.

Previous research has provided strong evidence about the monetary value of saving high-risk youth (Cohen and Piquero 2009). Children involved in school bullying are undoubtedly youth at risk, with significantly higher probabilities of following an antisocial path. What remains unanswered is the identification of protective factors that interrupt the continuity from school bullying to later adverse outcomes and confer resiliency on this special category of high-risk youth (Ttofi and Farrington 2012).

Glossary

- CI** Confidence interval
M Mean
OR Odds ratio
SD Standard deviation

Related Entries

- ▶ [Juvenile Violence](#)
- ▶ [Risk Assessment, Classification, and Prediction](#)
- ▶ [Risk factors for Adolescent Sexual Offending](#)
- ▶ [School Crime Statistics](#)
- ▶ [School Social Organization, Discipline Management, and Crime](#)
- ▶ [School Structural Characteristics and Crime](#)
- ▶ [School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis](#)

Recommended Reading and References

- Borenstein M, Hedges LV, Julian PTH, Rothstein HR (2009) Introduction to meta-analysis. Wiley, Chichester
- Boulton MJ, Smith PK, Cowie H (2010) Short-term longitudinal relationships between children's peer victimization/bullying experiences and self-perceptions: evidence for reciprocity. *School Psychol Int* 31(3):296–311
- Cohen MA, Piquero AR (2009) New evidence on the monetary value of saving a high risk youth. *J Quant Criminol* 25:25–49
- Farrington DP (1993) Understanding and preventing bullying. In: Tonry M (ed) *Crime and justice*, vol 17. University of Chicago Press, Chicago
- Farrington DP, Ttofi MM (2009) School-based programs to reduce bullying and victimization. *Campbell Systematic Reviews* 6
- Farrington DP, Ttofi MM, Lösel F (2011) Editorial: school bullying and later offending. *Crim Behav Mental Health* 21(2):77–79
- Farrington DP, Lösel F, Ttofi MM, Theodorakis N (2012) School bullying, depression and offending behaviour later in life: an updated systematic review of longitudinal studies. Swedish National Council for Crime Prevention, Stockholm
- Kim YS, Leventhal BL, Koh Y-J, Hubbard A, Boyce TW (2006) School bullying and youth violence: causes or consequences of psychopathologic behavior? *Arch Gen Psychiat* 63:1035–1041
- Kraemer HC, Lowe KK, Kupfer DJ (2005) *To your health: how to understand what research tells us about risks*. Oxford University Press, New York
- Lipsey MW (2003) Those confounded moderators in meta-analysis: good, bad, and ugly. *Ann Am Acad Polit Soc Sci* 587(1):69–81
- Lösel F, Bliesener T (2003) Aggression und delinquenz unter jugendlichen: Untersuchungen von kognitiven und sozialen Bedingungen. Luchterhand, Hermann
- Murray J, Farrington DP, Eisner MP (2009) Drawing conclusions about causes from systematic reviews of risk factors: the Cambridge quality checklists. *J Exp Criminol* 5(1):1–23
- Olweus D (1993) *Bullying at school: what we know and what we can do*. Blackwell, Oxford
- Rosenthal R (1979) The "File drawer problem" and tolerance for null results. *Psychological Bulletin* 86:638–641.
- Smith PK, Brain P (2000) Bullying in schools: lessons from two decades of research. *Aggressive Behavior* 26:1–9
- Ttofi MM, Farrington DP (2011) Effectiveness of school-based programs to reduce bullying: a systematic and meta-analytic review. *J Exp Criminol* 7(1):27–56
- Ttofi MM, Farrington DP (2012) Risk and protective factors, longitudinal research and bullying prevention. *New Direct Youth Develop* 133:85–98

- Ttofi MM, Farrington DP, Lösel F (2011a) Editorial: health consequences of school bullying. *J Aggress Conflict Peace Res* 3(2):60–62
- Ttofi MM, Farrington DP, Lösel F, Loeber R (2011b) Do the victims of school bullies tend to become depressed later in life? A systematic review and meta-analysis of longitudinal studies. *J Aggress Conflict Peace Res* 3(2):63–73
- Ttofi MM, Farrington DP, Lösel F, Loeber R (2011c) The predictive efficiency of school bullying versus later offending: a systematic/meta-analytic review of longitudinal studies. *Crim Behav Mental Health* 21(2):80–89
- Ttofi MM, Farrington DP, Lösel F (2012) School bullying as a risk marker for aggression and violence later in life: a systematic review and meta-analysis of prospective longitudinal studies. *Aggression and Violent Behavior* (Accepted, in press)
- Wilson D (2010) Email communication, October 25
- Wilson DB, Gottfredson DC, Najaka SS (2001) School-based prevention of problem behaviors: a meta-analysis. *J Quant Criminol* 17(3):247–272

implementing No Child Left Behind (NCLB, the most recent national education-reform initiative in the USA) stipulates that school systems must have programs in place to reduce levels of violence. The legislation authorizing NCLB has a specific provision that “persistently dangerous” schools be identified by the states and that students attending such schools be given the option of transferring to another school. In practice, only a few dozen schools nationwide have been identified as persistently dangerous, but that fact reflects politics more than the empirical reality.

As it turns out, obtaining reliable information about crime in schools is a challenge – for researchers as well as for state and federal officials. There are several sources of data in addition to the schools’ own reports, but each source is error prone. There are some rather remarkable differences among them with respect to estimated crime rates and patterns.

Based on our review of the evidence, we offer several lessons about crime in school and about the sources of statistics on crime. First, with respect to sources, we find that crime surveys with adolescent respondents give results that are remarkably sensitive to the survey method. For example, in-school victimization rates from the Youth Behavior Risk Surveillance System (YRBSS) are more than 20 times as high as comparable victimization rates estimated from the National Crime Victimization Survey (NCVS), and they produce qualitatively different patterns over time. It is always a good idea to validate estimates from one source against others, but that may be especially true with youth surveys. On the other hand, we find administrative data to be better than expected.

We are able to reach some conclusions about crime patterns that we feel are sufficiently robust to be believable. Among our conclusions are these:

- Victimization rates for 12–18-year-olds are as high in school as out (and higher for the 12–14-year-olds), despite the fact that youths spend many fewer waking hours in school.
- Victimization patterns are quite different in and out of school. Violent crimes tend to be

School Crime

- ▶ [School Security Practices and Crime](#)

School Crime Statistics

Philip J. Cook^{1,2}, Denise C. Gottfredson³ and Chongmin Na⁴

¹Sanford School of Public Policy, Duke University, Durham, NC, USA

²NBER, Cambridge, MA, USA

³Department of Criminology and Criminal Justice, University of Maryland, College Park, MD, USA

⁴School of Human Science and Humanities, University of Houston-Clear Lake, Houston, TX, USA

Overview

By rights, schools should be sanctuaries against criminal victimization, but the truth is otherwise. A threatening environment is not conducive to academic success. The US federal law

less serious in school (where only 1 % of homicide victimizations occur).

- Property crime victimization rates are remarkably homogeneous across race and also sex; alternative data sources are not consistent with respect to racial patterns of school violence.
- “Persistently dangerous schools” appear to be a real phenomenon and far more prevalent than are officially identified as such.

We begin this entry with a discussion of the important data sources and then go on to discuss victimization patterns for students. We then explore the issue of persistently dangerous schools using a unique data set. We conclude with an analysis of arrest patterns.

Data Sources

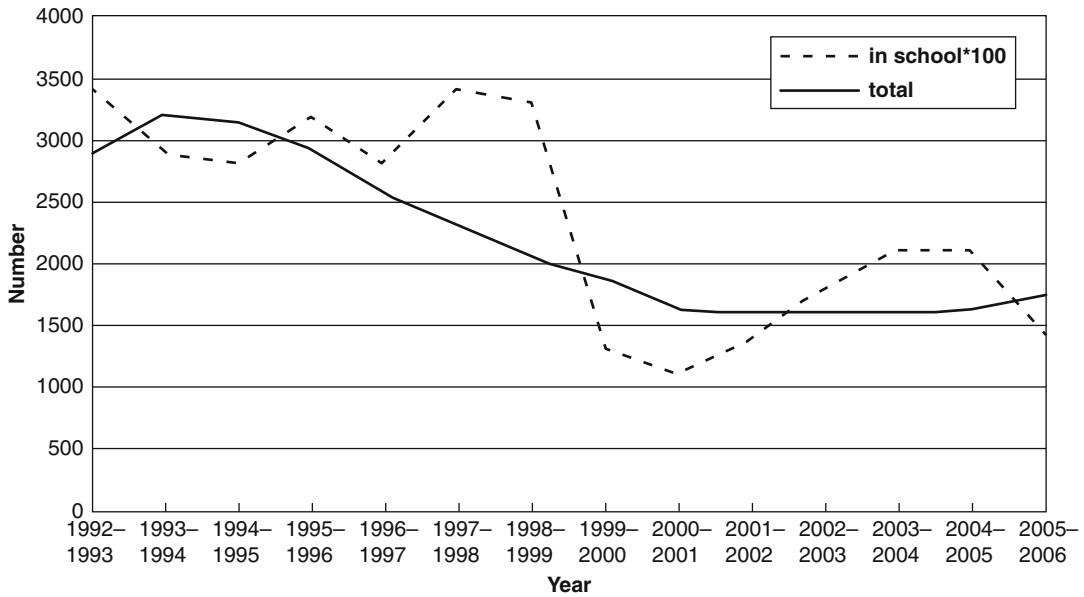
The primary source of the US crime data for many purposes is the FBI’s Uniform Crime Reports, compiled from crimes known to the police and reported by police departments. The UCR’s crime data do not provide information on the characteristics of victims and are of little help in estimating crime rates in schools. Some jurisdictions report crimes in much more detail through the National Incident-Based Reporting System (NIBRS): in this system police agencies submit a record of each known crime that includes the age, sex, and race of the victim, the location of the crime, and the characteristics of the perpetrator (when known). These data can be used to provide a detailed description of crimes involving school-aged youths, distinguishing, for example, between crimes on school grounds and elsewhere (Jacob and Lefgren 2003). There are two problems, however, with this source. First, participation rates are very low: only 20 % of police agencies, representing 16 % of the US population, were participating in NIBRS as of 2003 (<http://www.ojp.usdoj.gov/bjs/nibrstatus.htm>, accessed November 19, 2008). And second, crimes committed on school property may be less likely to become known to the police than crimes occurring elsewhere.

As a result of the limitations of police data, school crime statistics are usually generated from

school reports or surveys. In the School Survey on Crime and Safety (SSoCS), a sample of about 3,000 public-school principals report to the US Department of Education the number of violent incidents and thefts and indicate how many of these incidents were reported to the police. In addition there are several recurrent sample surveys: the National Crime Victimization Survey and the biannual School Crime Supplement to this Survey (sponsored by the National Center for Education Statistics, or NCES) and the Youth Risk Behavior Surveillance System (YRBSS), sponsored by the Centers for Disease Control and Prevention (CDC). The NCES compiles data from all these sources into a report called the *Indicators of School Crime and Safety* (e.g., Dinkes et al. 2007). When the estimates from these alternative sources are compared, there emerge some rather dramatic differences, leaving the investigator with the challenge of deciding where the truth lies.

Youthful Victimization in School and Out

Here we report crime victimization rates for school-aged youths, comparing, when possible, the rates at school and at other locations. We begin with murder, which is the only crime for which the statistics are reasonably accurate. Figure 1 depicts the trend in murders on school property for youths ages 5–18, compared with the overall murder count for that age group. There were about 30 school murders of youths each year from 1992–1993 to 1998–1999, a period notorious for the series of school rampage shootings that culminated with Columbine High School on April 20, 1999. During that event 12 students and a teacher were murdered and 23 students injured, before the shooters committed suicide. In the year following Columbine, the national in-school murder count dropped sharply and has remained relatively low since then. The overall murder rate for the same age group follows a similar pattern, though the decline began earlier and is less abrupt. The most important lesson from these data is that only about 1 in 100 murders of this age group occurs in school. That was true during the peak years of the early 1990s and also true a decade later. By this measure, then, school



School Crime Statistics, Fig. 1 Number of homicides involving young victims, in school and out, 1992–1993 to 2005–2006. Note: “In school” includes on school property, on the way to or from regular sessions at school, and while attending or traveling to or from a school-sponsored event (Source: Data on number of homicides in school are from School-Associated Violent Deaths Surveillance

Study (SAVD), tabulated in Indicators of School Crime and Safety (Dinkes et al. 2007, p. 68); and data on number of homicides total are from National Center for Injury Prevention and Control and Web-based Injury Statistics Query and Reporting System Fatal (WISQARS™ Fatal), retrieved Nov 2008 from <http://cdc.gov/ncipc/wisqars>)

appears much safer than other locations for school-aged youths.

However, schools have a much larger share of the nonfatal crimes with school-aged victims. Figure 2 depicts the trend for victimization rates of youths aged 12–18, including both theft and violence. Youths who have completed 12 years of school are excluded from this tabulation. The rates per 1,000 follow the trend for youth homicide (as well as the national trend for criminal victimization for all age groups) – a sustained and rather dramatic reduction, so that the 2005 figures are about one-third of the peak in 1993. For our immediate purpose here, the important thing to notice is that the victimization rate in school is about the same as out of school. That parity is the net result of theft, which has higher rates at school, and violence, which for most of the period has lower rates at school (although in-school and out-of-school rates of violence converged in 2004). Note that since youths spend over 80 % of their waking hours during a calendar year out

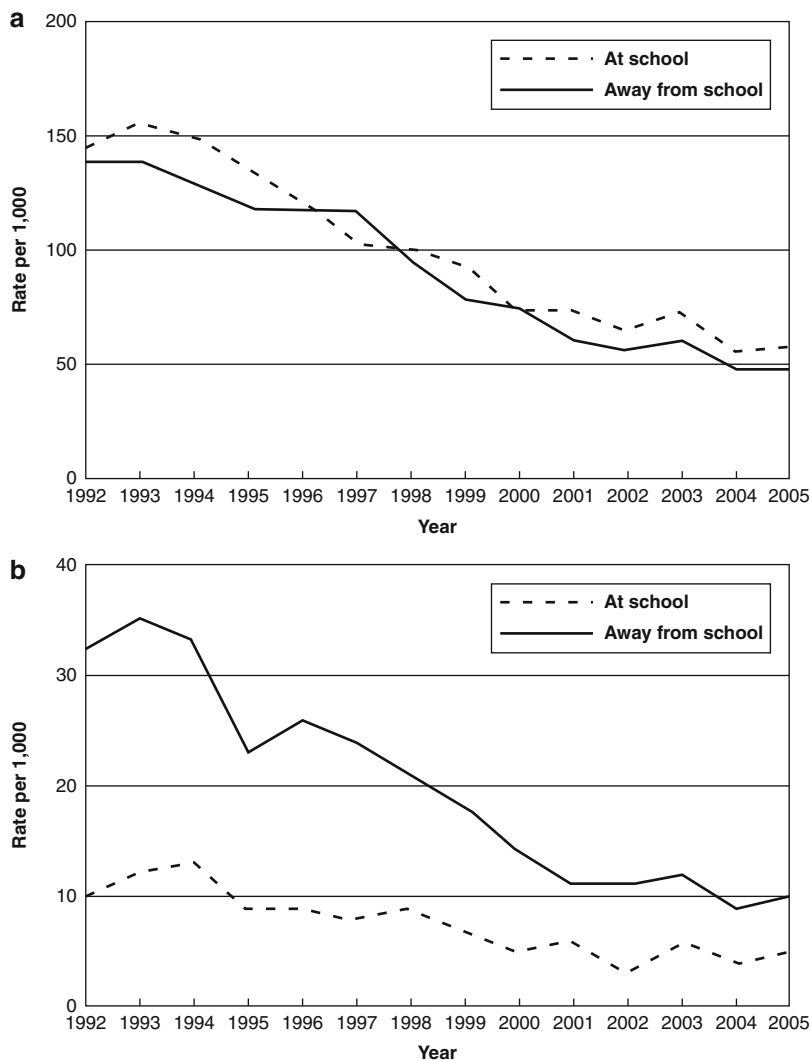
of school (Gottfredson 2001, p. 21), the parity in victimization rates implies that youths are far more likely to be victimized during an hour in school than an hour elsewhere.

For the serious violent crimes of rape, robbery, and aggravated assault, NCVS victimization rates are twice as high away from school as at school during recent years, as shown in Panel B. Since the corresponding ratio for murder is 100 to 1, we conclude that serious violent crimes committed out of school are far more likely to become murders than is true for similar crimes in school.

These NCVS results may well be misleading and should be validated. Youth survey data on crime are notoriously unreliable. In particular, crime survey results are exquisitely sensitive to the details of how the data are collected. One survey that provides an alternative to NCVS for estimating victimization rates is the Youth Risk Behavior Surveillance System (YRBSS), sponsored by the CDC. This survey yields estimates of victimization rates for serious violent crime

School Crime Statistics, Fig. 2 Panel A:

Victimization rates at school and out for youths ages 12–18, 1992–2005: theft and violence. *Panel B:* Victimization rates at school and out for youths ages 12–18, 1992–2005, serious crimes of violence. Note: Theft includes purse snatching, pickpocketing, and all attempted and completed thefts except motor vehicle thefts. Theft does not include robbery in which threat or use of force is involved. Violence includes serious crimes of violence and simple assault. Serious crimes of violence include rape, sexual assault, robbery, and aggravated assault. “At school” includes inside the school building, on school property, or on the way to or from school (Source: National Crime Victimization Survey (NCVS), tabulated in Indicators of School Crime and Safety (Dinkes et al. 2007, p. 70))



that are an order-of-magnitude higher than the NCVS rates. For example, in the 2005 YRBSS, 8 % of students in grades 9–12 reported being threatened or injured with a weapon on school property during the previous 12 months. That compares with the serious-violence victimization rate at school for 15–18-year-olds in the NCVS of 0.4 %. Thus, the YRBSS rate is 20 times as high, even though logic suggests that it should be *less*, given that the YRBSS refers to prevalence of victimization and the NCVS figure is overall incidence (so that multiple victimizations reported by the same respondent are included in computing the rate). Further, the NCVS category

of “serious violence” encompasses more types of crime than the YRBSS category of “threatened or injured with a weapon.”

What could account for this vast difference in results? First, the NCVS sample is interviewed every 6 months, and the previous interview serves as a bracket to help the respondent place events in time. Thus, the NCVS sample members are asked to report on events that occurred since the previous interview. The YRBSS, on the other hand, is a one-shot survey with no natural bracket on the time interval; respondents are asked to report on the previous 12 months, which creates the likelihood that some will report on serious

events that occurred outside the designated period (a phenomenon known as “telescoping”). A second important difference is that all YRBSS respondents are asked the specific question about whether they were threatened or injured with a weapon on school property, whereas the only NCVS respondents who are asked about such an incident are those who first respond affirmatively to a more general screener question. Third, the NCVS questionnaire is administered to the respondent (in person or over the telephone) at home, whereas the YRBSS questionnaire is self-administered by the respondent while in school. These and other differences, none of which are relevant in a literal sense, appear to be hugely important to the respondents’ answers in practice. Cook (1987) notes that the Safe Schools study (which had much in common with the YRBSS with respect to administration) estimated one million robberies in schools, compared with the estimate of 30,000 in the NCVS for the same period.

Given the disparate results from youth surveys, it is of interest to consider administrative data. The SSoCS gathers reports from public-school principals about crimes occurring during school hours. For the 2005–2006 school year, principals for middle and high schools reported a total of 928,000 violent crimes and 206,000 thefts (see Table 1). These counts are not precisely comparable to the NCVS results for 12–18-year-olds. Unlike the NCVS, the SSoCS is limited to public schools. The NCVS age range of 12–18 is roughly but not exactly comparable to the SSoCS category of “middle and high school.” Nevertheless, the estimates should be close. In fact, the violence reports are higher (by half) in the SSoCS than in the NCVS for 12–18-year-olds, while the SSoCS theft reports are much lower. It is not surprising that school officials do not know about many of the thefts that occur on school property, but the fact that they are aware of more violence that shows up in the NCVS defies ready explanation.

Thus, the truth about crime in school – or even a rough approximation of the truth – is elusive. Our inclination is to believe that the SSoCS reports provide a reliable lower bound for the

School Crime Statistics, Table 1 Comparison of SSoCS and NCVS crime counts

	SSoCS crime count: public middle and high schools 2005–2006 school year	NCVS crime count in school ages 12–18 2005
Violent crimes	928,000	628,000
Theft	206,000	868,000

Sources: <http://nces.ed.gov/pubs2007/2007361.pdf>; http://nces.ed.gov/programs/crimeindicators/crimeindicators2007/tables/table_02_1.asp?referrer=report

“true” volume of crime, understating the true total to the extent that officials are never made aware of some crimes, and may generally be inclined to underreport in order to make their schools look as safe as possible. If true, then the NCVS appears to provide a notable underestimate of the volume of violence in schools – but the difference is nothing like that suggested by the very high YRBSS results. We are inclined to believe that the NCVS data are superior to the YRBSS because the method of administration discourages exaggeration by respondents, and the bracketing provides some discipline on memory. We also note that the downward trend in NCVS rates (shown in Fig. 2) reproduces well-documented trends during that period for the entire US population and hence is credible. The YRBSS victimization rates, on the other hand, exhibit no such trend during this period, showing if anything an upward tilt since 1993. For those reasons we report additional NCVS results in what follows, even though we are willing to believe that these are also far off the mark.

Table 2 summarizes demographic patterns in victimization rates at school for youths aged 12–18. Note that these data exclude the responses of students who have already completed 12 years of schooling. They do not exclude school dropouts. The rates shown here are averaged over the three most recent years of the School Supplement of the NCVS. Note that “theft” and “violence” sum to the total – “serious violence” is included in “violence.”

Theft rates are remarkably uniform across all demographic categories, averaging 41/1,000.

School Crime Statistics, Table 2 At-school victimization rates/1,000 for youths age 12–18

	Total	Theft	Violence	Serious violence
Male	73	41	32	7
Female	61	41	21	4
Ages 12–14	75	42	33	6
Ages 15–18	61	40	21	5
Urban	75	41	34	9
Suburban	67	43	24	5
Rural	58	37	22	2
White	72	45	27	4
Black	64	38	27	5
Hispanic	55	29	26	5
Other	53	33	20	2
Overall	67	41	27	6

Source: NCVS results averaged for 2001, 2003, and 2005

Violence rates are a bit lower overall and more textured, although the differences among groups are still not as large as one might expect. Males are half again as likely to be victims of violence as females, and youths 12–14 are half again as likely as older youths. Urban schools experience a higher per capita rate of violent incidents than suburban or rural schools. Most surprising is that whites, blacks, and Hispanics report virtually the same rates of violence and serious violence.

The same NCVS data provide estimates for victimization rates away from school. The patterns are not much different, with two exceptions. First, blacks report a higher rate of serious violent crimes (17/1,000) than whites and Hispanics (both at 10/1,000). Second, and perhaps most intriguing, is that the age pattern away from school is the reverse of the age pattern at school. The younger group, aged 12–14, has somewhat higher victimization rates at school than the older group, but the older group has much higher victimization rates than the younger group away from school. The results are depicted in [Fig. 3](#). The explanation may in part be due to the fact that the older group includes a number of school dropouts who, since they are not attending school, are unlikely to be victimized on school property. Perhaps more important is that older youths have greater mobility and freedom outside of school and thus more opportunity to get into trouble.

In assessing the credibility of these results, we are inclined to believe the NCVS-based comparisons between in-school and out-of-school victimizations, since it is the same respondents and hence the same biases are likely to apply. The surprising homogeneity of theft and violence victimization between blacks, whites, and Hispanics can be checked against interschool patterns in the SSoCS. We do that below and find support for the theft results but an apparent contradiction with the results on violence.

Finally, we note the high prevalence of bullying in school. While not necessarily a crime, bullying can greatly color the school experience for some children. The NCVS School Crime Supplement found that in 2005, 28 % of youths ages 12–18 reported being bullied in school – of those, 79 % said they were bullied inside school, 28 % outside on school grounds, and 8 % on the school bus (Dinkes et al. 2007, p. 95).

Differences Among Schools

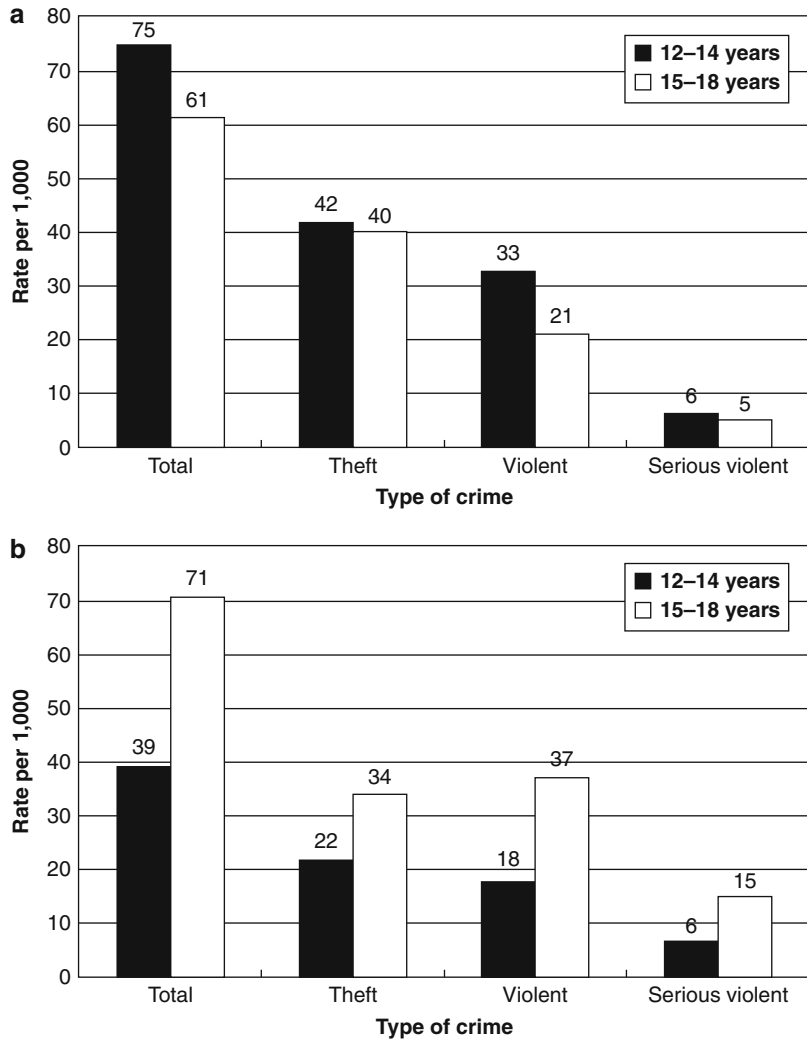
Up until this point, we have described crime patterns primarily with respect to the characteristics of the victims. From another perspective, school crime is a characteristic of the school, and there is strong evidence that school characteristics and policies influence crime victimization rates (Cook et al. 2010).

The 2005–2006 SSoCS classifies schools by grade level, enrollment size, urbanicity, and percent minority enrollment. The rate of violent incidents reported by principals is much higher for middle schools than either elementary or high schools, somewhat higher for city schools than those in suburban or rural communities, and higher in predominantly minority schools than those with less than half minority. Notably, there is little relationship between the size of the school and the violence victimization rate. The results for theft tend to be less patterned. [Table 3](#) summarizes the results.

In one respect these patterns are at odds with NCVS victimization patterns. It appears that the relatively high rate of violence in minority schools is at odds with the NCVS finding that there is little difference in victimization rates by race. One possibility is that black respondents are

School Crime Statistics, Fig. 3 Panel A:

Victimization rates at school for youths ages 12–14 and 15–18. **Panel B:** Victimization rates away from school for youths 12–14 and 15–18. Note: Total crimes include theft and violent crimes. Theft includes purse snatching, pickpocketing, and all attempted and completed thefts except motor vehicle thefts. Theft does not include robbery in which threat or use of force is involved. Violent crimes include serious violent crimes and simple assault. Serious violent crimes include rape, sexual assault, robbery, and aggravated assault. “At school” includes inside the school building, on school property, or on the way to or from school. NCVS results are averaged for 2001, 2003, and 2005 (Source: National Crime Victimization Survey (NCVS), tabulated in Indicators of School Crime and Safety (DeVoe et al. 2003, pp. 55–66; DeVoe et al. 2005, pp. 72–73; Dinkes et al. 2007, pp. 70–71))



less likely to report violent victimizations than white respondents in the NCVS or alternatively (as reported in Kinsler 2009) that school administrators in predominantly black schools are relatively likely to record violent crimes. The same source, SSoCS, reports information on gang-related crime. In 2005–2006, 11 % of middle schools and 16 % of high schools reported at least one crime that was gang related. Gang-related crimes were concentrated in large, urban, and predominantly minority schools (Nolle et al. 2007, Table 4). Some confirmation for these patterns comes from the NCVS School Supplement data. Students were asked about gangs in their schools. Affirmative responses

were much more likely by black and Hispanic students and by students in urban areas.

Another way to illustrate the variability across schools in crime levels is by comparing the distribution of the number of crimes reported per school with a distribution that would be generated under a random process. This type of exercise is common in criminological research. In a study of more than 23,000 boys born in two birth cohorts in Philadelphia, Tracy et al. (1990), for example, discovered that a small fraction of the boys (5–6 %) committed a majority of the delinquent acts. Similarly, researchers who study neighborhood crime have observed that it, too, is clustered at certain addresses. Weisburd et al. (2004),

School Crime Statistics, Table 3 Crime rates by school characteristic

	Violence rate/1,000 students	Theft rate/1,000 students
Level		
Primary	25.2	1.6
Middle	51.6	7.8
High school	25.7	8.7
Enrollment		
<300	34.5	4.3
300–499	34.0	3.3
500–999	30.9	4.5
1,000 or more	28.6	7.2
% minority enrollment		
<5 %	26.9	4.8
5–20	22.9	5.2
20–50	28.4	5.5
50 or more	39.9	4.8

Source: Nolle et al. (2007), extracted from Table 1

School Crime Statistics, Table 4 Number of “dangerous” and “persistently dangerous” schools, 2004, 2006, and 2008

	Dangerous schools	Persistently dangerous schools ^a
2003–2004	1,547 (1.9 %)	449
2005–2006	1,542 (1.9 %)	447
2007–2008	1,624 (2.0 %)	471

Source: Original tabulations from data files from the School Survey on Crime and Safety

^aAssumes 29 % of schools dangerous in 1 year are “persistently dangerous” (see text)

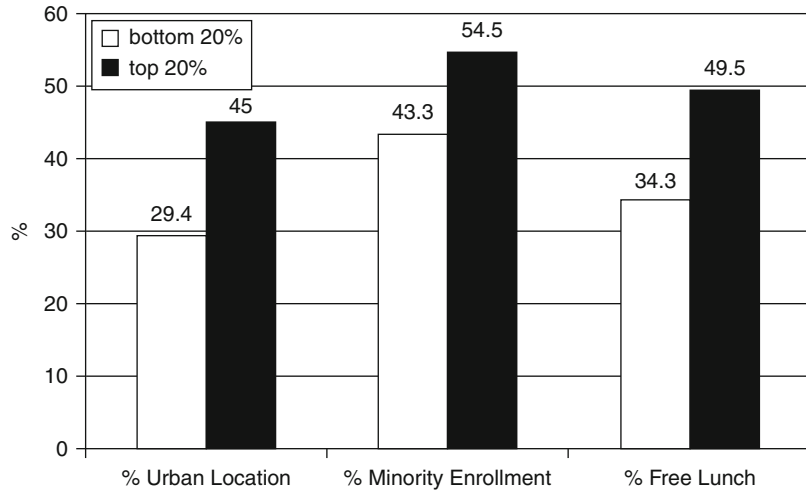
studying street segments in Seattle over a 14-year period, found that 4–5 % of the street segments account for about 50 % of incidents in each of the years examined. The distribution of crime at the school level is also concentrated, although not to the same degree.

We examined the distribution of school crime in the 2007–2008 SSocS data using schools with enrollments of at least 1,000 students (N = 8,843), because we assumed rank orders based on crime rates would be unstable for small schools. These analyses used sample weights provided by NCES and also weighted

by student enrollment. In these data, the 22.5 % of large schools with the highest *per capita* crime rates account for 50 % of all crimes reported by principals. The top 20 % of schools experience much higher crime rates than does the bottom 20 %: the victimization rates per 1,000 students for students in the most and least safe 20 % of schools are 9.8 and 111.3 for all crimes. The difference is even more striking for rates of violent crime – more than 13 times higher in the top 20 % of schools (4.9 vs. 65.7 per 1,000). These results suggest that although the average school is indeed a relatively safe place, a minority are notoriously unsafe. Figure 4 contrasts the safest 20 % of schools with the most dangerous 20 %. The more dangerous schools are much more likely to be urban schools, schools serving a high percentage of minority students, and schools serving more disadvantaged student populations.

The NCLB Act requires states to establish a mechanism for identifying such unsafe schools. Once a school has been labeled through this process as “persistently dangerous,” parents must be given the option of transferring their child to a safer school. Every state has responded to this requirement by defining “persistently dangerous” and establishing a procedure for identifying such schools. The definitions differ considerably from state to state according to the specific offenses considered to be indicators of dangerousness; whether or not an official response (e.g., expulsion, arrest, conviction) to the incident is required before it is considered a dangerous offense and whether or not a dangerous offense occurring in or around the school has to have been committed by a student; the cut-point for the number or rate of offenses above which a school is to be considered dangerous; and the number of years in the dangerous status required to be considered “persistently” dangerous. Also, some states use a multistage process for identifying unsafe schools. For example, Florida’s policy calls for a first stage in which a school is considered potentially dangerous if, for 3 consecutive years, it (a) has a federal Gun-Free School Act violation and (b) expels 1 % or more of a student body that is greater than 500 students, or 5

School Crime Statistics, Fig. 4 Comparison of the characteristics of the schools with the lowest and the highest 20 % of crime rate. Based on schools with enrollment of at least 1,000 students. Weighted by student enrollment and sample weight (Source: Original tabulation from data files. School Survey on Crime and Safety (SSoCS) (2007–2008))



students if the student body is 500 students or less, for homicide, battery, sexual battery, or weapons possession-related offenses. Any school so identified must, in a second stage, conduct an anonymous school-wide survey of students, parents, and school personnel. Only schools in which 51 % or more of survey respondents judge the school to be unsafe are labeled persistently dangerous (Education Commission of the States 2004).

Beginning with the 2003–2004 school year, the number of persistently dangerous schools reported by all states combined to the US Department of Education has ranged from 36 to 49, averaging 44. That low count raises the question of just how effective is the NCLB legislation for encouraging states to identify unsafe schools. We used the SSoCS survey data to identify schools that might reasonably be considered persistently dangerous. As noted above, there is no standard definition to guide this effort. Almost all states, however, include incidents involving serious violent behavior and weapons-related offenses, and many states also include incidents involving illegal drug use, possession, and distribution. Almost all states require that the rate of incidents per 100 enrolled students be between 1 and 3 before a school will be considered dangerous. Most states require that schools remain in this status for 3 years, but some require only two. We therefore classified schools in the SSoCS sample as

dangerous during the 2007–2008 school year if, according to their principals’ reports, the number of serious violent offenses (e.g., rape, sexual battery other than rape, physical attack or fight with a weapon, threat of physical attack with a weapon, and robbery with or without a weapon) plus the number of weapons-related incidents (those involving firearms, explosive devises, knives, or sharp objects) exceeded 2 per 100 enrolled students. We repeated this analysis adding the number of incidents involving illegal drugs (distribution, use, or possession) to the classification of dangerous. Doing so increases the number of schools identified as dangerous. Our conclusion about the under-identification of unsafe schools using the NCLB procedure would therefore be stronger if these drug offenses were included.

Although the SSoCS survey is cross-sectional by design, it is administered to a random sample of US school every 2 years. The sample design is stratified and over-samples middle and high schools. Some schools happen to be included in multiple years just by chance. By merging the 2007–2008 survey data with data from the two prior SSoCS surveys (2003–2004 and 2005–2006), we created a longitudinal sample containing 475 schools. This longitudinal sample overrepresents secondary schools, large schools, and schools in areas that are not located in rural areas but is nevertheless helpful for examining

the extent to which dangerousness persists across time.

Table 4 shows that applying our definition of dangerousness to the entire sample of schools in each of the 3 years resulted in identifying between 1.9 % and 2.0 % of the nation's schools in each survey year. The associated point estimates for the numbers of dangerous schools identified were 1,547 (2003–2004), 1,542 (2005–2006), and 1,624 (2007–2008). Using the longitudinal sample, we determined that 29 % of schools identified as dangerous in 1 year were also identified as such 2 years earlier. This exercise suggests that between 447 and 471 schools might be considered persistently dangerous in any given year. That estimate is about ten times as high as the average number of persistently dangerous schools identified to the US Department of Education each year, just 44.

Concluding Thoughts

There are a variety of sources of statistics on crime in schools, which are not entirely consistent with respect to levels, patterns, and trends. Anyone wishing to make sense of the available statistics should first become informed on the details of how the data are generated and consider the likely biases and as much as possible compare different sources.

We believe that the homicide statistics are accurate but that other police data on school crime are not to be trusted. For nonfatal crimes, we place some credence in the NCVS for students, which is a recurrent survey implemented by the US Census Bureau. What one learns from this source is that crime victimization in schools for students followed the downward trend in national crime rates during the 1990s and remains at a relatively low level since 2000. That there would be a common trend that makes sense and is one illustration of a more general result that crime in schools is closely linked to crime in the community.

Another credible result is that there is a great deal of crime in schools perpetrated by and against students. However, the estimated rates

and patterns differ widely among the various sources in common use. According to the NCVS, in-school victimization rates are similar to rates experienced outside of school. But we found order-of-magnitude differences in school crime rates from NCVS and YRBSS, CDC's recurrent youth survey in which youths self-report their own victimizations. The SSoCS, based on reports from school principals, also produces counts of violent crimes that are 50 % higher than the NCVS counts. We would nominate SSoCS as the most reliable source for violent crime, but there is no basis for comparison with out-of-school crimes. We can be confident that homicide is very rare in school (relatively and absolutely). All sources confirm that a much higher percentage of minor assaults occur in schools than serious assaults. It also seems reasonable to conclude that school crime is far from randomly distributed across schools. A relatively small number of schools experience inordinately high rates of theft and violence. Rates of crime in the least safe 20 % of schools are more than ten times higher than the rates of crime in the safest 20 %. Students in these schools are likely to avoid school for fear of their own safety, teacher turnover is likely to be high, and the quality of time dedicated to learning is likely to be diminished. Unfortunately, no reliable mechanism has been established to identify which of these schools should be identified as "persistently dangerous." The consequences for being labeled "persistently dangerous" under NCLB are sufficiently punitive to create a strong incentive for principals to underreport and for state boards of education to establish policies that minimize application of that label.

What should be done to advance knowledge about school crime and to provide a reliable mechanism for identifying schools requiring assistance? The place to begin the research agenda on crime in schools is with a close look at the quality of the data in current use. Despite the many problems we identified earlier, it is too often true that users do not investigate the quality of the data or check one source against another. It would be a useful service to all users if there were a comprehensive investigation of the differences

in crime rates and patterns across these data sets, together with an investigation of the sources of disagreement. One thing that is clear is that survey results with adolescent subjects are exquisitely sensitive to where and how the questions are administered.

Beyond acquiring a better understanding of the sources of discrepancy across different methods for obtaining school crime data, a serious effort is needed to develop shared understandings of what constitutes a dangerous school. Rather than asking each state to develop its own idiosyncratic definition and identification procedure, we suggest that a higher level dialogue is needed to understand the cut-point above which the school environment is truly dangerous to students and staff or seriously hinders learning. A discussion about how school safety might be monitored more systematically is needed. We suggest that rather than counting the number of certain incidents or responses to certain offenses, which are likely to be extremely rare and confounded with characteristics of school's surveillance and reporting practices, a more direct measure of school safety might come from student and staff surveys of perceptions of safety and the extent to which they avoid the school or certain areas in the school for fear of their own safety. Such survey measures have been a staple of school climate research since the 1970s and have proven to be highly reliable and stable measures of the school environment. At least one state (Florida) already uses such a strategy to identify persistently dangerous schools.

We also suggest that once a reliable mechanism is in place for identifying schools with problematic crime levels, efforts should be directed towards making the schools safer. Rigorous research (Gottfredson et al. 2002; Hahn et al. 2007; Wilson and Lipsey 2007) has identified a variety of school-based prevention and intervention practices that are likely to reduce the level of crime experienced in these schools.

Notes and Acknowledgments This entry is based on a more detailed chapter (Cook et al. 2010) on school crime control and prevention.

Related Entries

- ▶ [Crime](#)
- ▶ [School Social Organization, Discipline Management, and Crime](#)
- ▶ [School Structural Characteristics and Crime](#)
- ▶ [School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis](#)

References

- Centers for Disease Control and Prevention (2008) Youth Risk Behavior Surveillance – United States, 2007 MMWR
- Cook PJ (1987) Robbery violence. *J Crim Law Criminology* 70:357–376
- Cook PJ, Gottfredson DC, Na C (2010) School crime control and prevention. In: Michael T (ed) *Crime and justice: a review of research*, vol. 39. University of Chicago Press, Chicago, Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1368292
- DeVoe JF, Peter K, Kaufman P, Ruddy SA, Miller AK, Planty M, Snyder TD, Rand MR (2003) Indicators of school crime and safety: 2003. US Dept. of Justice, Office of Justice Programs, Bureau of Justice Statistics NCJ 210257.
- DeVoe JF, Peter K, Noonan M, Snyder TD, Baum K (2005) Indicators of school crime and safety: 2005. US Dept. of Justice, Office of Justice Programs, Bureau of Justice Statistics NCJ 210697.
- Dinkes R, Cataldi EF, Lin-Kelly W (2007) Indicators of school crime and safety (NCES 2008-021/NCJ 219553). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice, Washington, DC
- Education Commission of the States (2004) Persistently dangerous school criteria, Updated September 2004 <http://www.ecs.org/clearinghouse/52/98/5298.pdf> (downloaded 25 Sept. 2009)
- Gottfredson DC (2001) *Schools and delinquency*. Cambridge University Press, New York
- Gottfredson DC, Wilson DB, Najaka SS (2002) School-based crime prevention. In: Sherman LW, Farrington DP, Brandon WC, MacKenzie DL (eds) *Evidence-based crime prevention*. Routledge, London
- Guerino P, Hurwitz MD, Noonan ME, Kaffenberger SM (2006) *Crime, violence, discipline, and safety in U.S. Public schools: findings from the school survey on crime and safety: 2003–04* (NCES 2007-302rev). U.S. Department of Education, National Center for Education Statistics, Washington, DC
- Hahn RA, Fuqua-Whitley DS, Lowy J et al (2007) A review of the effectiveness of universal

school-based programs for the prevention of violence.

Am J Prev Med 33:S114–S129

Hernandez N (2007) No child's data on violence skewed: each state defines 'dangerous school' (November 18, p. A01). *Washington Post*. 77, 67–74

Jacob BA, Lefgren L (2003) Are idle hands the devil's workshop? Incapacitation, concentration, and juvenile crime. *Am Econ Rev* 93:1560–1577

Jekielek S, Brown B, Marin P, Lippman L (2007) Public school practices for violence prevention and reduction: 2003–04 (NCES-2007-010). U.S. Department of Education, National Center for Education Statistics, Washington, DC

Kinsler J (2009) School discipline: a source or salve for the racial achievement gap? Unpublished manuscript, University of Rochester, New York

Nolle KL, Guerino P, Dinkes R (2007) Crime, violence, discipline, and safety in US public schools: findings from the school survey on crime and safety: 2005-6 (NCES 2007-361). U.S. Department of Education, National Center for Education Statistics, Washington, DC

Tracy PE, Wolfgang ME, Figlio RM (1990) *Delinquency careers in two birth cohorts*. Plenum Press, New York

Weisburd D, Bushway S, Lum C, Yang SM (2004) Trajectories of crime at places: a longitudinal study of street segments in the city of Seattle. *Criminology* 42(2):283–321

Whitaker C, Bastian LD (1991) Teenage victims: a national crime survey report (NCJ-128129). U.S. Department of Justice, Bureau of Justice Statistics, Washington, DC

Wilson SJ, Lipsey MW (2007) School-based interventions for aggressive and disruptive behavior: update of a meta-analysis. *Am J Prev Med* 33:S130–S143

School Resource Officers

► [Police School Services](#)

School Resource Officers (SROs)

► [Police Officers in Schools](#)

School Security

► [School Security Practices and Crime](#)

School Security Practices and Crime

Allison Ann Payne and Ronald J. Eckert
Department of Sociology and Criminal Justice,
Villanova University, Villanova, PA, USA

Synonyms

[School crime](#); [School discipline](#); [School security](#)

School Discipline

► [School Security Practices and Crime](#)

School Police

► [Police School Services](#)

School Programs

► [Bullying Prevention: Assessing Existing Meta-Evaluations](#)

Overview

The crime control model of school security has led to an environment of school prisonization and student criminalization. The increasing use of restrictive security practices contrasts with both the decline in school crime and violence as well as the research suggesting that these practices may not be effective. Additionally, this intensification runs the risk of invading students' privacy and creating a negative school environment. Much research suggests that it would be better for schools to focus on other evidence-based strategies to reduce school disorder and respond to student misbehavior. If schools continue to use restrictive security practices that mirror the criminal justice system,

they may, ironically, ensure that certain students are found in that system in the future.

The Prisonization of School Security

It is abundantly clear that most modern schools, particularly urban public schools, utilize a crime control model when defining and managing student behavior. This has occurred largely in response to school shootings and other visible instances of student violence (Addington 2009). More and more schools have implemented prison-like practices in an effort to improve their security (Giroux 2003), resulting in the use of many restrictive security practices and procedures intended to deter school crime and ensure the safety of students and staff. Indeed, recent statistics indicate that at least 55 % of all schools, and nearly all urban schools, now have security and surveillance programs (Devoe et al. 2005; Gottfredson and Gottfredson 2001).

Examples of restrictive security practices seen in many schools range from requiring formal dress codes (Gottfredson and Gottfredson 2001) to installing metal detectors (Brooks et al. 2000). To further detect the presence of weapons and other contraband, schools may perform regular locker searches, require students to carry clear book bags, and use drug-sniffing dogs (Brooks et al. 2000; DeVoe et al. 2005; Gottfredson and Gottfredson 2001). Visitors to campuses are generally required to sign in before entering school buildings and student identification badges are often mandated to facilitate immediate identification of rule-breakers and to deter defiance and delinquency (Brooks et al. 2000). Hallways are often supervised by school staff and administrators (Devoe et al. 2005) or even by uniformed security guards or uniformed and armed security resource officers (SROs; Giroux 2003). Additionally, many schools have chosen to install security cameras that provide constant surveillance to most areas within a school (Devoe et al. 2005). Most institutions have locked or monitored doors and gates to prevent

unauthorized individuals from entering school grounds and to prohibit students from leaving campus (Devoe et al. 2005; Gottfredson and Gottfredson 2001). Schools are even investigating the use of cutting-edge technology to enhance security, such as iris recognition software, webcams, and radio frequency identification tags on ID badges (Addington 2009). Ironically, even with the increased implementation of these prison-like practices, parents and school boards have continued to call for stricter measures of control (Brooks et al. 2000) to manage the fear and anxiety surrounding school.

The intensification of school security has also led to an increase in the criminalization of students, such that schools often manage and punish student behavior in a way that is analogous to the treatment of adult criminals (Giroux 2003; Kupchik and Monahan 2006; Tredway et al. 2007). For example, the actions of students who violate school rules are often described with criminal justice language (Tredway et al. 2007): “suspects” or “repeat offenders” are subjected to “investigations,” “interrogations,” and “searches” by dogs or SROs and may then be involved in “lineups” and school “courts.” The students are then punished in ways similar to the sentences received by adult criminals; zero tolerance policies, for instance, function as the school equivalent of mandatory minimum criminal sentencing statutes. These policies have increased the use of harsh discipline, such as student exclusion through expulsion and suspension, even though these disciplinary responses have various negative impacts on both students and schools (Welch and Payne 2011). Excluded students are more likely to experience school failure, drop out of school, and engage in delinquency and drug use both in and out of school. Unfortunately, just as the use of more restrictive school security measures has increased, so has the use of these and other harsh disciplinary techniques.

Although it could be assumed that the intensification of school security is a result of increased school crime and violence, evidence suggests this is not the case (Welch and Payne 2011). In fact,

these changes have occurred despite a documented decline in student delinquency and drug use, violent victimization in schools, and school-related deaths (Brooks et al. 2000; Devoe et al. 2005). Other possible explanations for this intensification include popular anxiety about high-profile instances of school violence, termed the “Columbine effect” (Addington 2009); increased school accountability for the academic performance of students (Hirschfield 2008); concerns about possible litigation in response to violent incidents on campus (Hirschfield 2008); and responses to the growing proportion of minority students (Welch and Payne 2011). While many of these explanations fall short in certain ways (Hirschfield 2008), it is likely that a combination of these factors have allowed this intensification to continue, leading schools to become less oriented around education and more like prisons focused on punishment (Giroux 2003).

Trends in School Security

The use of restrictive security practices in schools has been increasing over the past decade. Using data from the School Survey on Crime and Safety, Payne and Eckert (2011) discussed the percentage of public school principals who reported using various security measures during the years 1996, 1999, 2003, 2005, and 2007 (Table 1). The School Survey on Crime and Safety is a national questionnaire administered to public school principals by the National Center for Education Statistics. Approximately 3,500 principals from schools at all levels are asked questions on topics such as school disorder and crime, safety and discipline, and prevention programs and policies. Between the years 1996 and 2007, the imposition of school uniforms increased from 3 % to 17.5 %, drug testing of certain students increased from 4.1 % to 6.4 %, and the use of identification badges for students increased from 3.9 % to 7.6 %. Other measures show even more dramatic increases, such as the use of identification badges for faculty (25.4–58.3 %), video

School Security Practices and Crime, Table 1 Percentage of schools reporting use of security measures, School Survey on Crime and Safety

Security measure	1996	1999	2003	2005	2007
School uniforms	3.0	11.8	13.8	13.8	17.5
Drug testing of students	–	4.1	5.3	5.0	6.4
Badges or IDs for students	–	3.9	6.4	6.1	7.6
Badges or IDs for faculty	–	25.4	48.0	47.8	58.3
Video surveillance	–	19.4	36.0	42.8	55.0
Controlled access to school grounds	24.0	33.7	36.2	41.1	42.6
Controlled access to buildings	53.0	74.6	83.0	84.9	89.5
Drug sweeps	19.0	21.0	21.3	23.0	21.5
Random sweeps for contraband	–	11.8	12.8	13.1	11.4
Random metal detector checks	4.0	7.2	5.6	4.9	5.3
Pass through metal detectors daily	1.0	0.9	1.1	1.1	–
Closed during lunch	80.0	90.0	66.0	66.1	65.0
Clear book bags or banning of bags	–	5.9	6.2	6.4	6.0
Visitor sign in	96.0	96.6	98.3	97.6	98.7

Source: Payne and Eckert 2011

surveillance (19.4–55 %), and controlled access to school grounds (19.4–55 %). Finally, Table 1 shows that more and more schools controlled access to their buildings over this time, from 53 % in 1996 to 89.5 % in 2007.

Similar trends are seen when Payne and Eckert (2011) analyzed student reports of security measures from the National Crime Victimization Survey: School Crime Supplement during the school years of 1999, 2001, 2003, 2005, and 2007 (Table 2). Designed by the National Center for Education Statistics and the Bureau of Justice Statistics, the School Crime Supplement is an occasional addendum to the annual NVCS that gathers specific information from individual students regarding school-related disorder and victimization on a national level. Data measuring students’ perceptions of safety and crime at school are collected from approximately 6,500 12-through 18-year-old students attending public and private schools.

School Security Practices and Crime, Table 2 Percentage of students reporting use of security measures, School Crime Supplement

Security measure	1999	2001	2003	2005	2007
Security cameras	–	38.5	47.9	57.6	66.0
Security guards	54.1	63.6	69.6	67.9	68.8
Locked doors during day	38.1	48.8	52.8	54.2	60.9
Metal detectors	9.0	8.7	10.1	10.7	10.1
Badges and picture ID	–	21.2	22.5	24.7	24.3
Locker checks	53.3	53.5	53.0	52.9	53.6
Staff/adult supervision	85.4	88.3	90.6	89.8	90.0
Code of student conduct	–	95.1	95.3	95.1	95.9
Visitor sign in	87.1	90.2	91.7	92.7	94.3

Source: Payne and Eckert 2011

Of particular note is the increase of various measures between 1999 and 2001: A clear upward shift was reported in locked entrances and exit doors (38.1 % in 1999 to 48.8 % in 2001) and school security guards or law enforcement personnel (54.1 % in 1999 to 63.6 % in 2001). This trend continued through 2007, with the biggest increases in security in the areas of locked doors (38.1 % in 1999 to 60.9 % in 2007), security guards or law enforcement personnel (54.1 % in 1999 to 68.8 % in 2007), and surveillance cameras (38.5 % in 2001 to 66 % in 2007).

When schools are grouped by characteristics such as level, location, and size, differences in the type and amount of school security measures are seen (Payne and Eckert 2011). For instance, elementary schools are generally the least likely to implement restrictive security practices, followed by middle schools then high schools. One exception, however, is controlled access to buildings and school grounds, which is more often seen in elementary schools. In addition, larger schools are more likely than smaller schools to use most enhanced security measures, as are urban schools when compared with schools in towns or rural areas. Finally, schools with a greater percentage of minority students and students who receive free or reduced-price lunches are more likely to implement these practices (Payne and Eckert 2011).

The Effectiveness of Security Measures

Very little research exists that evaluates the effectiveness of school security measures. The pervasiveness of these measures combined with the lack of knowledge regarding their impact is of concern; if these practices are ineffective, they could allow schools to feel secure when they are not and may even create a dangerous environment by diverting money and resources from measures that do actually work (Addington 2009).

Several studies that have been conducted examine the effectiveness of security measures based on perceptions of specific practices rather than actual impact. Overall, school community members perceive particular measures as positive and believe that they prevent crime and disorder. For instance, both students and staff view SROs as effective; similar beliefs are held by school administrators regarding security cameras. Unfortunately, there is little evidence to support these perceptions (for a full review, see Addington 2009).

The small amount of evaluation research that has examined more objective outcomes related to the use of security measures is inconclusive. An early study of school security examined the impact of metal detectors in three program and 12 comparison high schools in New York City (Ginsberg and Loffredo 1993). Surveys were completed by 13,999 students in June 1992. Students in schools with metal detectors were less likely to carry weapons inside the school building or to and from school. However, there were no student differences in weapon carrying in other locations nor in threats or physical fights anywhere (Ginsberg and Loffredo 1993).

Mayer and Leone (1999) examined the relationships among school security, discipline management, student self-protection, and school disorder. Data were taken from 6,947 public school students in grades 7–12 who were interviewed as part of the 1995 School Crime Supplement. The structural equation model that was estimated contained four constructs: secure building, which included security measures such as guards, visitor sign in, metal detectors, and

locked doors; system of law, which reflected students' perceptions of rules and rule enforcement; school disorder, which included gang presence, drug availability, and personal attack and theft; and individual self-protection, which contained student avoidance of school location and fear of attack. Results suggested that schools which used physical security measures, such as metal detectors and locked doors, experienced more crime and disorder, as did schools that used personnel-based measures, such as security guards and hall monitors. In contrast, schools that focused on communication of schools rules and consequences for rule infractions experienced less disorder.

Finally, Chen (2008) analyzed a model of school crime that included two forms of physical school security: the number of ways a school controls access to campus and the number of ways in which a school monitors student activities. Both measures were hypothesized to be negatively related to the number of crimes that had occurred in the school in the past 12 months. Other factors in the model included urbanicity, community crime, student socioeconomic status, school size, student mobility, student misbehavior, and serious disciplinary penalties such as transfers and suspensions. Data were taken from 712 secondary schools whose principals had participated in the 2000 School Survey on Crime and Safety. Bivariate correlations showed that the number of ways a school controls access to their buildings was positively correlated with the number of crimes, while the number of ways a school monitors student activities was not significantly related to crimes. When the full model was estimated using structural equation modeling, the path between school security (a latent variable containing both building access control and student activity monitoring) and number of criminal incidents was negative but nonsignificant, contrary to the study's hypothesis. These findings suggest that, while school security measures may not increase crime as reported by Mayer and Leone (1999), they do not appear to reduce crime either (Chen 2008).

Ultimately, there is little evidence supporting the effectiveness of security measures. In addition, the few studies that do exist suffer from severe limitations. First, several of the studies measure effectiveness through perceptions of school community members, rather than actual impact on school disorder (see Addington 2009). In addition, all of these studies rely on cross-sectional data with no baseline measures of school disorder, thus making it impossible to truly establish the temporal order of the implementation of security measures and the level of school disorder (Ginsberg and Loffredo 1993; Mayer and Leone 1999; Chen 2008).

Student Civil Liberties

An unintended consequence of enhanced school security that has generated much discussion is infringement on student civil liberties (Addington 2009; Berger 2003), which can be separated into two related concerns: suspicionless searches and privacy encroachments. Searches of students' persons, lockers, and belongings have been the subject of many court cases regarding the fourth Amendment, including several Supreme Court rulings. One landmark case, *New Jersey v. TLO* (1985), established that searches conducted by school administrators require a far lower standard of suspicion than police searches of citizens on the street, because the intrusion on students' privacy is deemed minor and is overshadowed by public health and safety concerns (Berger 2003). The rationale for this standard is the "special needs" doctrine that emerged from this case, stating that a special need exists such that schools must maintain discipline in order to have an environment conducive to learning. Therefore, it is not necessary for administrators to have probable cause before searching students and their belongings (Berger 2003), including the use of metal detectors and locker searches. Essentially, although students have the right to privacy, this right is counterbalanced by the "special need for an immediate response to behavior that threatens either the safety of schoolchildren and teachers

or the educational process itself” (New Jersey v. TLO 1985, p.353). Another landmark case, *Vernonia School District 47J v. Acton* (1995), applied this doctrine to drug testing of student athletes; later, this was expanded to allow schools to drug test all students participating in any extracurricular activity (Board of Education v. Earls 2002). Thus, although students do not waive their fourth Amendment rights in school, the special needs doctrine offers school officials a large amount of discretion and flexibility in order to maintain a safe environment (Berger 2003).

Although these Supreme Court rulings applied the special needs doctrine to searches by school administrators for contraband in order to reduce or prevent drug use, lower court rulings have expanded the application to include searches by law enforcement personnel and security measures designed to prevent violence (Addington 2009; Berger 2003). Originally, searches allowed through *New Jersey v. TLO* and similar cases were limited to those conducted by school personnel rather than police officers, based on the rationale that the mission of school officials is to educate rather than collect evidence for a criminal prosecution. However, public anxiety over school crime and violence has led to a new view of student searches by police officers as acceptable under the special needs doctrine because these searches are seen as minimal, nonintrusive, and within the realm of reasonable suspicion (Berger 2003). In addition, decisions by lower appellate courts have also expanded this power to security searches designed to prevent violence, such as metal detector screenings (Addington 2009).

There is also a sense of these security measures infringing on students’ privacy, even beyond the legal exploration of suspicionless searches (Addington 2009). This can be seen particularly with tactics that monitor student behavior in public areas, such as security cameras and the presence of SROS, which are also the measures that tend to be implemented most often. Concerns appear when these measures are used in ways for which they are not originally intended (Addington 2009). For example, if

security cameras are originally installed for violence prevention, is it legitimate to then use them to thwart vandalism? Additionally, the use of cutting-edge technology to enhance security, such as webcams and radio frequency identification tags on ID badges, has the potential for even greater privacy infringement (Addington 2009).

The Impact on School Community

This invasion on student privacy and civil liberties may lead to another consequence of the intensification of school security: the altering of the school environment. Much research (summarized in the “School Social Organization, Discipline Management, and Crime” entry in this encyclopedia) has illustrated the importance of school climate and social organization. Schools with healthy supportive environments generally have more effective teachers who enjoy their jobs more and have more positive perceptions of the school administration (Gottfredson et al. 2005). In addition, students who attend these types of schools are more attached to teachers, more committed to school, and have stronger belief in school norms (Payne et al. 2003). This positive school community and subsequent student bonding, in turn, leads to a beneficial learning environment and less school crime and disorder (Gottfredson et al. 2005; Payne et al. 2003).

But what happens to the school community and subsequent student behavior when restrictive security practices are implemented? Restrictive measures may first lead to an oppositional relationship between students and staff, as students could protest these practices with petition drives and class boycotts (Berger 2003). Even without this type of adversarial response, aggressive security measures may negatively impact the school environment by interfering with the educational process, disrupting the learning environment, and wasting valuable class time (Berger 2003). For example, Devine (1996) described how teachers in a New York City school could not teach due to “loud noises

coming from the walkie-talkies in the corridors” supervised by security staff (p. 89), while Glazer (1992) detailed how administrators in another New York City school took close to 3 h to “funnel all 3,000 students into the gym, where they are frisked with hand-held metal detectors and their book bags are probed” (p. 790).

In addition, the implementation of many of these intrusive security practices often creates a climate of fear and resentment, resulting in negative attitudes toward school (Addington 2009; Hyman and Peronne 1998). For example, (Hyman and Peronne 1998) detail how searches result in detrimental consequences for both students and staff, such as lower student morale, distrust for staff, and alienation for law enforcement authorities; other studies have shown similar student alienation and mistrust as a consequence of restrictive security measures (Noguera 1995). Metal detectors, security patrols, and lock-down drills make schools feel like prisons (Noguera 1995) and often make students more afraid rather than less (Devine 1996). As Devine (1996) describes, aggressive security measures in urban schools result in “a climate of fear that indoctrinates youth into a culture of violence and dictates that only those exhibiting a ‘tough’ demeanor will survive” (p.179). As fear increases, so does student resentment and hostility (Hyman and Peronne 1998), leading to lower student bonding and a negative school environment (Addington 2009). Thus, it is possible that intrusive security might have the opposite effect than intended, leading students to increasingly break school rules. Ultimately, there is a strong chance that the intensification of school security leads to students feeling as though they do not belong in their schools, which, in turn, may decrease learning and increase the likelihood of these students engaging in deviant and delinquent behavior.

The Future of School Security

Beyond a doubt, far more “evidence-based research is...needed to evaluate the costs and effectiveness of school security measures”

(Berger 2003, p. 351). The lack of studies clearly shows that evaluations must be conducted to examine the impact of these practices on the levels of misbehavior and violence in schools. Further, the little research that has been done is severely limited in several ways. First, several studies examine perceptions of school community members as an outcome, rather than actual effectiveness on school crime and violence (for a full review, see Addington 2009). In addition, the few studies that do examine the impact of these measures on school disorder rely on cross-sectional data with no baseline measures of crime and deviance (Ginsberg and Loffredo 1993; Mayer and Leone 1999; Chen 2008). This makes it impossible to truly establish the temporal order of the relationship between restrictive school security and school disorder; that is, it is possible that these practices were implemented in schools that were already displaying high levels of crime, rather than leading to more crime after implementation, as suggested by Mayer and Leone (1999). This lack of research is surprising given the large amount of evidence supporting the use of other policies and programs designed to prevent student misbehavior (see “School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis” entry in this encyclopedia). Future studies should use longitudinal data and include baseline measures of crime to examine the impact of restrictive measures on school disorder.

In addition, the financial costs of these practices need to be studied. Little data exist on the cumulative cost to schools of metal detectors, SROs, security cameras, and other such measures (Addington 2009). Along with a gathering of costs, cost-benefit analyses should be conducted to evaluate whether the financial price is worth the effectiveness. Included in this consideration should be the examination of the possibility of “budgetary trade-offs” that may happen in order to pay for the enhanced security; it is possible that the cost of these measures leads to fewer resources for actual learning, such as books and staff (Addington 2009, p. 1440).

Even if they are found to be both effective and cost-effective, the use of these restrictive security measures should be balanced against the costs to student civil liberties and the school environment. Although more research is needed, studies do suggest that these measures infringe upon student civil rights (Berger 2003). As Addington (2009) discusses, it appears as though no one is protecting the students' rights, as no one with power is questioning the impact of enhanced security measures on student privacy and civil liberties. Students may want to change policies but they do not have the ability to do so, while most parents believe this security keeps their children safe and school officials are likely to give in to the demand for strong school security. With the possibility of technology, such as radio frequency identification tags, enhancing security even further, it is imperative that the impact of these measures on students' rights is considered. While student safety is clearly important, school officials should work toward developing and implementing strategies that keep student safe without infringing on their fourth Amendment rights (Berger 2003).

It is also possible that enhanced security measures create a negative learning environment, filled with fear and hostility. The few studies that have been conducted suggest that implementing restrictive measures such as metal detectors, locker checks, and security personnel may lead to an "unwelcome, almost jail-like, heavily scrutinized environment [that] may foster the violence and disorder school administrators hope to avoid" (Mayer and Leone 1999, p. 349; Chen 2008). The findings suggest, instead, that more attention should be paid to communicating school rules and norms and fairly and consistently enforcing these rules with appropriate consequences (Mayer and Leone 1999), a practice that is well supported by previous research (Berger 2003; Gottfredson et al. 2005).

This leads to a broader discussion on what schools can do to effectively reduce disorder and violence. Fortunately, a large body of research exists that establishes effective

strategies. These practices are discussed in two related encyclopedia entries: "School Social Organization, Discipline Management, and Crime" and "School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis."

Conclusion

As the crime control model continues to guide security, "schools grow more like prisons than institutions of education" (Giroux 2003: 553). Reliance on this model has intensified security practices to the point of creating a picture of school prisonization and student criminalization, despite a clear decline in school crime and violence. The little research that has been conducted suggests that, although school officials appear to be pleased with enhanced security, these practices may not be effective for reducing student misbehavior. In addition, this intensification runs the strong risk of invading students' privacy and creating a negative school environment, both of which may, in fact, increase school disorder. Until evidence supporting the use of these practices is provided, schools would benefit from focusing on other evidence-based strategies to reduce school disorder and respond to student misbehavior, rather than continuing to use restrictive security practices that mirror the criminal justice system. If changes are not made, the school-to-prison pipeline may, ironically, ensure that certain students are found in that system in the future.

Notes and Acknowledgements This entry is based on a more detailed manuscript (Payne and Eckert 2011).

Related Entries

- ▶ [Police Officers in Schools](#)
- ▶ [School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis](#)
- ▶ [School Social Organization, Discipline Management, and Crime](#)

Recommended Reading and References

- Addington LA (2009) Cops and cameras: public school security as a policy response to columbine. *Am Behav Sci* 52(10):1426–1446
- Berger R (2003) The best of both worlds: school security and the disappearing fourth amendment rights of students. *Crim Justice Rev* 28:336–354
- Brooks K, Schiraldi V, Ziedenberg J (2000) School house hype: two years later. Policy report of Justice Policy Institute and Children’s Law Center, Inc
- Chen G (2008) Communities, students, schools, and school crime: a confirmatory study of crime in U.S. high schools. *Urban Educ* 43:301–318
- Devine J (1996) Maximum security: the culture of violence in inner-city schools. University of Chicago Press, Chicago
- Devoe JF, Peter K, Noonan M, Snyder TD, Baum K (2005) Indicators of school crime and safety: 2005. U.S. Departments of Education and Justice, Washington
- Ginsberg C, Loffredo L (1993) Violence-related attitudes and behaviors of high school students—New York city, 1992. *J Stud Health* 63(10):438–439
- Giroux HA (2003) Racial injustice and disposable youth in the age of zero tolerance. *Int J Qual Stud Educ* 16:553–565
- Glazer S (1992) Violence in schools. *CQ Res* 34:785–808
- Gottfredson GD, Gottfredson DC (2001) What schools do to prevent problem behavior and promote safe environments. *J Educ Psychol Consult* 12:313–344
- Gottfredson GD, Gottfredson DC, Payne AA, Gottfredson ND (2005) School climate predictors of school disorder: results from the national study of delinquency prevention in schools. *J Res Crime Delinq* 42:412–444
- Hirschfield PJ (2008) Preparing for prison? The criminalization of school discipline in the USA. *Theor Criminol* 12:79–101
- Hyman IA, Peronne DC (1998) The other side of school violence: educator policies and practices that may contribute to student misbehavior. *J Sch Psychol* 36:7–27
- Kupchik A, Monahan T (2006) The New American School: Preparation for post-industrial discipline. *British Journal of Sociology of Education* 27:617–632
- Mayer MJ, Leone PE (1999) A structural analysis of school violence and disruption: implications for creating safer schools. *Educ Treat Child* 23:333–356
- Noguera PA (1995) Preventing and producing violence: A critical analysis of responses to school violence. *Harvard Educational Review*, 65(2), 189–212
- Noguera PA (2008) The trouble with black boys...and other reflections on race, equity, and the future of public education. Jossey-Bass, San Francisco
- Payne AA, Eckert RJ (2011) School security: How effective is it? Unpublished manuscript
- Payne AA, Gottfredson DC, Gottfredson GD (2003) Schools as communities: the relationships among communal school organization, student bonding, and school disorder. *Criminology* 41:749–777
- Tredway L, Brill F, Hernandez J (2007) Taking off the cape: the stories of novice urban leadership. *Theory Pract* 46:212–221
- Welch K, Payne AA (2011) Exclusionary school punishment: the effect of racial threat on expulsion and suspension. *Youth Violence Juvenile Justice* 10:155–171

School Social Organization, Discipline Management, and Crime

Denise C. Gottfredson¹, Philip J. Cook^{2,3} and Chongmin Na⁴

¹Department of Criminology and Criminal Justice, University of Maryland, College Park, MD, USA

²Sanford School of Public Policy, Duke University, Durham, NC, USA

³NBER, Cambridge, MA, USA

⁴School of Human Science and Humanities, University of Houston-Clear Lake, Houston, TX, USA

Overview

School violence, drug use, vandalism, gang activity, bullying, and theft are costly and interfere with academic achievement. Student misbehavior interferes with teaching and learning and is one of the primary sources of teacher turnover in our nation’s schools. Gallup polls from the past 20 years show that the percentage of parents who report being concerned about the physical safety of their children while at school has ranged from 15 % to 55 %, with the highest percentages registering just after the infamous school shootings at Columbine High School in 1999. Reducing crime rates has become an increasingly high priority for America’s schools.

There has also been an important trend in the official response to school crime. The response has become increasingly formal over the last 20 years, with greater recourse to arrest and the juvenile courts rather than school-based discipline. To some extent, this trend has been furthered by federal law that has imposed

zero-tolerance rules for some offenses and has subsidized the hiring of uniformed officers to police the schools. The shift has been from administrative discretion to mandatory penalties and from in-school discipline to increasing use of suspension or arrest. At the same time, there has been a considerable investment in the use of surveillance cameras and metal detectors.

While the increasing formality in school response to crime has coincided with the declining crime rates, there is no clear indication of whether the new approach gets any of the credit. Indeed, the evaluation literature that we review here has very little to say about the likely effects of these changes. As so often happens, there appears to be a disconnect between policy and research.

There are alternatives to the get-tough approach with its reliance on deterrence and exclusion. We know that some schools do a much better job than others in controlling the behavior of their students. Characteristic of successful schools in this respect is that they are close-knit communities where rules of acceptable behavior are clearly communicated and consistently (if not harshly) enforced. In addition to good management practices, there is much that can be done in the classroom that has demonstrated effectiveness in improving behavior.

In this entry, we summarize evidence showing that how the school is organized and managed (in general) influences problem behavior and school safety. We then discuss two specific aspects of school organizational climate – discipline management and school culture, respectively – and how they can be manipulated to reduce crime and related problem behaviors. We conclude with recommendations to guide future evaluation research on school-based interventions. Throughout the entry, our focus is on how schools can and do influence the behavior of students while they are enrolled.

Does School Organization Matter?

How the school is organized and managed influences problem behavior and school safety. In an

early national study of school disorder, Gottfredson and Gottfredson (1985) showed that even after controlling for input characteristics of students and communities in which schools were located, characteristics of schools accounted for an additional 12 % (junior high) and 18 % (senior high) of variance in teacher victimization rates. More recent national studies have replicated these findings and extended them to show that school characteristics account for a substantial amount of variance not only in teacher victimization but also in student reports of victimization and delinquency (Gottfredson et al. 2005).

Which aspects of the way schools are organized and managed influence crime and disorder? Cook et al. (2010) discuss school system decisions that influence the demographic composition of schools and the number and types of other students to whom a child is exposed. Schools and school districts have a good deal of control over the makeup of the student body. Schools can be based on neighborhood residential patterns or integrated across race and class. The grade span for elementary and middle schools can be adjusted. Truancy and dropout prevention programs can be pursued with more or less vigor, and troublesome students reassigned. Whether failing students are retained in grade or given a social promotion influences the extent of age homogeneity within classrooms. Students who are enrolled in the school can be tracked on the basis of academic potential or mixed together. These decisions influence the characteristics of other students to whom youths will be exposed. Importantly, these decisions determine the pool of youths from which highly influential peers will be selected as well as the dominant peer culture in the school.

School and school district decisions about curricular content and teaching methods are also important. These decisions determine student success in school and decisions to persist in school. As summarized in the Wilson entry elsewhere in this section, the use of specialized prevention curricula directly influences the level of problem behavior. Below, we discuss two additional characteristics of schools that influence crime and disorder: (1) policies and procedures

governing discipline management that directly affect the extent to which formal sanctions are applied and the effectiveness of these sanctions and (2) aspects of the school social organization that affect the nature of interactions among teachers and students (and hence the application of social controls) and the school culture.

Discipline Management

Cook et al. (2010) summarize findings from 12 studies that looked at the association of discipline management practices with school crime. The results show remarkable consistency: when schools monitor students and control access to the campus, and when students perceive that school rules are fair and consistently enforced, schools experience lower levels of problem behavior. Inclusion of students in establishing school rules and policies for dealing with problem behaviors has also been found to be related to lower levels of problem behavior, most likely because students are apt to internalize school rules if they have helped to shape them. On the other hand, severity of sanctions is not related to a reduction in problem behaviors. These findings conform to the main findings from deterrence research that the certainty of punishment has greater deterrent effect than the severity of punishment.

Of course, there has been considerable policy attention to school disciplinary practices, especially in response to the spate of school shootings experienced in the 1980s and 1990s. Most schools employ security and surveillance strategies aimed at keeping intruders out and preventing weapons from coming into the schools. Common practices include controlled entry and identification systems, metal detectors, security personnel or volunteers who challenge intruders, or doors fitted with electromagnetic locks. Since the late 1990s, school resource officers (SROs) have also been especially popular in secondary schools as a way to prevent violence, encouraged by federal subsidies. The Payne and Eckert and Na and Gottfredson entries elsewhere in this section discuss what is known about the effectiveness of these practices.

A closely related discipline strategy is the use of zero-tolerance policies in schools – another “tough on crime” practice engendered by the epidemic of youth violence in the late 1980s and the school rampage shootings of that decade and the next. The US Congress adopted the Gun-Free Schools Act in 1994, mandating that students be suspended for 1 year if they brought a gun to school. A large majority of school districts adopted zero-tolerance policies for alcohol, tobacco, drugs, and violence. The use of suspension, especially long-term suspension, is thought to have disproportionate impact on minority and special education populations, whose behavior places them more at risk for suspension. Civil liberty advocates have argued that zero-tolerance policies rob youths of their right to a public education.

As with other security-related school policies, little high-quality evidence is available to guide decisions about which discipline management policies produce the most desirable outcomes. The issue is complex, requiring consideration of the trade-offs between in-school and out-of-school crime, the welfare of the youths who perpetrate the school-based offenses versus that of the other youths in the school, and long-term versus short-term outcomes. Clearly, removing troublemakers from school helps to maintain an environment more suitable for learning for these remaining students. But the costs of doing so for the offenders and society are not well understood. A complete analysis of the effect of zero-tolerance policies on youth crime would consider the displacement of crime from school to the community as well as the consequences for the suspended youths’ long-term criminal and academic careers. As youths lose more days of school to suspension, promotion to the next grade becomes less likely. And as youths fall further behind grade, they become much less likely to graduate, which is likely to increase subsequent crime. Clearly, although zero-tolerance policies benefit the classmates of troublesome youths, a rational discipline policy would also have to consider the broader consequences of such policies for the community.

More consistent with the research on effective crime deterrents are school discipline policies that emphasize the certainty of response to misbehavior over the severity of the response. Among the most effective school-based strategies for reducing youth violence, aggression, and problem behavior are behavioral interventions that target specific behaviors, systematically remove rewards for undesirable behavior, and apply contingent rewards for desired behavior or punishment for undesired behavior. These interventions are often applied to the high-risk youths who are most at risk for being suspended from school under zero-tolerance policies and as such could be incorporated into school routines for discipline management. Gottfredson et al. (2002) meta-analysis reported average effect size on measures of antisocial behavior and aggression of 0.34 ($p < .05$) across 12 studies of this type of behavioral intervention.

Examples of particularly effective behavioral interventions currently in use in schools are the “Good Behavior Game” (GBG; Kellam et al. 2008) and “home-based reinforcement” (Schumaker et al. 1977). The GBG is a classroom-based application of behavioral principles in which elementary school children are divided into small teams, and the teams are rewarded when the classroom behavior of the entire team meets or exceeds a preestablished standard. The GBG is played several times per week throughout the school year. The intervention was evaluated through a randomized trial involving 19 schools in Baltimore, with posttests conducted immediately following the intervention, as well as 6 and 14 years later. The results of this study indicate that participation in GBG is related to immediate reductions in aggressive behavior, rates of diagnosed antisocial personality disorder, and long-term effects (14 years later) on drug and alcohol use and smoking. Home-based reinforcement (HBR), applied to individual students displaying behavior problems, requires cooperation between teachers and parents in the management of the child’s behavior. After agreeing upon specific child behaviors to be extinguished or encouraged and establishing a baseline for these behaviors, teachers

systematically record data on the target behavior on a “daily report card” that goes home to the parents. The parents, who generally have access to a wider array of reinforcers and punishments than do the teachers, use the teacher’s information to guide the application of rewards and punishments. As the desired behavior emerges, the frequency of reports for home is reduced, and the schedule of contingencies is relaxed. In the earliest research on HBR, application of this technique to junior high school students showed that school rule compliance, teacher satisfaction with the student, and academic performance improved as a result of participation in an HBR program (Schumaker et al. 1977). A recent review of 18 empirical studies of “school-home collaboration” interventions (Cox 2005) concluded that behavioral interventions using the daily report card strategy had the strongest effects on problem behavior. Lasting effects on crime are unknown.

These relatively simple and inexpensive behavioral interventions represent a potentially potent school-based prevention strategy that might be incorporated into routine school practice. The 1997 reauthorization of the Individuals with Disabilities Education Act (IDEA; P.L. 105–17) required functional assessment and behavioral intervention procedures to be implemented in the disciplining of students with disabilities. The evidence-based programs described here would meet these federal requirements.

Behavioral principles have also been incorporated into school-wide discipline management systems. These systems are typically designed to clarify expectations for behavior. They establish school and classroom rules, communicate these rules as well as consequences for breaking them clearly to parents and students, establish systems for tracking both youth behavior and consequences applied by the schools, and monitor the consistency of the application of consequences for misbehavior. School-wide discipline management efforts, most often implemented by a school-based team of educators, are highly consistent with the research summarized earlier suggesting that students’ perceptions of school rules as fair and consistently enforced are related to reductions in problem behavior.

The meta-analysis described earlier (Gottfredson et al. 2002) also examined the effectiveness of this type of school-wide effort to improve discipline management and reported average effect size on measures of crime (0.27, $p < .05$) and alcohol and other drug use (0.24, $p < .05$). Among the studies included in the meta-analysis are two early studies of the effects of school-wide discipline management systems on problem behavior outcomes. Students in the intervention schools in the first of these efforts (Project PATHE implemented in nine Charleston, South Carolina, schools) reported less delinquent behavior and drug use and fewer punishments in school relative to the students in the comparison schools (Gottfredson 1986). A similar intervention was tested in a troubled Baltimore, Maryland, junior high school, with a special emphasis on replacing the school's reliance on out-of-school suspension with a wider array of consequences for misbehavior. This intervention, which added positive reinforcement for desired behavior to the mix of consequences routinely used, also showed positive effects on student delinquency and rebellious behavior (Gottfredson 1987a). This early research, although based on relatively small numbers of schools and lacking randomization to condition, suggested that behavioral principles could be incorporated into "normal" school disciplinary practices and that an emphasis on consistency of rule enforcement as opposed to severity of punishment provided an effective deterrent.

Contemporary approaches to discipline management incorporate behavioral principles into comprehensive systems that include school-wide discipline policies and practices as well as targeted behavioral interventions. One popular approach is School-Wide Positive Behavior Support (SWPBS). This system, adopted by over 5,600 schools throughout the United States, uses a school-team approach to apply behavioral interventions at different levels of intensity for students at different levels of need. Universal interventions focus on clarity of school and classroom rules and consistency of enforcement, and on screening for more serious behavior disorders. Group-based behavioral interventions are

employed with the 5–10 % of youths who do not respond to the universal interventions. In addition, intensive, individualized behavioral interventions are employed to manage the behavior of the small segment of the population that is especially at risk. Unfortunately, the research on the effectiveness of SWPBS is not as sophisticated as it should be for such a widely disseminated program. Although dozens of studies have demonstrated that problem behavior decreases after the intervention is put in place, only one (Sprague et al. 2001) compared change in the intervention school(s) with the change that might be expected in the absence of an intervention. Even this study is not useful for isolating the effects of the behavior management strategies because it also included the introduction of a prevention curriculum along with the school-wide behavioral supports. Higher-quality research is needed to assess the effects of this promising approach on crime both in and out of school.

School Culture

School culture is potentially the most potent aspect of school climate because it involves proximal interpersonal influences on student behavior. *School culture* refers to the quality of human relationships in the school and includes both peer culture and the extent to which the organization is communally organized. All of these dimensions influence youth crime and can be successfully manipulated to reduce it.

Behavioral Norms

Cultural norms, expectations, and beliefs influence all behaviors. Recent research on school culture (summarized in Cook et al. 2010) concurs that norms and expectations for behavior, of both peers and adult, are powerful determinants of behavior, net of the individual's own beliefs.

Of course, school "inputs" are key determinants of the predominant cultural beliefs in the school. School desegregation and retention policies, as well as the grade span, of the school can influence school culture by altering the mix of

students in the school. But several more programmatic attempts to alter school culture have also been studied. These programs have in common a focus on clarifying behavioral norms. That is, in contrast to the instructional programs described in the previous section that focus on teaching youths with specific social competency skills, these normative change programs focus on clarifying expectations for behavior. Some signal appropriate behavior through media campaigns or ceremonies; others involve youths in activities aimed at clarifying misperceptions about normative behavior; and still others increase exposure to prosocial models and messages.

Several studies of attempts to clarify norms for behavior have been reported. Gottfredson et al. (2002) summarized effects reported in 13 studies and concluded that such programs are effective for reducing crime, substance use, and antisocial behavior. Two of the better-known examples of programs in this category are the Bullying Prevention Program (Olweus et al. 1999) and the Safe Dates Program (Foshee et al. 1998).

Olweus's anti-bullying program includes school-wide, classroom, and individual components. School-wide components include increased adult supervision at bullying "hot spots" and school-wide discussions of bullying. Classroom components focus on developing and enforcing rules against bullying. Individual counseling is also provided to children identified as bullies and victims. A large-scale evaluation of this program in Norwegian schools demonstrated that it led to reductions in student bullying and victimization and decreases in the incidence of vandalism, fighting, and theft (Olweus et al. 1999). A very recent review of anti-bullying programs summarizing results from 59 studies conducted between 1983 and 2008 (Farrington and Ttofi 2009) confirmed that anti-bullying programs are effective for reducing bullying and student victimization and that Olweus's program is particularly effective.

The Safe Dates Program targets norms for dating violence among adolescents. The school portion of the intervention includes a theater production performed by peers; a ten-session curriculum addressing dating violence norms, gender

stereotyping, and conflict management skills; and a poster contest. The community portion of the intervention includes services for adolescents experiencing abuse and training for community service providers. Foshee et al. (1998) found that intervention students reported less psychological abuse and violence against dating partners than did control students.

Based on these and other relatively rigorous evaluations, Gottfredson et al. (2002) concluded that interventions aimed at establishing norms or expectations for behavior can be effective in preventing substance use, delinquency, aggression, and other problem behaviors. It should be noted, however, that evaluations of these programs seldom provide clean tests of the proposition that culture matters, since the programs more often than not combine attempts to alter norms with other components aimed at increasing levels of supervision and enforcement (e.g., Olweus) or improving social competency skills (Foshee).

We would be remiss if we failed to mention that sometimes school-based practices seek to clarify norms for behavior backfire. One example is a peer counseling program that deliberately mixed delinquent and nondelinquent youths in counseling sessions in which youths were encouraged to share their problems. The intent was that the negative beliefs and attitudes voiced by the delinquent youths would be corrected through interaction with the nondelinquent youths. A randomized experiment testing this program as implemented in the Chicago Public Schools (Gottfredson 1987b) reported predominantly harmful effects for high school students: high school treatment youths reported significantly *more* delinquent behavior than controls. A more recent large-scale evaluation of the Reconnecting Youth program (Cho et al. 2005) also found negative effects for a group counseling program for at-risk high school students. This program sought to "reconnect" truant, underachieving high school students (and to reduce their deviance and substance use) by developing a positive peer-group culture. Students were grouped together in classes of 10–12 students for a full semester during which a trained group leader (following a standardized

curriculum) attempted to develop a climate conducive to building trust. The evaluation reported only negative effects 6 months following the end of the intervention. Treatment students showed greater bonding to high-risk peers, lower bonding to school and conventional peers, lower grade point average (GPA), and higher anger than control students at the 6-month follow-up.

Communal Social Organization

A second aspect of school culture that has been studied extensively pertains to the affective bonds between students and teachers and among adults in the school. The concept of “communal social organization” (CSO) was first introduced as part of the effective schools debate in the 1980s and studied by Bryk and colleagues (Bryk and Driscoll 1988) mostly in the context of predictors of school achievement. Communally organized schools are schools characterized by high levels of social support, connectedness, common goals, and sense of shared purpose. Members of such schools are more likely to be involved and personally committed to the school. This aspect of school culture is especially important for school crime research because individual-level student affective bonds are an important predictor of delinquency, and it seems reasonable to hypothesize that schools high on CSO would produce higher levels of student bonding to school.

Research suggests that average student attachment to school and CSO more generally do inhibit student problem behaviors. The most comprehensive test of this linkage was provided by Payne et al. (2003) using data from the NSDPS. This study demonstrated that more communally organized schools experience less student delinquency and teacher victimization and that the effect of communal school organization on student delinquency is mediated by average student bonding.

This survey research dovetails nicely with an ambitious ethnographic study of school violence conducted for the National Research Council. In 2003, the Committee to Study Youth Violence in Schools of the National Research Council published its report on the circumstances

surrounding several incidents involving extreme lethal violence that had occurred in the nation’s schools (National Research Council 2003). The report was based on detailed case studies of six schools and communities that had experienced school shootings resulting in death. Among the committee’s several insights into the factors leading to the incidents is the following:

the sense of community between youth and adults in these schools...was lacking. In the worst example, the school allowed a school newspaper to print an article that humiliated one of the students who became a shooter. The adults involved may have been too distant from the students to prevent some social processes leading to the potential for violence or resulting in an intolerable humiliation from some potentially vulnerable youth. (p. 256)

This observation is consistent with the research on more mundane forms of school violence just summarized. It suggests that strategies that increase social bonds between students and others in their schools will reduce misbehavior by increasing informal controls. Students who care what adults in the school think about them will be less likely to act in ways that jeopardize their positive regard. More concretely, students who have close ties to the adults in the school will be more likely to report on rumors of impending attacks. But how can such bonds be built or maintained? Possibilities include organizing the school so that the typical teacher interacts with fewer students, reducing class size, and creating more “communal” social environments in which members are more tightly joined together by common goals and in which members are held in place by the support and positive regard of others in the organization. Reorganizing schools to create a smaller feel to the schooling experience is an effective strategy for increasing youths’ sense of connection and that enhanced connectedness should hold criminal behavior in check.

A less drastic intervention with the same objectives is mentoring. Youth mentoring programs often target youths at risk of behavioral problems, assigning them to an adult mentor who spends time with the young person, provides support and guidance, and offers general guidance. Evaluations of such programs have been mixed,

but often null or weak results can be attributed to implementation failure. As with any voluntary program, mentoring programs in practice are often not as intensive as intended (e.g., Karcher 2008). However, a recent meta-analysis of mentoring programs (Eby et al. 2007) demonstrated small but positive effects of mentoring programs on several behaviors of interest in this entry: withdrawal behaviors (e.g., school dropout, truancy – 18 studies), deviance (e.g., suspension from school, aggressive behavior, property crime – 15 studies), and substance use (7 studies). This review included a wide range of types of mentoring programs, but outcomes for youth mentoring programs were as strong on these outcomes as were the other types of mentoring programs (academic and workplace mentoring) included in the review.

One of the better-known models for adult mentoring, the Big Brothers Big Sisters program (BBBS), is a community-based program identified by BVP as a model primarily on the basis of evidence from a large-scale randomized trial that found that mentored youths were 46 % less likely than control youth to initiate drug use, 27 % less likely to initiate alcohol use, and almost one-third less likely to hit someone during the study period (Tierney et al. 1995). Community-based mentoring involves meetings between the mentor and mentee at times and places selected by the pair. Many schools now provide “school-based mentoring,” (SBM), which involves meetings primarily in school during the school day. A recent evaluation of the BBBS SBM model, also involving random assignment of a large number of youths, shows that although it is not as effective as the community-based alternative, SBM does improve academic performance, reduce truancy, and reduce serious school infractions (Herrera et al. 2007) at least during the first year of mentoring. Consistent with results from smaller-scale randomized trials of SBM showing positive effects on connectedness and social support (Karcher 2008); Herrera et al. (2007) found that mentored youths reported more often than controls the presence of a nonparental adult in their life who provides social supports. At the end of the second year of the study during which

minimal SBM was provided, the positive program effect on truancy was sustained but the other positive effects were not. Herrera et al. (2007) conclude that although the SBM model is promising, it needs to be strengthened to ensure longer and higher-quality mentor/mentee matches than are typically found in schools.

Discussion and Conclusions

In this brief essay, we summarized research on school discipline management policies and practices and showed that they are important determinants of school crime. Research consistently shows that in schools in which students report that the school rules are clearly stated, fair, and consistently enforced, and in schools in which students have participated in establishing mechanisms for reducing misbehavior, students are much less likely to engage in problem behaviors. We showed that evaluations of specific school-based programs that employ behavioral strategies to monitor and reinforce student behavior are effective both for controlling behavior in school and for reducing subsequent crime. Also, altering school-wide discipline management policies and practices to incorporate behavioral principles, clarify expectations for behavior, and consistently enforce rules reduces problem behavior. We discussed popular “get-tough” approaches to school discipline such as zero-tolerance policies. Although the effects of these policies on crime are not known, we argued that they might actually increase crime outside of school. There is a clear need for rigorous research on the effects of these policies.

Finally, we summarized research showing that perceptions of social norms for behavior are related as expected to problem behavior, net of individuals’ personal beliefs. In schools in which the prevailing norm is to condone delinquent activities, students are more likely to do so regardless of their own personal dispositions to engage in these behaviors. But we showed that schools can intervene to change perceptions of norms and expectations for behavior and that doing so reduces delinquency, although attempts

to do so sometimes backfire. We also reported on evidence suggesting that in schools in which students feel an emotional attachment to the adults in the school, their misbehavior is restrained. We discussed several strategies that might increase communal social organization and that show promise for increasing youths' sense of connection to the school. We reviewed research on school-based mentoring programs and showed that they also hold considerable promise for crime prevention. Although research documents positive effects of these programs on social relations outcomes, more work is needed to test the full potential of more potent models of school-based mentoring than have been tested to date.

Given the limitations of the evidence base, we are more confident in making recommendations about research priorities than about effective policy. Indeed, this field is burdened by a lack of timely policy research and a tendency to launch major initiatives without first (or ever!) doing a high-quality evaluation. Note in this regard the various "get-tough" policies that have been encouraged by the federal government and adopted nationwide since the 1990s, the widespread use of SROs, or the School-Wide Positive Behavior Support package that has been adopted by 5,500 schools.

We have several recommendations to guide evaluation research on interventions. The first recommendation is to actually do such research, as suggested above. An impediment to learning about the effects of many school reforms is that the reforms tend to be implemented in all schools in the affected jurisdiction at once. This hinders rigorous evaluation because it leaves no schools in which to measure what would happen in the absence of the reform. A smarter approach would be to randomly assign schools to different phase-in periods, allowing for comparison during the first few years of the schools who implement the reform early and those who will implement it in the future.

Other recommendations are to measure effects on crime and other forms of misbehavior in evaluations conducted of interventions intended to improve academic performance, to capture the most serious forms of crime in evaluations rather

than only less serious misbehavior, and to assess effects of prevention practices and policies on the entire student population rather than only on the students who are targeted.

Finally, it is important to identify programs to create more cohesive, communal, personalized environments. Many approaches to creating such environments seem plausible, but no rigorous research has yet established that such changes can be accomplished and that doing so results in a reduction in crime. This appears to be the next large challenge facing research on school-based prevention.

Notes and Acknowledgments This entry is based on a more detailed chapter (Cook et al. 2010) on school crime control and prevention. Portions of this work are also included in a forthcoming chapter entitled "School-Based Crime Prevention" in the *Oxford Handbook on Crime Prevention* (Brandon C Welsh and David P. Farrington, editors).

Related Entries

- ▶ [Police Officers in Schools](#)
- ▶ [School Security Practices and Crime](#)
- ▶ [School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis](#)

Recommended Reading and References

- Bryk AS, Driscoll ME (1988) The high school as community: contextual influences and consequences for students and teachers. University of Wisconsin, National Center on Effective Secondary Schools, Madison
- Cho H, Hallfors DD, Sánchez V (2005) Evaluation of a high school peer group intervention for at-risk youth. *J Abnorm Child Psychol* 33:363–374
- Cook PJ, Gottfredson DC, Na C (2010) School crime control and prevention. In: Tonry M (ed) *Crime and justice: a review of research*, vol 39. University of Chicago Press, Chicago, Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1368292
- Council NR (2003) *Deadly lessons: understanding lethal school violence*. National Academy Press, Washington, DC
- Cox DD (2005) Evidence-based intervention using home-school collaboration. *School Psychol Quart* 20:473–497

- Eby LT, Allen TD, Evans SC, Ng T, DuBois DL (2007) Does mentoring matter? A multidisciplinary meta-analysis comparing mentored and non-mentored individuals. *J Vocat Behav* 72:254–267
- Farrington DP, Ttofi MM (2009) Reducing school bullying: evidence-based implications for policy. In: Michael T (ed) *Crime and justice: a review of research*, vol 38. University of Chicago Press, Chicago
- Foshee VA, Bauman KE, Arriaga XB, Helms RW, Koch GG, Linder GF (1998) An evaluation of safe dates, an adolescent dating violence prevention program. *Am J Public Health* 88:45–50
- Gottfredson DC (1986) An empirical test of school-based environmental and individual interventions to reduce the risk of delinquent behavior. *Criminology* 24:705–731
- Gottfredson DC (1987a) An evaluation of an organization development approach to reducing school disorder. *Eval Rev* 11:739–763
- Gottfredson GD (1987b) Peer group interventions to reduce the risk of delinquent behavior: a selective review and a new evaluation. *Criminology* 25:671–714
- Gottfredson GD, Gottfredson DC (1985) *Victimization in schools*. Plenum Press, New York
- Gottfredson DC, Wilson DB, Najaka SS (2002) School-based crime prevention. In: Sherman LW, Farrington DP, Brandon Welsh C, MacKenzie DL (eds) *Evidence-based crime prevention*. Routledge, London
- Gottfredson GD, Gottfredson DC, Payne AA, Gottfredson NC (2005) School climate predictors of school disorder: results from a national study of delinquency prevention in schools. *J Res Crime Delinq* 42:412–444
- Herrera C, Grossman JB, Kauh TJ, Feldman AF, McMaken J, Jucovy LZ (2007) *Making a difference in schools: the big brothers big sisters school-based mentoring impact study*. Public/Private Ventures, Philadelphia
- Karcher MJ (2008) The study of mentoring in the learning environment (SMILE): a randomized evaluation of the effectiveness of school-based mentoring. *Prev Sci* 9:99–113
- Kellam SG, Hendricks Brown C, Poduska J, Ialongo N, Wang W, Toyinbo P, Petras H, Ford C, Windham A, Wilcox HC (2008) Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug Alcohol Depend* 95:S5–S28
- Olweus D, Limber S, Mihalic S (1999) *Blueprints for violence prevention: bullying prevention program*. Center for the Study and Prevention of Violence, Boulder
- Payne AA, Gottfredson DC, Gottfredson GD (2003) Schools as communities: the relationships among communal school organization, student bonding, and school disorder. *Criminology* 41:749–777
- Schumaker JB, Hovell MF, Sherman JA (1977) An analysis of daily report cards and parent-managed privileges in the improvement of adolescents' classroom performance. *J Appl Behav Anal* 10:449–464
- Sprague J, Walker H, Golly A, White K, Myers D, Shannon T (2001) Translating research into effective practice: the effects of a universal staff and student intervention on key indicators of school safety and discipline. *Educ Treat Child* 24:495–511
- Tierney JP, Grossman JB, Resch NL (1995) *Making a difference: an impact study of big brothers big sisters*. Public/Private Ventures, Philadelphia

School Structural Characteristics and Crime

Philip J. Cook^{1,2}, Denise C. Gottfredson³ and Chongmin Na⁴

¹Sanford School of Public Policy, Duke University, Durham, NC, USA

²NBER, Cambridge, MA, USA

³Department of Criminology and Criminal Justice, University of Maryland, College Park, MD, USA

⁴School of Human Science and Humanities, University of Houston-Clear Lake, Houston, TX, USA

Overview

Other essays in this encyclopedia summarize school-based prevention strategies that have been demonstrated to be effective for reducing crime and discuss the importance for schools of having close-knit communities where rules of acceptable behavior are clearly communicated and consistently (if not harshly) enforced. Schools' capacity to implement effective prevention strategies and to establish well-ordered communities is likely to be greater in schools that are not overwhelmed by having a high proportion of the students at risk. In this essay, we discuss decisions primarily made outside of the school building that influence the demographic composition of schools and the number and types of other students to whom a child is exposed. Externally determined factors (e.g., finances, physical features of building, school size, and student/teacher ratio) also determine resources available in the school and

define patterns of interaction broadly speaking. Here we will focus on school-district policies influencing the student composition of schools and school size.

Specifically, we will consider the extent to which school or school-district decisions regarding how students are organized for instruction (e.g., academic or behavioral tracking or departmentalization) further narrow the characteristics of other students to whom youths will be exposed. Importantly, these decisions determine the pool of youths from which highly influential peers will be selected as well as the dominant peer culture in the school. These structural characteristics potentially influence crime-producing mechanisms.

We find that the concentration of different types of students in a school has important implications for the amount of crime in the school: The grade levels included in the school or average age of the students in the school, the percentage male students, the social class composition, and the racial and ethnic composition of the schools are related to measures of problem behavior. Further, some policies that alter the composition of classes, grades, or schools have been shown to reduce problem behavior: Keeping sixth graders in elementary schools as opposed to moving them to middle schools reduces disciplinary infractions, and retaining students in grade increases conduct problems both for the “old-for-grade” students and for non-retained students. We find little support for the idea that the size of the school per se matters for reducing school crime.

We begin this essay with a summary of research related to school size and student demographics as they relate to school crime and disorder. We then discuss school policies and practices that are likely to alter the types of students with whom a student is likely to be exposed. We end with a discussion of implications for policy and research.

School Ecology

One category of school structural characteristics is “school ecology.” Very few studies have

examined the influence of several potentially important aspects of school ecology including school finances and the physical features of the school building. By contrast, many studies include a measure of school size – number of students in the school or in the grade. Our discussion will therefore focus on school size, providing a summary of the literature and some new results.

School size is thought to have a major influence on the internal organization of schools and on subsequent student outcomes. Lee et al. (1993) suggest that larger schools are likely to have increased capacity to tailor programs and services to meet the diverse needs of students in the school. The extreme example of low specialization is a one-room schoolhouse in which one teacher teaches all students all day. In small schools, the typical teacher teaches a smaller number of different students and gets to know these students well. Students in such schools may develop a greater sense of trust in the adults and be more likely to communicate potentially dangerous situations to them. Large schools are likely to be organized more bureaucratically and to involve more formalized social interactions among members of the school population. As a result, communication may be less frequent or less direct, cohesiveness may be reduced, management functions (including the management of discipline) may become less nuanced, and individuals may share less of a common experience in the school. Alienation, isolation, and disengagement may result. All of these mechanisms are plausible but speculative.

As it turns out, school size has not received much focused attention in research on schools and crime. However, many studies have included a measure of school size as a control variable when focusing on the effects of other aspects of school climate. Cook et al. (2010) summarized the associations between measures of school size and problem behavior in school-level and multilevel studies. They found nine school-level studies based on data from seven different data sources. In the studies that reported an unambiguous association between school size and a measure of problem behavior, the conclusions

differed depending among other things on the measure of problem behavior used. Positive associations between school size and measures of minor misbehavior were reported for the High School and Beyond high school data and the National Education Longitudinal Study [NELS] eighth graders, but the association with more serious forms of misbehavior was not statistically significant. In another data source (Safe School Study), school size was not significantly related to student victimization but was positively related to teacher victimization. That study also showed that the effect of school size on teacher victimization is mediated by aspects of the school social organization and culture. No significant relationship with school size was found in a study of middle schools in Philadelphia.

The multilevel studies summarized provided no support for the “smaller is better” viewpoint. Fifteen different research reports based on nine different data sources were summarized. In these studies, which generally controlled for community characteristics as well as characteristics of the students who attend the school, only one data source (NELS tenth graders, as reported in Stewart 2003) produced a significant positive association between school size and a measure of problem behavior, and the measure of problem behavior used in this study was unusual because it contained mainly school responses to misbehavior (e.g., being suspended or put on probation) rather than actual youth behavior. Hoffmann and Dufur (2008) also reported on the association of school size and a broader measure of problem behaviors including substance use, arrest, and running away using the NELS tenth grade sample and found no significant association. Reports from a sample of Israeli schools containing seventh and 11th grades documented a positive association between average *class* size and student victimization, but no significant association with school size. One of the multilevel studies reported a significant negative association between school size and student victimization, but this sample included only rural schools located in New Brunswick, Canada, whose average size

was 39 and 53 students, respectively, for sixth and eighth grade. A recent study by Gottfredson and DiPietro (2011) found that, net of individual-level risk factors and confounding characteristics of schools and their surrounding communities, more students per teachers and a higher number of different students taught by the average teacher were related to higher student victimization rates, but larger school size was significantly related to lower student victimization rates. Cook et al. (2010) concluded that most of the multilevel studies suggest that school size is not reliably related to student problem behavior once characteristics of the students who attend the schools are controlled.

However, the studies summarized often reported on the association between school size and problem behavior from models that may provide too conservative a test. The student characteristics that are controlled in the multilevel studies are often exactly those student characteristics that Lee et al. (1993) hypothesized to be influenced by school size (e.g., school attachment, involvement, perceived positive social climate). Also, most of the associations with school size reported in the studies are from models that partial out influences not only of the communities in which the schools are located and the average demographic characteristics of the students attending the schools but also of other school climate characteristics such as school culture and the administration/management of discipline, hypothesized to mediate the influence of school size on student outcomes. For example, Hoffmann and Dufur’s (2008) study reported a negative association between school size and delinquency in the NELS data, but the equation also contained measures of “school quality,” a composite measure assessing youth perceptions of their school as fair and their teachers and fellow students as caring and trustworthy. Unfortunately, many of these reports did not report the association of school size with problem behavior in models that do not control for potential effects of school size.

New results. We analyzed data from the 2007–2008 School Survey on Crime and Safety

(SSCS; Neiman and DeVoe 2009) in an attempt to establish baseline descriptive results on how school size relates to school crime. Principals were asked how many incidents of various types of crime had occurred at school during the last school year. They were asked about violent crimes (rape, sexual battery other than rape, physical attack or fight with or without a weapon, threat of physical attack with or without a weapon, and robbery with or without a weapon) and about theft and larceny. We calculated rates per 1,000 students for each school. Because school size is highly related to location and level, it is necessary to look at the association of crime rates and enrollment while controlling for these factors.

Figure 1 shows median rates per 1,000 students for theft/larceny and violent crimes, according to school principals. The figures make clear that crime rates are not systematically related to school size within level and location. However, there is some suggestion that the association between size and principal reports of crime differs according to type of crime, level, and location: In urban locations, principals in smaller elementary and middle schools report more violent crimes. This is not the case in rural/suburban schools or in urban high schools.

We conclude that school size is not generally related to principal reports of school and that whatever differences observed favor larger schools over smaller schools.

It is likely that the ratio of adults to students rather than the actual number of students in the school is related to problem behavior. Five of the studies summarized in Cook et al. (2010) looked at the association of problem behavior to student/teacher ratio. Only one of the five studies reported a significant relationship: Gottfredson and DiPietro (2011), in a multilevel study using data from the National Study of Delinquency Prevention in Schools, reported a positive association of student/teacher ratio and personal but not property crime victimization. That study reported that higher levels of social capital, as measured by student consensus about normative beliefs, partially mediated the effects of student/

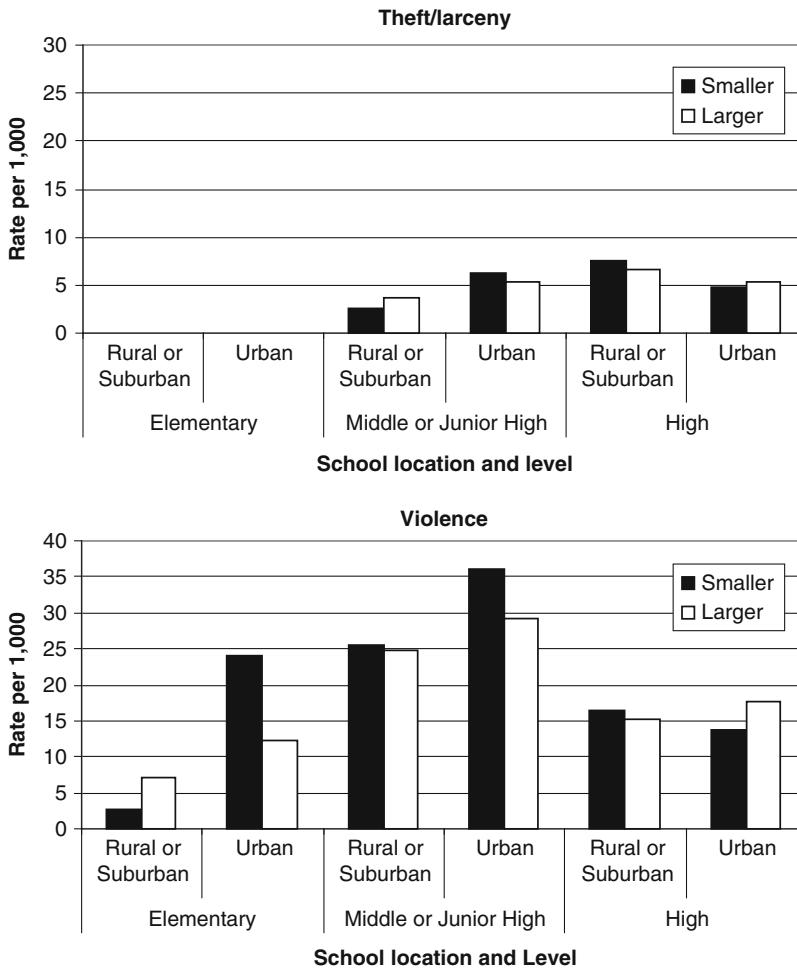
teacher ratio on personal victimization. We believe that a more sensitive measure of adult presence would be the ratio of all adults (rather than just teachers) to students. Many schools use parent volunteers and teacher aides in addition to teachers to help maintain order. The ratio of the total number of adults to students would reflect variability in the use of such auxiliaries. Unfortunately, no studies have reported on this association.

As far as we know, there are no intervention-based studies of how school size affects school crime. Case studies of instances in which an established school is divided in smaller units are available, but they almost never assess effects on crime, and they do not provide a clean test of the effects of changing school size because other factors (such as the curriculum, aspects of the physical space, and school finances) are always altered simultaneously.

Milieu

The SSCS data show that rates of problem behavior differ with demographic characteristics. Middle schools have higher rates of delinquency than elementary or high schools (the partial exception is substance use, which increases through high school). Schools with 50 % or more minority enrollment experience higher rates of violence than majority white schools (Nolle et al. 2007). Socioeconomic status of the student body is also associated with delinquency rates (Gottfredson et al. 2005).

More interesting from a policy perspective is the extent to which the mix of students in the school or the classroom influences the likelihood that any given student will misbehave. The mechanisms of deviant peer influence are both direct and indirect. The direct effects may arise as a result of deviant peer influence: learning and imitation, social reinforcement for deviant acts, and the creation of opportunities for deviant activities (Dishion and Dodge 2006). It may also be due to the presence of social norms that support (or at least appear to youth to support) delinquent behavior. All these mechanisms are



School Structural Characteristics and Crime, Fig. 1 Principal reports of crimes recorded by school, by school enrollment, location, and level – median rates per 1,000 students. Note: Size is split at the median total enrollment for each school level: 452, 585, and 885 for elementary, middle, and high, respectively. One outlier (a small alternative school for delinquent youths with a very high crime rate) is excluded. Violent crimes include rapes, sexual batteries other than rape, robberies with or without a weapon, physical attacks or fights with or without

a weapon, and threats of physical attack with or without a weapon. Theft includes taking things worth over \$10 without personal confrontation and includes pocket picking, stealing a purse or backpack (if left unattended or no force was used to take it from owner), theft from a building, theft from a motor vehicle or [of] motor vehicle parts or accessories, theft of a bicycle, theft from a vending machine, and all other types of thefts (Source: Original tabulation from data files. SSCS 2007–2008)

relevant for involvement with delinquency both in and out of school, including drugs and alcohol, and participation in gangs. The indirect effects may come about as a result of the dilution of authority – a teacher who can manage one or two disruptive students may lose control of the classroom when there are more than two. The same phenomenon can occur at the school level,

where a high “load” of troublesome students may swamp the mechanisms of control in the corridors, cafeteria, lavatories, and grounds.

Given the real possibility of peer influence, the actual behavior of youths with a given propensity to deviant or criminal activity may well depend on whom they encounter in their classes and in the other locations in the school. A variety of

policies are relevant to influencing the mix of students. At the level of the school district, the distribution of students among schools will be influenced by which grade spans are included in the middle schools, the extent to which low-performing students are held back, and whether school assignments are tied largely to place of residence or tailored to promote integration or parental choice. For a given pattern of assignments to schools, the number and characteristics of students who are actually in the building on a school day will depend on absenteeism and use of out-of-school suspension. And for a given population of students who are actually attending the school on any given day, social influence will likely be mediated by policies that influence the extent to which deviant students are concentrated, such as in-school suspension or academic tracking.

Cook et al. (2010) summarized the results from 18 multilevel studies based on 14 datasets showing associations between the milieu of the school and measures of problem behavior, controlling for individual-level demographics as well as for related characteristics of the communities from which the student body is drawn. These studies showed that the grade levels included in the school or average age of the students in the school, the percentage male of students, the social class composition, and the racial and ethnic composition of the schools are related to measures of problem behavior. These associations sometimes did not reach statistical significance, but they were nearly always in the expected direction. Here we discuss several of the strongest studies relevant to evaluating the impact of policy choices concerning grade span, grade retention, truancy prevention, racial segregation, and use of alternative schools.

1. *Grade span.* One recent study demonstrates that the grade composition in a middle school influences the rates of misbehavior of the students. A generation ago, most elementary schools included sixth grade, but now most sixth graders attend middle school. Using a quasi-experimental approach, Cook et al. (2008) compare the school records of North Carolina students whose sixth grade was

located in a middle school with those whose sixth grade was in elementary school (the sample of sixth grades in middle school was trimmed to match the sixth grades in elementary school in several dimensions). While the two groups of students had similar infraction rates in fourth and fifth grades, those who moved to a middle school for sixth grade experienced a sharp increase in disciplinary infractions relative to those who stayed in elementary school. More interesting, perhaps, is that the elevated infraction rate persisted through ninth grade. A plausible interpretation of these findings is that sixth graders are at a highly impressionable age, and if placed with older adolescents, they tend to be heavily influenced by their inclination to break the rules. This example of negative peer influence is quite large and extends to all types of infractions, including violence and drug violations.

2. *Retention policies.* The age mix in a school is closely related to the grade span, but that is not the only determinant; the school district's retention policies also play a role. In response to high-stakes accountability programs (including No Child Left Behind), many school systems have ended social promotion for students who fail end-of-grade tests, thus increasing the number of old-for-grade students. Entry-level at-risk students are often held back for a year before making the transition to second grade. The effect of retention on behavior of the retained students has been extensively studied. Most studies have focused on academic outcomes: Meta-analyses of this literature conclude that the long-term effect on academic achievement is null or negative, with a greatly elevated risk of dropping out (Jimerson et al. 2006). Hence, given the robust general finding that students with academic difficulties are more prone to antisocial behavior (Nagin et al. 2003), it is not surprising that grade retention appears to increase conduct problems. One of the most sophisticated studies, using Richard Tremblay's longitudinal data on Montreal

school children, found that the effect of grade retention on classroom physical aggression (as measured by teacher reports) is conditioned by the developmental history of the child: Those showing no aggression or chronic aggression levels were not affected, whereas those whose trajectory of aggression was declining over time increased their aggression more so if retained than if not retained (Nagin et al. 2003).

There has been less attention to the contextual effect of having old-for-grade students in the classroom and school. One exception is a recent study that uses a comprehensive data set of North Carolina students. Muschkin et al. (2008) conduct a cross-sectional analysis of infraction rates by seventh graders, finding that the prevalence and incidence of infractions increase with the prevalence of retained students (students who were retained at least once in the previous 3 years) and the prevalence of old-for-grade students who were not retained during that 3-year period. These results hold after they controlled for various characteristics of the student body and the schools and the inclusion of district fixed effects. The authors also find evidence that susceptibility differs among types of students; in particular, the old-for-grade seventh graders were themselves especially susceptible to the influence of the concentration of other old-for-grade students in their school. Similar results were found when the outcome variable was the likelihood of being suspended.

3. *Truancy prevention.* The mix of students who are in the school building on any given day will be affected by absenteeism and tardiness. School attendance laws require that youths between specified ages (e.g., 7–16 in North Carolina and 5–18 in New Mexico) attend school, with possible exceptions for home schooling. This is a legal obligation for which both the child and parents are liable. In many school districts, however, these laws are widely flouted. For example, the absentee rate in DC public high schools in the 2006–2007 school year averaged 17 %.

The rate of unexcused absence determines not only the number of students in the school building but also the behavioral propensities of those students. Chronic truants are not a representative sample of the student body but rather tend to come from dysfunctional families and be at risk for delinquency, violence, and substance abuse (McCluskey et al. 2004). It is also true that chronic truancy engenders academic problems and is associated with failure to graduate from high school and a variety of poor life outcomes, including involvement with serious crime. As a logical matter, then, programs that are effective in improving attendance rates may have several effects. First, if they get delinquent youths off the street and into school, the result may be reduced crime rates in the community. Indeed, communities concerned about the daytime crimes committed by truants have increasingly enlisted the police and the juvenile court to combat truancy. Second, effective truancy-prevention programs may come at the cost of higher crime rates within the school. And third, if at-risk youths are persuaded to attend school more faithfully, the long-term result may be to improve their chances of graduation and subsequent success.

A number of school-based programs have been evaluated in part by their effect on school attendance. Unfortunately, there are no studies, insofar as we know, that evaluate the effect of attendance-promoting programs on school crime rates or overall (community plus school) crime rates.

4. *School desegregation.* In compliance with the 1954 Supreme Court ruling *Brown v. Board of Education*, federal courts issued a series of desegregation orders to public school districts. These orders forced a considerable increase in the extent to which African American students attended school with whites during the 1960s and 1970s. A vast literature on the effects of school segregation and desegregation has focused on academic outcomes. The results of this research offer support to the conclusion that integrated schools promote

black achievement and increase black high school graduation rates, college attendance and graduation rates, and occupational success (LaFree and Arum 2006). A persuasive quasi-experimental study of the effect of desegregation plans found that they reduced black dropout rates by 2–3 % points, with no detectable effect on whites (Guryan 2004). Given the tight link between academic success and school behavior, it is entirely plausible that the degree of segregation has a direct influence on delinquency in schools. But we are not aware of any direct evidence on the subject; segregation studies have not used school crime as an outcome variable. There have been two persuasive studies concerning the effects of segregation on crime outside of school. LaFree and Arum (2006) analyzed the incarceration rates for black males who moved to a different state following school. For any given destination state, they found that those who moved from a state with well-integrated schools had a substantially lower incarceration rate than those who moved from a state where the schools were more segregated. A more recent study (Weiner et al. 2009) utilizes a quasi-experimental approach in which the court desegregation orders serve as the experimental intervention: They report that these orders reduced black and white homicide victimization rates for 15–19-year-olds. The authors explore several mechanisms that may account for this result, including both the direct effects of changing the mix of students in the schools and indirect effects associated with police spending and relocation of some white students. In any event, since all but a handful of these homicides occurred outside of school, we are still waiting for direct evidence on crime in schools.

5. *Alternative schools.* A recent survey found that 39 % of public school districts administered at least one alternative school for students at risk of educational failure. As of October 2000, 613,000 students were enrolled in these schools, 1.3 % of nationwide

total enrollment. Urban districts, large districts (those with 10,000 or more students), districts in the Southeast, districts with high minority student enrollments, and districts with high poverty concentrations were more likely than other districts to have alternative schools and programs for at-risk students. Despite the widespread use of these schools as a means of removing antisocial and violent students from the regular classrooms, there have been no systematic studies of the effects on school crime rates. The effects on the behavior of youths who are given alternative-school placements have been studied, with mixed results. Indeed, there is unlikely to be any generic answer, since effects will depend on the quality of the programming and on which students are selected. Best-practice judgments tend to rely on expert opinion rather than on evaluation studies with strong designs.

6. *Grouping within schools.* Academic tracking is nearly universal in US secondary education. The attraction of separating students into tracks that are more or less demanding academically is the belief that this is the best way to tailor coursework to the differing background, ability, and motivation of the students. Tracking tends to have the result of concentrating minorities and students from lower socioeconomic status households in certain classrooms. Given the strong association between academic success and delinquency involvement, it also has the effect of concentrating crime-prone students, setting the stage for negative peer influence.

As a device to improve academic progress, tracking has had more detractors than advocates among education specialists. The evidence base is thin: Most notably, Mosteller et al. (1996) identified only ten randomized experimental evaluations comparing the performance of students in tracked (homogeneous) and untracked (heterogeneous) classrooms. When these studies were combined, the best estimate was a zero difference in average academic performance. Still thinner is any evidence on how tracking affects misbehavior. Thus, we conclude that

the possibility of deviant peer influence due to tracking is plausible but unproven.

Conclusion

School reform is typically shaped by theories of how to improve students' academic performance. But to the extent that school safety is an important goal, somewhat distinct from academic progress, the potential impacts on safety should be considered in any evaluation.

One of the most prominent reform efforts since 2001 has been the campaign funded largely by the Gates Foundation to create small high schools. While that effort was abandoned in 2008 as a result of disappointing results on academic progress, it would also be of interest to know the effect on school crime and juvenile delinquency. Given the fact, reported here, that small schools are not systematically safer than large schools (controlling for urbanicity and grade level), it appears doubtful that smaller is better in this domain.

There is very good reason to believe that the mix of students who are assembled in a school or any one classroom may influence the behavior of all. Two relevant mechanisms are deviant peer influence and "resource swamping," both implying that overall crime rates within school may increase in nonlinear fashion with the addition of deviant students to the mix (Cook and Ludwig 2006). This concern is relevant in evaluating policies regarding grade retention, truancy prevention, use of suspension and expulsion, use of alternative high schools, and even academic tracking. In each case, however, we found that relevant evaluations were lacking.

Implications for Research and Policy

Crime rates differ widely among schools. Experimental evidence suggests that the crime involvement of any given student at risk is influenced by the school that he or she attends. That fact motivates our scientific quest to find the school structural determinants of criminal activity by school-aged youths.

In another entry in this volume, we summarized findings from studies of school climate and crime and discussed numerous additional studies that reported on attempts to manipulate aspects of school climate. The school climate studies revealed sturdy associations between measures of school climate and measures of student delinquency, victimization, substance use, or other forms of problem behavior.

A starting point in accounting for interschool differences in crime is the criminal propensities of the students. Schools in which many of the students are active delinquents outside school start with a far greater challenge than those in which the students are largely law-abiding. The school crime rate of a student body with high crime propensity may be greater than the sum of the parts for two reasons. First, if the school lacks the adult resources to manage the "load" of misbehavior, then the school may become progressively more chaotic, spinning out of control. Further, delinquent and deviant youths may have a negative influence on each other and other students as well, further amplifying the problem. As noted in earlier discussion of the mechanisms involved in the production of crime and the features of schools that might influence these mechanisms, decisions that influence the demographic composition of schools are important because they determine the prevailing cultural beliefs in the school as well as the pool of youths from whom friends can be selected. In short, the crime rate in school is not just the sum of the parts but does reflect the ecological effects of the mix of students in the building.

Schools and school districts have a good deal of control over the makeup of the student body. Schools can be based on neighborhood residential patterns or integrated across race and class. The grade span for elementary and middle schools can be adjusted. Truancy and dropout prevention programs can be pursued with more or less vigor and troublesome students reassigned. Whether failing students are retained in a grade or given a social promotion influences the extent of age homogeneity within classrooms. Students who are enrolled in the school can be tracked on the basis of academic potential or

mixed together and so forth. This array of policy choices has the potential to influence the “load” on teachers and other adults and the opportunity for deviant peer influence. Some of these policies have been evaluated for these ecological effects, but the evidence base is quite thin.

We also note that most of the evaluations of policies that affect the mix of students – truancy and dropout prevention, alternative schools, tracking, grade retention of failing students, and so forth – only consider the effect on the students who are targeted and fail to consider the ecological effects. But secondary effects on other students may be quite important and should be included when it is possible to implement a comprehensive study.

Finally, we noted that evidence does not support the conclusion that smaller schools are more effective for limiting problem behaviors than larger schools, but research reported in the “School Social Organization, Discipline Management, and Crime” entry of this volume suggests that conditions that make a school environment “feel” smaller and more communally organized are related to levels of problem behavior.

Notes and Acknowledgments This essay is based on a more detailed chapter (Cook et al. 2010) on school crime control and prevention.

Related Entries

- ▶ [School Social Organization, Discipline Management, and Crime](#)
- ▶ [School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis](#)

References

- Carrell SE, Hoekstra ML (2008) Externalities in the classroom: how children exposed to domestic violence affect everyone’s kids. National Bureau of Economic Research (NBER) working paper no.14246. Available at <http://www.nber.org>
- Cook PJ, Gottfredson DC, Na C (2010) School crime control and prevention. In: Tonry M (ed) *Crime and justice: a review of research*. The University of Chicago Press, Chicago
- Cook PJ, Ludwig J (2006) Assigning youths to minimize total harm. In: Dodge KA, Dishion TJ, Lansford JE (eds) *Deviant peer influences in programs for youth: problems and solutions*. The Guilford Press, New York, pp 122–140
- Cook PJ, MacCoun R, Muschkin C, Vigdor J (2008) The negative impact of starting middle school in sixth grade. *J Policy Anal Manage* 27:104–121
- Dishion TJ, Dodge KA (2006) Deviant peer contagion in interventions and programs: an ecological framework for understanding influence mechanisms. In: Dodge KA, Dishion TJ, Lansford JE (eds) *Deviant peer influences in programs for youth: problems and solutions*. The Guilford Press, New York, pp 14–43
- Gottfredson DC, DiPietro SM (2011) School size, social capital, and student victimization. *Sociol Educ* 84(1):69–89
- Gottfredson GD, Gottfredson DC, Payne AA, Gottfredson NC (2005) School climate predictors of school disorder: results from a national study of delinquency prevention in schools. *J Res Crime Delinq* 42:412–444
- Guryan J (2004) Desegregation and black dropout rates. *Am Econ Rev* 94:919–943
- Hoffmann JP, Dufur MJ (2008) Family and school capital effects on delinquency: substitutes or complements? *Sociol Perspect* 51:29–62
- Jimerson SR, Pletcher SMW, Graydon K, Schnurr BL, Nickerson AB, Kundert DK (2006) Beyond grade retention and social promotion: promoting the social and academic competence of students. *Psychol Sch* 43:85–97
- LaFree G, Arum R (2006) The impact of racially segregated schooling on the risk of adult incarceration rates among U.S. cohorts of African-Americans and whites since 1930. *Criminology* 44:73–103
- Lee VE, Bryk AS, Smith JB (1993) The organization of effective secondary schools. In: Darling-Hammond L (ed) *Review of research in education*. American Educational Research Association, Washington, DC, pp 171–267
- McCluskey CP, Bynum TS, Patchin JW (2004) Reducing chronic absenteeism: an assessment of an early truancy initiative. *Crime Delinq* 50:214–234
- Mosteller F, Light RJ, Sachs JA (1996) Sustained inquiry in education: lessons from skill grouping and class size. *Harv Educ Rev* 66:797–842
- Muschkin C, Glennie B, Beck A (2008) Effects of school peers on student behavior: age, grade retention, and disciplinary infractions in middle school. *J Policy Anal Manage* 27:104–121
- Nagin DS, Pagani L, Tremblay RE, Vitaro F (2003) Life course turning points: the effect of grade retention on physical aggression. *Dev Psychopathol* 15:345
- Neiman S, DeVoe JF (2009) Crime, violence, discipline, and safety in U.S. public schools: findings from the school survey on crime and safety: 2007–08 (NCES 2009-326). National Center for Education Statistics,

- Institute of Education Sciences, U.S. Department of Education, Washington, DC
- Nolle KL, Guerino P, Dinkes R (2007) Crime, violence, discipline, and safety in US public schools: findings from the school survey on crime and safety: 2005-6 (NCES 2007-361). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC
- Stewart EA (2003) School social bonds, school climate, and school misbehavior: a multilevel analysis. *Justice Q* 20:575–604
- Weiner DA, Lutz BF, Ludwig J (2009) The effects of school desegregation on crime. National Bureau of Economic Research (NBER) working paper no.15380. Available at <http://www.nber.org>

School Violence Prevention

- [School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis](#)

School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis

Sandra Jo Wilson
Peabody Research Institute, Vanderbilt
University, Nashville, TN, USA

Synonyms

[School violence prevention](#)

Overview

Schools are an important location for interventions to prevent or reduce aggressive behavior. They are a setting in which much interpersonal aggression among children occurs and the only setting with almost universal access to children. There are many prevention strategies from which school administrators can and do choose, including surveillance (e.g., metal detectors, security guards), deterrence (e.g., disciplinary rules, zero

tolerance policies), and psychosocial programs. Indeed, over 75 % of schools in one national sample reported using one or more of these prevention strategies to deal with behavior problems (Gottfredson et al. 2000). Other reports similarly indicate that more than three-fourths of schools offer mental health, social service, and prevention service options for students and their families (Brenner et al. 2001). Among psychosocial prevention strategies, a broad array of programs can be implemented in schools. These include packaged curricula and homegrown programs for use school-wide and others that target selected children already showing behavior problems or deemed at risk for such problems. Most psychosocial prevention programs address a range of social and emotional factors assumed to cause aggressive behavior or to be instrumental in controlling it (e.g., social skills or emotional self-regulation), and it is these psychosocial programs that are the focus of this meta-analysis.

Key Issues

Various resources are available to help schools identify programs with proven effectiveness. Among these resources are the Blueprints for Violence Prevention; the Collaborative for Academic, Social, and Emotional Learning (CASEL); and the National Registry of Evidence-based Programs and Practices (NREPP) administered by the Substance Abuse and Mental Health Services Administration (SAMSHA). There is, however, little indication that the evidence-based programs promoted to schools through such sources have been widely adopted or that, when adopted, they are implemented with fidelity (Gottfredson and Gottfredson 2002).

While lists of evidence-based programs can provide useful guidance to schools about interventions likely to be effective in their settings, they are limited by their orientation to distinct program models and the relatively few studies typically available for each such program. A meta-analysis, by contrast, can encompass virtually all credible studies of such interventions and yield evidence about generic intervention

approaches as well as distinct program models. Perhaps most important, it can illuminate the features that characterize the most effective programs and the kinds of students who benefit most. Since many schools already have prevention programs in place, a meta-analysis that identifies characteristics of successful prevention programs can inform schools about ways they might improve those programs or better direct them to the students for whom they are likely to be most effective. Thus, the purpose of the meta-analysis reported here is to investigate which program and student characteristics are associated with the most effective treatments.

In 2003, we published a meta-analysis on the effects of school-based psychosocial interventions for reducing aggressive and disruptive behavior aimed at identifying the characteristics of the most effective programs (Wilson et al. 2003). That meta-analysis included 172 experimental and quasi-experimental studies of intervention programs, most of which were conducted as research or demonstration projects with significant researcher involvement in program implementation. Though not necessarily representative of routine practice in schools, those programs showed significant potential for reducing aggressive and disruptive behavior, especially for students whose baseline levels of antisocial behavior were already high. Different intervention approaches appeared equally effective, but significantly larger reductions in aggressive and disruptive behavior were produced by those programs with better implementation, that is, more complete delivery of the intended intervention to the intended recipients. In 2007, we published an updated version of that meta-analysis that included 249 studies of school-based prevention programs (Wilson and Lipsey 2007). For that analysis, we separated the programs into four mutually exclusive groups, characterized by the general service format and target population of the programs. The four program groups were universal programs, pullout programs for high-risk populations, comprehensive programs, and special school programs. In addition to being distinctive in terms of service format, programs in each of the groups had

a number of methodological, subject, and dosage differences that made it unwise to combine them in a single analysis. Furthermore, school decision-makers typically make choices from within a format category, rather than across them. Thus, there was little utility in lumping all the school-based prevention programs into one larger set.

Results of the 2007 meta-analysis found positive overall intervention effects on aggressive and disruptive behavior and other relevant outcomes. The most common and most effective approaches were universal programs and pullout programs targeted for high-risk children. Comprehensive programs did not show significant effects and those for special schools or classrooms were marginal. Different program approaches (e.g., behavioral, cognitive, social skills) produced largely similar effects. Effects were larger for better implemented programs and for those involving students at higher risk for aggressive behavior.

Since the publication of our earlier work, the analytic methods for handling dependent effect size estimates in meta-analysis have improved. Multiple effect size estimates from a single study sample can now be included in a meta-analysis using robust standard errors that account for the statistical dependencies present when multiple effect size estimates are used (Hedges et al. 2010). This entry provides an opportunity to take the school-based prevention programs included in the 2003 and 2007 papers and reanalyze their results using these new statistical methods.

Method

Criteria for Including Studies in the Meta-Analysis

Studies were selected for this meta-analysis based on a set of detailed criteria, summarized as follows:

1. The study was reported in English no earlier than 1950 and involved a school-based program for children attending any grade, prekindergarten through 12th grade.

2. The study assessed intervention effects on at least one outcome variable that represented either (a) aggressive or violent behavior (e.g., fighting, bullying, person crimes), (b) disruptive behavior (e.g., classroom disruption, conduct disorder, acting out), or (c) both aggressive and disruptive behavior.
3. The study used an experimental or quasi-experimental design that compared students exposed to one or more identifiable intervention conditions with one or more comparison conditions on at least one qualifying outcome variable. To qualify as an experimental or quasi-experimental design, a study was required to meet at least one of the following criteria:
 - Students or classrooms were randomly assigned to conditions.
 - Students in the intervention and comparison conditions were matched and the matching variables included a pretest for at least one qualifying outcome variable or a close proxy.
 - If students or classrooms were not randomly assigned or matched, the study reported both pretest and posttest values on at least one qualifying outcome variable or sufficient demographic information to describe the initial equivalence of the intervention and comparison groups.

Search and Retrieval of Studies

An attempt was made to identify and retrieve the entire population of published and unpublished studies that met the inclusion criteria summarized above. The primary source of studies was a comprehensive search of bibliographic databases, including PsycINFO, Dissertation Abstracts International, ERIC (Education Resources Information Center), US Government Printing Office publications, National Criminal Justice Reference Service, and MEDLINE. Second, the bibliographies of meta-analyses and literature reviews on similar topics were reviewed for eligible studies. Finally, the bibliographies of retrieved studies were themselves examined for candidate studies. Identified studies were retrieved from the library, obtained via

interlibrary loan, or requested directly from the author. We obtained and screened more than 95 % of the reports identified as potentially eligible through these sources. Note that a new search was not performed for this entry and that the literature reviewed here is current through 2004.

Coding of Study Reports

Study findings were coded to represent the mean difference in aggressive behavior between experimental conditions at the posttest measurement. The effect size statistic used for these purposes was the standardized mean difference, defined as the difference between the treatment and control group means on an outcome variable divided by their pooled standard deviation (Lipsey and Wilson 2001). In addition to effect size values, information was coded for each study that described the methods and procedures, the intervention, and the student samples. Coding reliability was determined from a sample of approximately 10 % of the studies that were randomly selected and recoded by a different coder. For categorical items, intercoder agreement ranged from 73 % to 100 %. For continuous items, the intercoder correlations ranged from 0.76 to 0.99. A copy of the full coding protocol is available from the author.

General Analytic Procedures

All effect sizes were multiplied by the small sample correction factor, $1 - (3/4n - 9)$, where n is the total sample size for the study, and each effect size was weighted by its inverse variance in all computations (Lipsey and Wilson 2001). The inverse variance weights were computed using the subject-level sample size for each effect size. Because many of the studies used groups (e.g., classrooms, schools) as the unit of assignment to intervention and control conditions, they involved a design effect associated with the clustering of students within classrooms or schools that reduces the effective sample size. The respective study reports provided no basis for estimating those design effects or adjusting the inverse variance weights for them, so cluster adjustments were not made in the analyses reported here. This should not greatly affect the

effect sizes estimates or the magnitude of their relationships to moderator variables but does assign them somewhat smaller standard error estimates and, hence, larger inverse variance weights than is technically correct. A dummy code identifying these cases was included in the analyses to reveal any differences in findings from these studies relative to those using students as the unit of assignment. The effect sizes for cluster assigned studies were not statistically different from the student-assigned studies in any analysis.

Examination of the effect size distributions for each of the four program format groups identified a small number of outliers with potential to distort the analysis. Outliers were defined as values that fell more than three interquartile ranges (IQR) above the 75th percentile or below the 25th percentile of the effect size distribution. Outliers identified using this procedure were Winsorized to the next closest value. In addition, several studies used unusually large samples. Because the inverse variance weights chiefly reflect sample size, those few studies would dominate any analysis in which they were included. Therefore, the extreme tail of the sample size distribution was Winsorized using the procedure described above, and the inverse variance weights were recomputed for those effect sizes. These adjustments allowed us to retain outliers in the analysis, but with less extreme values that would not exercise undue influence on the analysis results.

In the original publications (Wilson et al. 2003; Wilson and Lipsey 2007) only one effect size from each subject sample was used in the analyses to maintain independence of the effect size estimates. When more than one was available, the effect size from the measurement source most frequently represented across all studies (e.g., teachers' reports, self-reports) was selected. We wanted to retain informant as a variable for analysis, so we did not elect to average across effect sizes from different informants; if there was more than one effect size from the same informant or source, however, their mean value was used. However, techniques for computing robust standard errors for dependent effect sizes

have since been developed (Hedges et al. 2010) and are employed in the analyses reported below. Thus, when a study reported intervention effects for more than one aggressive or disruptive behavior outcome, all effect sizes were used in the analyses. Dummy codes indicating the type of aggressive behavior measured and/or the informant or source of the outcome measure were examined for their influence on the effect sizes in the analyses.

Finally, many studies provided data sufficient for calculating mean difference effect sizes on the outcome variables at the pretest. In such cases, we adjusted the posttest effect size by subtracting the pretest effect size value. This information was included in the analyses presented below to test whether there were systematic differences between effect sizes adjusted in this way and those that were not.

A small number of studies were missing data on the method, participant, or program variables used in the final analyses; missing values were imputed using an expectation-maximization (EM) algorithm in SPSS.

Analysis of the effect sizes was conducted separately for each program format (described below) and done in several stages. We first produced the random effects mean effect size for each group of effect sizes and examined the homogeneity of the effect size distributions using the Q-statistic (Hedges and Olkin 1985). Moderator analyses were then performed to identify the characteristics of the most effective programs using weighted mixed effects multiple regression with the aggressive/disruptive behavior effect size as the dependent variable. When the number of effect sizes and studies was sufficient, three models were produced in order to examine the added contribution of different sets of variables. In the first stage of these analyses, we examined the influence of study methods on the effect sizes. Influential method variables were carried forward as control variables for the next stage of analysis, which examined the relationships between student characteristics and effect size. The third stage in these analyses involved adding important treatment characteristics to the models. Because a large number of potential

moderators were available from our coding manual, moderators were selected based on the size of their weighted zero-order correlations with the effect size while being mindful of collinearity among the different moderator variables. That is, if two potential moderators were highly correlated with each other and with effect size, the one with the largest correlation with effect size was selected for the meta-regression models. Random effects analysis was used throughout but, in light of the modest number of studies in some categories and the large effect size variance, statistical significance was reported at the $\alpha=0.10$ level as well as the conventional 0.05 level.

Results

Program Format and Treatment Modality

The literature search and coding process yielded data from 283 independent study samples. Note that many studies reported results separately for different subgroups of students (most commonly, results for boys and girls were reported separately). We treated each subgroup as a separate study sample. The 283 study samples participated in a variety of prevention and intervention programs. For purposes of analyzing their effects on student aggressive/disruptive behavior, we divided the programs into four groups according to their general service format. Programs differ across these groups on a number of methodological, participant, and intervention characteristics that made it unwise to combine them in a single analysis. The four intervention formats are as follows:

- *Universal programs.* These programs are delivered in classroom settings to all the students in the classroom; that is, the children are not selected individually for treatment but, rather, receive it simply because they are in a program classroom. However, the schools with such programs are often in low socioeconomic status and/or high crime neighborhoods, and thus, the children in these universal programs may be considered at risk by virtue of their socioeconomic background or neighborhood context.
- *Pullout programs for high-risk students.* These programs are provided to students who are specifically selected to receive treatment because of conduct problems or some risk factor (typically identified by teachers for social problems or classroom disruptiveness). Most of these programs are delivered to the selected children outside of their regular classrooms (either individually or in groups), although some are used in the regular classrooms but targeted on the selected children.
- *Special schools or classes.* These programs involve special schools or classrooms that serve as the usual educational setting for the students involved. Children are placed in these special schools or classrooms because of behavioral or academic difficulties that schools do not want to address in the context of mainstream classrooms. Included in this category are special education classrooms for behavior disordered children, alternative high schools, and schools-within-schools programs.
- *Comprehensive/multimodal programs.* These programs involve multiple distinct intervention elements (e.g., a social skills program for students and parenting skills training) and/or a mix of different intervention formats. They may also involve programs for parents or capacity building for school administrators and teachers in addition to the programming provided to the students. Within the comprehensive service format, some programs are delivered universally while others are targeted toward high-risk groups. All but one of the programs in this subcategory includes services for both students and their parents.

General Study Characteristics

Nearly 90 % of the studies were conducted in the United States with over 75 % run by researchers in psychology or education. Table 1 shows additional characteristics of the studies, broken down by program format. Among universal programs, nonrandomized designs were most common. However, cluster randomization, in which schools or classrooms were randomly assigned to treatment conditions, was utilized with some

School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis, Table 1 Study and method characteristics for each program format

Variable	Universal		Pullout		Comprehensive		Special	
	k	%	k	%	k	%	k	%
<i>Method of assignment</i>								
Individual randomization	6	6.7	91	71.1	5	23.8	20	44.4
Cluster randomized	32	36.0	11	8.6	10	47.6	9	20.0
Nonrandom	51	57.3	26	20.3	6	28.6	16	35.6
<i>Year of publication</i>								
1970s	8	9.0	36	28.1	1	4.8	8	17.8
1980s	17	19.1	33	25.8	2	9.5	18	40.0
1990s	33	37.1	48	37.5	14	66.7	13	28.9
2000s	31	34.8	11	8.6	4	19.0	6	13.3
<i>Publication type</i>								
Journal article	57	64.0	75	58.6	17	81.0	22	48.9
Thesis, dissertation, other	32	36.0	53	41.4	4	19.0	23	51.1
Total studies	89		128		21		45	

k = number of independent study samples

frequency for both the universal and comprehensive programs, as would be expected given how participants were typically recruited for such studies. Individual random assignment was the norm for the pullout programs and common for the special school programs, though cluster randomization and nonrandomized designs were also present for these formats. Overall, fewer than 20 % of the studies were conducted prior to 1980 and most were published in peer-reviewed journals (60 %), with the remainder reported as dissertations, theses, conference papers, and technical reports.

The student samples reflect the diversity in American schools (see Table 2). Most were comprised of a mix of boys and girls, but there were some all boy samples (19 %) and a few all girl samples (6 %). Formats targeting higher-risk children (i.e., pullout and special school programs) tended to have larger proportions of boys. Minority children were well represented with over a third of the studies having primarily minority youth; nearly 30 % of the included studies, however, did not report ethnicity information, making it difficult to examine differential program effects for different ethnic groups. Interestingly, missing information on ethnicity was more common among the special school and pullout programs, both of which tended to serve higher-risk groups of students. All school ages

were included, from preschool through high school; the average age was around 9–10 for universal, pullout, and comprehensive programs and slightly older (about 13) for special school programs. A range of risk levels was also present, from generally low-risk students to those with serious behavior problems, and different risk levels were associated with the different program formats. Socioeconomic status was not widely reported, but a range of socioeconomic levels was represented among those studies for which it was reported.

Characteristics of the delivery and dosage of the programs, as well as the different program types, are shown in Table 3. Most studies were conducted as research or demonstration projects with relatively high levels of researcher involvement. Routine practice programs implemented by typical service personnel with less researcher involvement were less common. Program length varied by format, with comprehensive programs longest. Pullout and special school programs were the shortest, averaging 15 and 18 weeks, respectively. Universal programs averaged about 24 weeks. Service frequency also varied, with comprehensive and special school programs having larger proportions of daily contacts. Programs were generally delivered by teachers or the researchers themselves, though comprehensive programs tended to involve multiple types of

School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis, Table 2 Subject sample characteristics for each program format

Variable	Universal		Pullout		Comprehensive		Special	
	k	%	k	%	k	%	k	%
<i>Gender mix</i>								
No males	9	10.1	6	4.7	2	9.5	0	0.0
<50 % male	16	18.0	12	9.4	4	19.0	3	6.7
50–60 % male	51	57.3	30	23.4	6	28.6	12	26.7
>60 % male	3	3.4	47	36.7	6	28.6	23	51.1
All males	10	11.2	33	25.8	3	14.3	7	15.6
<i>Age</i>								
Average age	m	range	m	range	m	range	m	range
	9.2	4–16.5	10.3	4–16.0	9.1	5–16	12.7	4–16
<i>Predominant ethnicity</i>								
White	37	41.6	35	27.3	9	42.9	13	28.9
Black	23	25.8	30	23.4	5	23.8	11	24.4
Hispanic	10	11.2	6	4.7	3	14.3	1	2.2
Other minority	7	7.9	6	4.7	0	0.0	0	0.0
Missing	12	13.5	51	39.8	4	19.0	20	44.4
<i>Socioeconomic status</i>								
Mostly low SES	45	50.6	36	28.1	12	57.1	17	37.8
Middle + SES	44	49.4	39	30.5	8	38.1	15	33.3
Missing	0	0.0	53	41.4	1	4.8	13	28.9
<i>Risk level</i>								
General population	86	96.6	13	10.1	11	52.4	0	0.0
At-risk	0	0.0	80	62.5	8	38.1	30	66.7
Exhibiting problems	3 ^a	3.3	35	27.3	2	9.5	15	33.3

^aThree universal programs were delivered to entire classrooms, but results were presented only for the students exhibiting behavior problems

k = number of independent study samples

other delivery personnel. About 35 % of the reports mentioned some difficulties with the implementation of the program. This information, when reported, presented a great variety of relatively idiosyncratic problems, for example, attendance at sessions, dropouts from the program, turnover among delivery personnel, problems scheduling all sessions or delivering them as intended, wide variation between different program settings or providers, results from implementation fidelity measures, and the like. This necessitated use of a rather broad coding scheme in which we distinguished no problems indicated versus possible (some suggestion of difficulties but little explicit information) or definite problems with implementation.

The treatment modalities used in the four program formats varied. However, cognitive approaches and social skills training were common across all four service formats. Social cognitive strategies were cognitively oriented and focused on changing thinking patterns (e.g.,

hostile attributions) and developing social problem-solving skills (e.g., I Can Problem Solve; Shure and Spivack 1980). Anger management programs were also cognitively oriented but tended to focus on changing thinking patterns around anger and developing strategies for controlling angry impulses or coping with frustration (e.g., Coping Power; Lochman and Wells 2002). Social skills training focused on learning constructive behavior for interpersonal interactions, including communication skills and conflict management (Gresham and Nagle 1980). Also relatively common among the modalities were behavioral strategies that manipulated rewards and incentives (e.g., the Good Behavior Game, Dolan et al. 1993). Traditional counseling for individuals, groups, or families was also represented and tended to use a variety of therapeutic techniques (Nafpaktitis and Perlmutter 1998). The treatment coding was not mutually exclusive for any of the four program formats because many programs involved more than one

School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis, Table 3 Program characteristics for each program format

Variable	Universal		Pullout		Comprehensive		Special	
	k	%	k	%	k	%	k	%
<i>Routine practice</i>								
Research program	39	43.8	89	69.5	2	9.5	21	46.7
Demonstration program	38	42.7	24	18.8	18	85.7	17	37.8
Routine program	12	13.5	15	11.7	1	4.8	7	15.6
<i>Delivery personnel</i>								
Teacher	44	49.4	15	11.7	1	4.8	19	42.2
Researcher	17	19.1	57	44.5	1	4.8	12	26.7
Other personnel	28	31.5	48	37.5	19	90.5	14	31.1
<i>Program length</i>								
Average weeks	m	sd	m	sd	m	sd	m	sd
	24.3	32.6	15.4	17.1	54.0	34.5	17.9	12.6
<i>Program frequency</i>								
1× week or less	30	33.7	31	24.2	9	42.9	8	17.8
1–2× week	31	34.8	73	57.0	3	14.3	4	8.9
2–3× week	12	13.5	15	11.7	2	9.5	7	15.6
3–4× week	8	8.9	13	10.2	2	9.5	1	2.2
5× week	9	10.1	11	8.6	5	23.8	25	55.6
<i>Implementation</i>								
Problems = yes	32	36.0	36	28.1	12	57.1	12	26.7
<i>Program elements</i>								
Mean # of elements	m	sd	m	sd	m	sd	m	sd
	2.4	1.5	1.8	1.1	3.6	1.5	2.1	1.0
<i>Format</i>								
In-class (vs. out)	89	100.0	0	0.0	12	57.1	34	75.6
<i>Program types</i>								
Social cognitive program	58	65.1	39	30.5	7	33.3	11	24.4
Social skills training	42	47.2	28	21.9	11	52.4	11	24.4
Anger management	25	28.1	35	27.3	7	33.3	11	24.4
Behavioral program	8	9.0	31	24.2	6	28.6	14	31.1
Counseling	6	6.7	38	29.7	7	33.3	12	26.7
Academic services	2	2.2	4	3.1	6	28.6	17	37.8
Cognitive behavioral program	1	1.1	6	4.7	0	0.0	3	6.7

k = number of independent study samples

strategy. While many of the programs involved multiple program strategies, the comprehensive programs tended to have more strategies on average than the other three formats, and the selection of different components was distinctive for the comprehensive programs. The comprehensive programs tended to have multiple distinct components (e.g., a school-based cognitive component and a family-based component) and often had nonschool-based elements, while programs in the other three format categories that had multiple components tended to combine similar elements, like social cognitive and anger management components; they also tended to be solely school based.

The analyses reported below allow the 238 study samples to contribute multiple effect sizes

with different characteristics, which are shown in Table 4. The aggressive behavior outcomes were grouped into six categories. Violence outcomes involved clear-cut violent behavior, including hitting. Delinquency was generally not aggressive or violent and included arrests, police contacts, and self-reported criminal behavior. Aggressive behavior outcomes were those that assessed clear physically aggressive behavior. Measures of aggressive behavior were similar to those for violence, but generally less serious. Ideally, we would have liked to examine program effects on relatively distinct groups of outcome constructs. However, very few of the measures that called themselves aggressive behavior focused solely on physically aggressive interpersonal behavior. Many included disruptiveness,

School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis, Table 4 Effect size and measurement characteristics for each program format

Variable	Universal		Pullout		Comprehensive		Special	
	n = 175 n (ES)	k = 89 %	n = 272 n (ES)	k = 128 %	n = 72 n (ES)	k = 21 %	N = 139 n (ES)	K = 45 %
<i>Type of dependent variable</i>								
Violent behavior	13	7.4	11	4.0	11	15.3	12	8.6
Delinquency	17	9.7	26	9.6	11	15.3	36	25.9
Aggressive behavior	38	21.7	31	11.4	12	16.7	13	9.4
Aggression and disruptiveness mixed	53	30.3	93	34.2	26	36.1	60	43.2
Externalizing behavior	21	12.0	40	14.7	6	8.3	6	4.3
Undifferentiated problem behavior	33	18.9	71	26.1	6	8.3	12	8.6
<i>Informant, source of outcome</i>								
Self-report	50	28.6	34	12.5	29	40.3	41	29.5
Teacher report	68	38.9	126	46.3	15	20.8	41	29.5
Archive, record	18	10.3	43	15.8	2	2.8	33	23.7
Observation	19	10.9	31	11.4	9	12.5	11	7.9
Parent report	12	6.9	27	9.9	11	15.3	8	5.8
Other	8	4.6	11	4.0	6	8.3	5	3.6
<i>Measurement timing</i>								
Immediately after tx (vs. all later)	129	73.7	236	86.8	61	84.7	98	70.5
<i>Effect size computation</i>								
Means and sds (vs. all other)	130	74.3	199	73.2	50	69.4	105	75.5
<i>Effect size estimation</i>								
No estimate (vs. some)	155	88.6	227	83.5	59	81.8	117	84.2
<i>Effect size adjustment</i>								
Adjusted for pretest	145	82.9	232	85.3	56	77.8	86	61.9
<i>Attrition</i>								
Average pre- to post-attrition	13 %	15 %	10 %	14 %	14 %	17 %	18 %	19 %

n = number of effect sizes, k = number of independent study samples

acting out, and other forms of behavior problems that are negative, but not necessarily aggressive. These measures were placed in the aggression and disruption mixed category. Externalizing measures generally involved disruptive and acting behaviors, but did not include physical aggression. Finally, the undifferentiated problem behavior category was comprised of those measures that included both internalizing and externalizing behaviors in the same score (e.g., the Child Behavior Checklist). Self-reports and teacher reports were the most common sources of information about aggressive and disruptive behavior, but observations, archival information, and parent reports were also used. Treatment effects were typically measured immediately after treatment, but some studies had longer follow-up periods. Effect sizes were generally computed directly from means and standard deviations, but other statistics were used in

about 25 % of the cases. Direct computation from means and standard deviations or mathematically equivalent statistics was possible in over 80 % of the cases, and over 80 % of effect sizes had pretest effect sizes that allowed for adjustment. Attrition was moderate and averaged 10–18 % across the four program formats.

Overall Effects for School-Based Prevention Programs

The random effects weighted mean effect sizes and confidence intervals for the four program formats are shown in Table 5. The mean effect sizes for the universal, pullout, comprehensive, and special school programs are all positive and statistically significant, indicating that the participants in the school-based programs represented here had significantly lower levels of aggressive and disruptive behavior after the programs than students in comparison groups. The pullout

School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis, Table 5 Random effects mean effect sizes and confidence intervals for each program format

	Mean	95 %	<i>n</i>	<i>k</i>	Q_E	τ^2
	ES	CI				
Universal programs	0.16	(0.11, 0.22)	175	89	233	0.03
Pullout programs	0.37	(0.28, 0.47)	272	128	392	0.20
Comprehensive programs	0.06	(0.02, 0.11)	72	21	40	0.01
Special Schools programs	0.13	(0.03, 0.23)	139	45	91	0.07

programs evidenced the largest mean effect size of the four groups. The comprehensive programs had the smallest mean effect size at 0.06. The homogeneity statistics for the four groups of programs suggest that there is greater variability in the distributions of effect sizes than would be expected from sampling error. This variation was expected to be associated with the nature of the interventions, students, and methods used in these studies. The next step, then, is to examine some of the method, subject, and program characteristics that may be associated with the variability in treatment effects using meta-regression models.

Results for Universal Programs

There were 89 studies of universal programs in the database, all delivered in classroom settings to entire classes of students. Many studies tested program effects on more than one relevant outcome. Thus, the universal programs contributed 175 effect size estimates to the analysis. The random effects weighted mean effect size for universal programs was 0.16 ($p < 0.05$) and the distribution of effect sizes evidenced heterogeneity. The moderator analysis focused first on the relationship between study methods and the intervention effects using random effects inverse variance weights estimated via the method of moments. As mentioned above, moderator variables were selected based on their weighted zero-order correlations with effect size. This analysis is shown as Model I in Table 6.

Neither informant nor type of outcome measure was related to the effect size, so these variables were not included in the final models. Most notable is the lack of significant relationships between the study design variables and effect size. Method of assignment, pretest adjustments and computations, and attrition were not significantly related to effect size.

Our next step was to identify student characteristics that were associated with effect size while controlling for method variables. The results of this analysis are presented as Model II in the table. Only two student variables were significantly associated with effect size: gender mix and socioeconomic status. Programs with larger proportions of boys showed larger effects from universal programming than those with larger proportions of girls. Since boys tend to be more likely to exhibit aggressive, disruptive, or acting out behavior, this result is not surprising. Students with low socioeconomic status achieved significantly greater reductions in aggressive and disruptive behavior from universal programs than middle-class students ($p < 0.10$).

Model III in Table 6 shows the addition of the program characteristics to the models. Again, moderators were selected based on their correlations with effect size. In Model III, nonrandom assignment becomes active when the program variables are included, indicating that some of the variability remaining after controlling for subject and program characteristics was associated with method of assignment. Gender mix and socioeconomic status were still influential and associated with effect size. Several other variables in this analysis were also significant. Teacher and researcher delivery of programs had a significant relationship with effect size, with programs delivered by teachers or researchers having larger effects than those delivered by other personnel. In general, the other personnel were not school affiliated and included laypeople, social workers, and counselors. Well-implemented programs showed significantly larger effect sizes than those experiencing implementation problems. Dummy codes for the most common program types were also included in the model and none of those were statistically

School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis, Table 6 Random effects meta-regression models for universal programs

	<i>Model I</i>			<i>Model II</i>			<i>Model III</i>		
	<i>b</i>	<i>se</i>	<i>95 % CI</i>	<i>b</i>	<i>se</i>	<i>95 % CI</i>	<i>b</i>	<i>se</i>	<i>95 % CI</i>
<i>Methodological characteristics</i>									
Nonrandom assignment	0.05	0.09	(-0.13, 0.23)	0.10	0.11	(-0.11, 0.31)	0.26*	0.15	(-0.03, 0.55)
Cluster randomized	0.00	0.10	(-0.20, 0.20)	0.04	0.11	(-0.18, 0.25)	0.14	0.14	(-0.13, 0.41)
Pretest adjusted effect size	-0.11	0.08	(-0.26, 0.05)	-0.08	0.07	(-0.22, 0.06)	-0.04	0.08	(-0.19, 0.11)
ES computed with means/sds	-0.06	0.06	(-0.19, 0.06)	-0.10	0.06	(-0.22, 0.03)	-0.09	0.07	(-0.23, 0.05)
Percent attrition	-0.01	0.13	(-0.26, 0.24)	-0.10	0.13	(-0.36, 0.16)	0.08	0.15	(-0.22, 0.38)
<i>Subject characteristics</i>									
Gender mix				0.04**	0.02	(0.01, 0.08)	0.03*	0.02	(0.00, 0.07)
Average age				-0.02	0.01	(-0.04, 0.00)	0.00	0.01	(-0.03, 0.02)
Predominantly low SES				0.10*	0.05	(0.00, 0.20)	0.18**	0.07	(0.04, 0.32)
<i>Program characteristics</i>									
Teacher delivery of program							0.11*	0.06	(-0.01, 0.23)
Researcher delivery of program							0.24**	0.10	(0.04, 0.44)
Implementation problems = yes							-0.15*	0.07	(-0.29, 0.00)
Frequency of tx sessions							0.00	0.02	(-0.03, 0.04)
Social cognitive program							-0.02	0.07	(-0.15, 0.11)
Social skills training							-0.10	0.07	(-0.25, 0.04)
Anger management program							0.07	0.08	(-0.08, 0.22)
Behavioral program							0.11	0.13	(-0.15, 0.37)

Mixed effects meta-regression models estimated with robust standard errors to account for dependent effect sizes estimates (n = 175; k = 89). b - unstandardized regression coefficient

**p* < 0.10

***p* < 0.05

significant. The different program types all produced positive effects on the outcomes and were not appreciably different from each other in effectiveness.

Results for Pullout Programs for High-Risk Students

There were 128 studies of pullout programs that targeted interventions to individually identified children. These studies contributed 272 effect sizes. The overall random effects mean effect size for these programs was 0.37 (*p* < 0.05), nearly twice as large as the average effect size for universal programs. The homogeneity test of the effect sizes showed significant variability across studies, and our analysis of

the relationships between effect size and methodological and substantive characteristics of the studies proceeded much the same as for the universal programs (see Table 7). First, Model I examined the influence of each method variable on the aggressive/disruptive behavior effect sizes. Here also the study design was not associated with effect size—random assignment studies did not show appreciably smaller or larger effects than nonrandomized studies. Note that for the pullout programs, the design contrast was primarily between individual-level randomization and nonrandomization; there were only 11 cluster randomized studies. The two method variables that did show significant relationships with effect size were the

School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis, Table 7 Random effects meta-regression models for pullout programs for high-risk students

	<i>Model I</i>			<i>Model II</i>			<i>Model III</i>		
	<i>b</i>	<i>se</i>	<i>95 % CI</i>	<i>b</i>	<i>se</i>	<i>95 % CI</i>	<i>b</i>	<i>se</i>	<i>95 % CI</i>
<i>Methodological characteristics</i>									
Aggressive behavior outcome	-0.27	0.17	(-0.60, 0.06)	-0.22	0.15	(-0.53, 0.08)	-0.17	0.16	(-0.48, 0.14)
Problem behavior outcome	0.20**	0.09	(0.01, 0.38)	0.20*	0.09	(0.02, 0.39)	0.25**	0.10	(0.06, 0.45)
Externalizing behavior outcome	0.09	0.12	(-0.14, 0.33)	0.17	0.14	(-0.11, 0.45)	0.29**	0.14	(0.01, 0.58)
Random assignment	-0.02	0.09	(-0.21, 0.16)	-0.10	0.09	(-0.29, 0.08)	-0.06	0.10	(-0.25, 0.14)
Percent attrition	-0.93**	0.28	(-1.48, -0.39)	-1.02**	0.29	(-1.59, -0.44)	-1.07**	0.32	(-1.69, -0.44)
<i>Subject characteristics</i>									
Gender mix				0.01	0.04	(-0.08, 0.10)	0.02	0.05	(-0.07, 0.12)
Average age				0.01	0.02	(-0.02, 0.05)	0.01	0.02	(-0.03, 0.05)
Risk level				0.16**	0.04	(0.07, 0.24)	0.18**	0.04	(0.10, 0.27)
<i>Program characteristics</i>									
Researcher delivery of program							-0.21*	0.12	(-0.45, 0.04)
Other tx personnel							-0.15	0.12	(-0.39, 0.08)
Routine practice program							0.08	0.07	(-0.05, 0.21)
Implementation problems = yes							-0.36**	0.10	(-0.56, -0.16)
Frequency of tx sessions							0.01	0.02	(-0.03, 0.05)
Anger management							0.04	0.15	(-0.25, 0.33)
Social cognitive program							-0.02	0.14	(-0.29, 0.24)
Counseling program							0.11	0.12	(-0.13, 0.34)
Behavioral program							0.10	0.12	(-0.13, 0.33)
Social skills training							0.24	0.15	(-0.07, 0.54)

Mixed effects meta-regression models estimated with robust standard errors to account for dependent effect sizes estimates ($n = 272$; $k = 128$). *b* - unstandardized regression coefficient

* $p < 0.05$

** $p < 0.1$

outcome measures of undifferentiated problem behavior and attrition. Outcome measures that included both internalizing and externalizing problem behavior tended to produce larger effect sizes than the other outcomes (all more closely associated with aggressive behavior). Attrition was associated with smaller effect sizes.

Model II in Table 7 shows the addition of the subject characteristics to the analysis. One student characteristic had a significant relationship with effect size in this model. Higher-risk subjects showed larger effect sizes than lower-risk subjects though, with the pullout programs, very few low-risk children were involved. The

distinction here is mainly between indicated students who are already exhibiting behavior problems and selected students who have risk factors that may lead to later problems. Nevertheless, programs with the highest-risk students tended to have larger effects.

Model III includes the characteristics of the intervention programs. In contrast to the universal programs, researchers and other non-teacher personnel as treatment delivery agents were less effective than teachers. Programs with higher-quality implementation were associated with larger effects. None of the different program types was significantly better or worse than any other.

Results for Comprehensive or Multimodal Programs

There were only 21 studies of comprehensive programs in the database, distinguished by their multiple treatment components and formats. These programs contributed 72 effect sizes indexing treatment effects on aggressive and disruptive behavior. The average number of distinct treatment components for these programs was nearly four, whereas the universal and pullout programs typically employed one or two treatment components. The studies of comprehensive programs tended to involve larger samples of students than the other program formats and a larger proportion of cluster randomizations as well. Thus, the statistical significance of the mean effect size of 0.06 is likely overstated. Comprehensive programs were generally longer than the universal and pullout programs. The modal program covered an entire school year and almost half of the programs were longer than 1 year. In contrast, the average program lengths for universal and pullout programs were 24 and 15 weeks, respectively.

Students who participated in comprehensive programs were slightly better off than students who did not, though the effect size was small. The Q has relatively low statistical power with small numbers of studies; therefore, despite the relative homogeneity in relation to the other program formats, we ran a single meta-regression model to examine the influence of some key study characteristics. This model is shown in Table 8. Because of the limited number of studies available in this category, only a few variables were tested. None of the method characteristics had correlations with effect size above 0.20 and were not tested. We also examined student risk level and found that it was not associated with effect size. Overlapping with the risk level variable was the variable we tested in the 2007 paper which compared universally delivered programs to those targeting higher-risk groups in a pullout format. Like the risk variable, this variable was not associated with effect size. Programs with implementation problems tended to have smaller effects than those without such problems, but the relationship was not significant. However, the

School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis, Table 8 Random effects meta-regression model for comprehensive programs

	<i>Model I</i>		
	<i>b</i>	<i>se</i>	<i>95 % CI</i>
<i>Methodological characteristics</i>			
Nonrandom assignment	-0.04	0.05	(-0.14, 0.06)
<i>Subject characteristics</i>			
Risk level	0.00	0.01	(-0.02, 0.03)
<i>Program characteristics</i>			
Implementation problems = yes	-0.04	0.03	(-0.12, 0.03)
Frequency of tx sessions	0.02*	0.01	(0.01, 0.03)
Mean # components	-0.04*	0.01	(-0.06, -0.02)

Mixed effects meta-regression models estimated with robust standard errors to account for dependent effect sizes estimates (n = 72; k = 21). b - unstandardized regression coefficient

*p < 0.05

frequency of treatment and the number of different treatment components were both associated with effect sizes. More frequent treatments produced larger effects on aggressive and disruptive behavior, while programs with more components produced smaller effects. The comprehensive programs with larger numbers of distinct components tended to be less effective than those with fewer components.

Results for Special Schools or Classes

There were 45 studies of programs delivered in special schools or classrooms. These programs generally involved an academic curriculum plus programming for social or aggressive behavior. The students typically had behavioral (and often academic) difficulties that resulted in their placement outside of mainstream classrooms. The mean aggressive/disruptive behavior effect size for these programs was 0.13 (p < 0.05). The Q-statistic (Q_e = 91) indicates that the distribution of effect sizes was heterogeneous. About 20 % of the studies of special programs assigned students to intervention and control conditions at the classroom level. As a result, there may be a design effect associated with the

School-Based Interventions for Aggressive and Disruptive Behavior: A Meta-Analysis, Table 9 Random effects meta-regression models for the special school programs

	<i>Model I</i>			<i>Model II</i>			<i>Model III</i>		
	<i>b</i>	<i>se</i>	<i>95 % CI</i>	<i>b</i>	<i>se</i>	<i>95 % CI</i>	<i>b</i>	<i>se</i>	<i>95 % CI</i>
<i>Methodological characteristics</i>									
Delinquency outcome	-0.19*	0.09	(-0.37, -0.01)	-0.18*	0.08	(-0.34, -0.01)	-0.15*	0.08	(-0.31, 0.00)
Random assignment	-0.02	0.10	(-0.22, 0.18)	-0.01	0.11	(-0.23, 0.21)	0.00	0.09	(-0.18, 0.18)
<i>Subject characteristics</i>									
Average age				0.00	0.02	(-0.03, 0.04)			
Risk level				0.14	0.13	(-0.12, 0.40)			
<i>Program characteristics</i>									
Implementation problems = yes							-0.22*	0.09	(-0.40, -0.04)
Length of program							-0.01	0.00	(-0.02, 0.00)
# of components							0.14*	0.06	(0.01, 0.27)
In-class versus out of class							0.11	0.09	(-0.08, 0.30)
Anger management							0.17	0.14	(-0.13, 0.46)
Social cognitive program							-0.24	0.15	(-0.55, 0.06)
Counseling							0.05	0.08	(-0.12, 0.21)

Mixed effects meta-regression models estimated with robust standard errors to account for dependent effect sizes estimates ($n = 139$; $k = 45$). b - unstandardized regression coefficient

* $p < 0.05$

clustering of students within classrooms that overstates the significance, though the overall effect size and the regression coefficients presented below should not be greatly affected.

The limited number of studies available for this program format necessitated that we examine fewer moderator variables in the meta-regression models (See Table 9). The two method characteristics with the highest weighted zero-order correlations with effect size were random assignment and delinquency outcomes. Studies employing individual randomization did not produce effect sizes that were smaller or larger than studies that did not use random assignment. When program effects were assessed on delinquency outcomes for the special school programs, effect sizes were significantly smaller than when any other outcome variable was used. Note that all of the other outcomes contained some form of aggressive, violent, or acting out/disruptive behavior, while the delinquency outcomes included arrests, police contacts, and self-reported crime.

In Model II, two subject characteristics were tested, average age and risk level. Neither age nor risk level were associated with effect size. These subject characteristics were dropped from Model III to allow us to test more different treatment characteristics. The length of the program did not have a significant relationship to effect size. In one form of the special school programs, students were assigned to special education classes or schools and the program was delivered entirely in the classroom setting. The other form involved students in special education classrooms who were pulled out of class for additional small group treatments. The programs delivered in classroom settings did not produce different effects than the pullout programs. Also, as in other analyses, better implemented programs showed larger effects. Three treatment modalities were tested in this model, anger management, social cognitive programs, and counseling. None were found to be significantly more effective than any of the other strategies.

Summary and Conclusions

The primary issue addressed in this entry is the effectiveness of programs for preventing or reducing such aggressive and disruptive behaviors as aggression, fighting, bullying, name-calling, intimidation, acting out, and disruptive behaviors occurring in school settings. The main finding is that, overall, the school-based programs that have been studied by researchers (and often developed and implemented by them as well) generally have positive effects for this purpose. A secondary purpose of this entry was to utilize new analytic techniques that allow for multiple dependent effect sizes from the same study to be included in an analysis (Hedges et al. 2010). Several differences from our original work are noteworthy. The mean effect size using the new techniques for universal programs was 0.16, while it was 0.21 in Wilson and Lipsey (2007). The mean effect size for the pullout programs was 0.37, quite a bit larger than the 0.29 published in Wilson and Lipsey (2007). It is well known that selecting a single effect size from each study is not ideal in meta-analysis; this reanalysis dramatizes how much difference such selections can make on results and conclusions.

The most common and most effective approaches are universal programs delivered to all the students in a classroom or school and pullout programs targeted for at-risk children who participate in the programs outside of their regular classrooms. The universal programs that were included in the analysis mainly used social cognitive approaches, anger management, and social skills training, and, like the original analysis, all strategies appeared about equally effective. In the 2007 meta-regression model for the universal programs, only SES was significant. The new analyses found that implementation quality and the type of delivery personnel were also important predictors of treatment effects.

Cognitively oriented approaches (both social cognitive programs and anger management) and social skills training were also the most frequent among the pullout programs, but many did use behavioral or counseling treatment modalities as well. In the original analysis, behavioral

strategies were associated with larger treatment effects. However, differential use of the different modalities within the pullout format was not associated with differential effects in the new analysis. This suggests that it may be the pullout program format that is most important but does not rule out the possibility that the treatment modalities used with that format are especially effective ones. It is also possible that the larger effect size with the pullout programs was due to the generally higher-risk students participating in these programs. Higher-risk students generally achieve better outcomes from psychosocial programming than lower-risk students, so the larger effect size for the more targeted pullout programs may be partly due to the students with more serious behavior problems having greater room for improvement.

The mean effect sizes of 0.16 and 0.37 for the universal and pullout programs, respectively, represent a decrease in aggressive/disruptive behavior that is not only statistically significant but likely to be of practical significance to schools as well. Suppose, for example, that approximately 20 % of students are involved in some version of such behavior during a typical school year. This is a plausible assumption according to the *Indicators of School Crime and Safety: 2005* which reports that 13 % of students age 12–18 were in a fight on school property, 12 % had been the targets of hate-related words, and 7 % had been bullied (DeVoe et al. 2004). Effect sizes of 0.16 and 0.37 represent reductions from a base rate prevalence of 20 % to about 14 % and 8 %, respectively, that is, 30–40 % reductions in aggressive and disruptive behavior. The programs of above average effectiveness, of course, produce even larger decreases.

The substantial similarity in effects across the different types of programs within the universal and pullout formats suggests that schools may choose from a range of such programs with some confidence that whatever they pick will be about as effective as any other choice. In the absence of evidence that one modality is significantly more effective at reducing aggressive and disruptive behavior than another, schools might benefit most by considering ease of

implementation, compatibility with school culture, and teacher and administrator training and preferences when selecting programs and focusing on implementation quality once programs are in place. Our coding of implementation quality, albeit crude, was associated with larger effect sizes for all four treatment formats. A very high proportion of the studies in this meta-analysis, however, were research or demonstration projects in which the researchers had a relatively large direct influence on the service delivery. Schools adopting these programs without such engagement may have difficulty attaining comparable program fidelity, a concern reinforced by evidence of frequent weak implementation in actual practice (Gottfredson and Gottfredson 2002). The best choice of a universal or pullout program for a school, therefore, may be the one they are most confident they can implement well. In addition, for the universal programs, teachers and researchers as delivery personnel were generally more effective than nonschool personnel. Selecting programs that fit with teachers' interests, training, and existing curricula may help to keep implementation fidelity as high as possible.

Another significant factor that cut across the universal and targeted pullout programs was the relationship of student characteristics to program effects. Larger treatment effects were achieved with higher-risk students, even among the already higher-risk samples in the targeted pullout programs. For the universal programs, the greatest benefits appeared for students from economically disadvantaged backgrounds and for groups with larger proportions of boys. Universal programs did not specifically select students with individual risk factors or behavior problems, though many students were of low socioeconomic status and there were most likely some behavior problem students in the classrooms that received universal interventions. These findings reinforce the fact that a program cannot have large effects unless there is sufficient problem behavior, or risk for such behavior, to allow for significant improvement.

The programs in the category we called comprehensive, in contrast to the universal and pullout programs, were surprisingly ineffective.

On the face of it, combinations of universal and pullout treatment elements and multiple intervention strategies would be expected to be at least as effective, if not more so, than less multifaceted programs. The small mean effect size for the comprehensive programs raises questions about their value. It should be noted, however, that there were few programs in this category and that most of these were long-term school-wide programs. It may be that this broad scope is associated with some dilution of the intensity and focus of the programs so that students have less engagement with them than with the programs in the universal and pullout categories. The comprehensive programs also suffered from implementation difficulties more frequently than the other formats (see Table 3), although implementation was not significantly associated with effect size in the reanalysis or in the original. However, the reanalysis found that the comprehensive programs with the largest number of treatment components were the least effective in this category, a finding that may be an indicator for implementation problems. This is an area that clearly warrants further study.

The most distinctive programs in our collection were those for students in atypical school settings. The mean effect size for these programs was modest, though statistically significant. As in the original analysis, implementation was important. The results also are somewhat anomalous. One of the signal characteristics of students in these settings is a relatively high level of behavior problems or risk for such problems; thus, there should be ample room for improvement. On the other hand, the special school settings in which they are placed can be expected to already have some programming in place to deal with such problems. The control conditions in these studies would thus reflect the effects of that practice-as-usual situation with less value added provided by additional programming of the sort examined in the studies included here. Alternatively, however, the add-on programs studied in these cases may have been weaker or less focused than those found in the pullout format, or the more serious

behavior problems of students in these settings may be more resistant to change. Here too are issues that warrant further study.

A particular concern of our original meta-analysis (Wilson et al. 2003) was the smaller effects of routine practice programs in comparison to those of the more heavily represented research and demonstration programs. Routine practice programs are those implemented in a school on an ongoing routine basis and evaluated by a researcher with no direct role in developing or implementing the program. Research and demonstration programs are mounted by a researcher for research or demonstration purposes with the researcher often being the program developer and heavily involved in the implementation of the program, though somewhat less so for demonstration programs. In the present meta-analysis, somewhat more studies of routine programs were included, and it is reassuring that their mean effect sizes were not significantly smaller. The zero-order correlation between routine practice and effect sizes was only large enough to be included in the meta-regression models for the pullout programs, and it did not turn out to have a significant relationship with effect size in the model in which it was tested. However, only 35 of the 283 studies in this meta-analysis examined routine practice programs. This number dramatizes how little evidence exists about the actual effectiveness, in everyday real-world practice, of the kinds of school-based programs for aggressive/disruptive behavior represented in this review.

Notes and Acknowledgments The research reported in this entry was supported by grants from the National Institute of Mental Health, the National Institute of Justice, and the Centers for Disease Control and Prevention.

A bibliography of studies used in the meta-analysis is available from the first author.

Related Entries

- ▶ Juvenile Violence
- ▶ School Crime Statistics
- ▶ School Security Practices and Crime

- ▶ School Social Organization, Discipline Management, and Crime
- ▶ School Structural Characteristics and Crime

Recommended Reading and References

- Brener ND, Martindale J, Weist MD (2001) Mental health and social services: results from the school health policies and program study. *J Sch Heal* 71:305–312
- DeVoe JF, Peter K, Kaufman P et al (2004) Indicators of school crime and safety: 2004 (Report Num. NCES 2005-002/NCJ 205290) Washington, DC: U.S. Departments of Education and Justice
- Dolan LJ, Kellam SG, Brown CH, Werthamer-Larsson L, Rebok GW et al (1993) The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *J Appl Dev Psychol* 14:317–345
- Gottfredson DC, Gottfredson GD (2002) Quality of school-based prevention programs: results from a national survey. *J Res Crime Delinquency* 39:3–35
- Gottfredson GD, Gottfredson DC, Czeh ER, Cantor D, Crosse S, Hantman I (2000) National study of delinquency prevention in schools. Final Report, Grant No. 96-MU-MU-0008. Ellicott City, MD: Gottfredson Associates, Inc. Available from: www.gottfredson.com
- Gresham FM, Nagle RJ (1980) Social skills training with children: responsiveness to modeling and coaching as a function of peer orientation. *J Consult Clin Psychol* 48:718–729
- Hedges LV, Olkin D (1985) *Statistical methods for meta-analyses*. Academic Press, San Diego
- Hedges LV, Tipton E, Johnson MC (2010) Robust variance estimation in meta-regression with dependent effect size estimates. *Res Syn Method* 1:39–65
- Lipsey MW, Wilson DB (2001) *Practical meta-analysis*. Sage, Thousand Oaks
- Lochman JE, Wells KC (2002) The coping power program at the middle-school transition: universal and indicated prevention effects. *Psychol Addict Behav* 16(suppl 4): S40–S54
- Nafpaktitis M, Perlmutter BF (1998) School-based early mental health intervention with at-risk students. *Sch Psychol Rev* 27:420–432
- Shure MB, Spivack G (1980) Interpersonal problem solving as a mediator of behavioral adjustment in pre-school and kindergarten children. *J Appl Dev Psychol* 1:29–44
- Wilson SJ, Lipsey MW (2007) School-based interventions for aggressive and disruptive behavior: update of a meta-analysis. *Am J Prev Med* 33(2S):S130–S143
- Wilson SJ, Lipsey MW, Derzon JH (2003) The effects of school-based intervention programs on aggressive and disruptive behavior: a meta-analysis. *J Consult Clin Psychol* 71:136–149

Scientific Basis of the Forensic Process

William J. Tilstone
Centre for Forensic Science, University of
Strathclyde, Glasgow, Scotland, UK

Overview

The chapter reviews the contributions to the History and Philosophy of Science (HPS) that impact on the scientific basis of the forensic process. There is a close parallel between the scientific method and the forensic process, in that each consists of hypothesis creation, testing and review based on observations and rigorous challenge. The leaders in HPS whose thoughts are most relevant are Popper, for his views on the limitations of inductive reasoning, Peirce, for introducing the concept of abduction, Kuhn for his view on “normal science” and the sociology of scientific developments, and Lakatos, for formalizing Popper’s views into what he termed research programs.

Having established the scientific foundation of the forensic process, the chapter turns to how it is used to convert “science” into knowledge in the form of evidence arising from an investigation and of relevance to the determination of facts by a tribunal. The fundamental types of forensic evidence, classified by purpose, are described, leading to a taxonomy of forensic evidence based on the principles of rhetoric.

Is It Science?

There is no definition of “forensic science” that addresses the question of its scientific basis. Most definitions are constructed around the concepts of “science applied to the law” and “working carefully”, but these beg more questions than they answer. Is parentage testing for immigration purposes forensic? What about dope testing of Olympic athletes, or race horses? Is investigation of environmental spills forensic? Where does

pre-employment drug screening fit? And “working carefully” encompasses pilots, accountants, and surgeons, but that clearly does not make them forensic.

Adopting the approach that it is something conducted by scientists in the laboratory does not help much either. For example, aspects of a homicide investigation that could be classed as “forensic” may be conducted at the scene by personnel as diverse as those who are serving police officers and whose qualification consists of successful completion of high school together with in-service training, and by toxicologists and DNA analysts with a Ph.D. degree working in a university or private laboratory.

Not too long ago, such questions would have seemed too academic and abstract for a work on forensic science, the forensic literature was dominated by case reports and papers describing new or improved technical methods. That should never have been so and is certainly less so today. The relatively recently published report by the National Academies of Science on forensic science in the US was highly critical of forensic science in general, and aspects of crime scene-related criminalistics such as tool marks and fingerprints in particular, for not being well-founded in science. The UK too has embraced basic scientific philosophy in evaluation of forensic evidence, where proposed legislation on the admissibility of expert evidence requires that a trial judge may refuse to admit expert evidence that is based on a hypothesis which has not been subjected to scrutiny (including experimental or other testing) or has failed to stand up to scrutiny.

However, the main unanswered question, other than a superficial passing mention in the NAS report, is what exactly **is** science?

This chapter addresses these issues by first of all exploring what science really is, then defining and exploring the forensic process, and thereby demonstrating how the forensic process is indeed fundamentally a scientific endeavor, with all the advantages and limitations that arise from that. The chapter ends by addressing the question “does it matter whether or not it is based in science” and explores some consequences.

Science, Technology, and the Scientific Method

The word “science” is derived from the Latin verb *scire*, to know, and its noun *scientia*, knowledge. Although the principles of the scientific method can and have been applied to other areas including psychology and the social sciences, the accepted usage today is knowledge of the natural world. Hence the generally accepted perspective of forensic science as a source of unqualified and independent information, but the history of the scientific method and the writings of the philosophers of science tell us that science is not absolute.

The ancient Chinese and early Muslim societies contributed many discoveries and inventions relevant to natural laws, but “knowledge”, including scientific knowledge, in western societies is built on the foundations provided by the ancient Greeks and Romans. As a result, up until the early sixteenth century, philosophers formulated what they regarded as laws of nature based on the universal belief that the Earth was the center of the Universe. Thus it was “known” that the Earth was at rest in the centre of the universe. This was a natural and inevitable conclusion arising from the form of deductive logic that prevailed at the time, namely that having accepted certain propositions we can increase our knowledge by using those to predict outcomes, and so new propositions come into being.

This changed dramatically when Nicolaus Copernicus (1473–1543) decided to test the accepted knowledge by carrying out observations on the heavenly bodies to see if they behaved as predicted. They did not, and the scientific method was born.

The next century reinforced the importance of inductive logic in developing our knowledge of the natural world. Francis Bacon (1561–1626) described what is still today accepted as the essence of the scientific method, namely the cycle of observations to determine facts, followed by application of inductive reasoning to formulate a hypothesis that explains the observations, followed by further experimentation to test the hypothesis, resulting in its acceptance, refinement or rejection. Bacon’s ideas were

central to the subsequent three centuries of scientific enlightenment, and stimulated interest in science as an area of study in its own right.

The core of Bacon’s scientific method is the role of the hypothesis. It is the link between observation and interpretation, and one that requires the utmost rigor in its application. Contrary to the popular expression, facts do not speak for themselves. The same observations can satisfy many interpretations, and many of the advances in what can be regarded as “scientific” are concerned with the relationship between observation, hypothesis, and interpretation. A fact is something known or proven to be true and may therefore be assumed to be the starting point or end point of the cycle, but “facts” themselves may indeed not be objective, proven, immutable entities, but may be subjective perceptions or interpretations of some consequence of a true fact.

The work of the American philosopher of science Charles Sanders Peirce (Peirce 1934) is central to the understanding of the strengths and weaknesses of the science that underpins forensic science, especially in regard to current controversies regarding contextual bias. Before Peirce, scientific enquiry was either *inductive*, that is, making inferences from a particular set of phenomena in order to predict and make propositions of more general character (empiricism), or *deductive*, that is, inferring from accepted truths and using general laws to deduce outcomes (rationalism). He introduced a third option, “abduction,” using the formulation and testing of hypotheses to explore possible explanations for a surprising event or occurrence, one in which there is no existing set of phenomena. Abduction therefore gives us an alternative entry point into the scientific method, and of course in this context criminal offences are almost always surprising events.

In practical terms the hypothesis is the scientist’s best guess of the character of phenomena that he or she is going to investigate. Intuitively, the validity or strength of any guess will be proportionate to the quality and extent of the information on which it is based. Abduction is therefore counter-intuitive but provides the most

fertile although least secure mode of inference. The hypothesis, being based on a “surprise event” and lacking a foundation of pre-existing phenomena, needs to have practical implications that can be rigorously tested.

Peirce himself was aware of the limitations of abduction as well as its advantages. He expressed it thus: “Deduction proves that something must be; Induction shows that something actually is operative; Abduction merely suggests that something may be.”

A functional definition of forensic science is that it is concerned with identifying objects at a crime scene that can be used to provide a credible reconstruction of the event (Tilstone et al. 2012), and induction, abduction and deduction are all essential elements of the process in deriving that reconstruction. The surprise event is the crime itself and trying to explain what happened begins with the formulation of an explanatory hypothesis, followed by its testing and refinement. However, the information available as a consequence of that one event, the crime, is an insufficient basis for the conventional inductive approach to hypothesis formulation. On the other hand abduction is an imaginative, “what if,” process that enjoys exploring the unexpected, captures the intuitive actions of the crime scene investigator, and can lead to formulation of one or more credible hypothesis. Peirce established abduction as a legitimate element of the scientific method, but the imaginative beginning must be balanced by rigorous, objective, scrutiny of alternate hypotheses and testing of their validity.

How then can a hypothesis be evaluated with a view to establish whether it is a matter of fact and not just a possible explanation? The cycle of the scientific method demonstrates that science is not an absolute, but a continuous process of searching for the truth, and therefore cannot prove any hypothesis to the exclusion of everything else. Karl Popper (Popper 1970) was a key figure in the debate. He held that all observation is from a point of view, and therefore any conclusion drawn from observations will be colored consciously or sub-consciously by our existing understanding. This by the way was 40–50 years

before the “discovery” of contextual bias in the forensic sciences.

Popper’s critical rationalism challenged the prevailing thinking in science, which held that a theory could be verified and thus acknowledged as being true on the basis of direct observations in the natural world. At its simplest, Popper’s view is that no amount of confirmation proves a theory but one reliable falsification disproves it.

This point was controversial in that the prevailing theory of science preceding Popper was one that relied on empirical observation as the basis of drawing scientific conclusions. If one could make direct observations in the world which supported one’s theory – then the theory was believed to be verified and thus could be acknowledged as true.

Popper used a simple example to illustrate his disagreement. Everyone in Europe and North America knew that all swans were white, but the exploration of Australasia was accompanied by the discovery of black swans, thereby challenging the accepted situation and leading to refinement of the properties that define “swan.” Using this and other examples, Popper successfully argued that using observation to verify a theory does not necessarily result in secure knowledge or empirical truth. He proposed that seeking to falsify theories rather than verify them was a much more reliable way in which to produce knowledge that could be trusted.

Falsification is a simple and powerful tool for testing scientific hypotheses. Sadly it is one that seems to have been beyond the wit of the then US Supreme Court Chief Justice, at the time of the Court’s decision in *Daubert –v- Merrel Dow*. It is itself not absolute, however, and apparent falsifications have to be subject to the same scrutiny as any other element of a hypothesis. Observations may produce information that apparently falsifies a hypothesis, but further testing may provide an explanation that maintains the integrity of the law. Such was the case with the discovery of the planet Neptune. The development of better telescopes led to the observation that the orbit of the planet Uranus was not as predicted by Newton’s laws. Rather than reject Newton’s laws which had stood the test of time, astronomers searched

for an explanation, the most likely of which was the presence of a hitherto undiscovered planet. And so it turned out when Neptune was observed in 1846, in exactly the predicted place.

These examples – the black swans and the discovery of Neptune – show that an essential element of a good theory is its ability to predict something or even better forbid it, and this is as true for the specific application of science in forensic investigations as it is for science in general.

Mention was made above of how Popper's falsification challenged the status quo of the scientific method at the time. This leads to an alternate approach to defining science, by focusing on the behaviors of scientists rather than seeking an explanation of "science" based on the methods used or framework within which they work.

This was the approach taken by Thomas Kuhn (Kuhn 1966) who argued that scientists work within a conceptual paradigm that strongly influences the way in which they see data. Kuhn was interested in the way that science advances, and differentiated between what he called normal science, which progresses incrementally, little by little, and revolutionary science, where no change occurs until there is overwhelming evidence to overcome the prevailing paradigm which is then cast aside and replaced by the new.

In conducting his research Kuhn had to grapple with the essential question of what "science" is in order to be able to discuss how changes come about in scientific theories. He concluded that it is defined by communities of practitioners sharing properties related to the manner in which they conduct the exploration of knowledge about some aspect of the physical world. It is the existence of these communities and their established norms that make the paradigm such a strong force, inevitably resulting in conservatism.

Imre Lakatos is the last of the scientific philosophers to be considered here, but by no means the least, certainly in considering the foundation principles of forensic science. Lakatos was concerned that a strict and unthinking interpretation of Popper's falsification concept meant a hypothesis would be rejected, but its core usefulness might still be valid but lost (Lakatos

1970). Any apparent falsification should not be accepted without question, but may instead point the way to further refinement and improvement of the apparently false hypothesis.

To remedy what he saw as a problem with Popper's theory Lakatos introduced the concept of "research programmes" as a solution. A research programme is the aggregate of several theories which share a core and that progress using a common set of methodologies. In Lakatos' approach, a scientist works within a "research program" that corresponds roughly with Kuhn's "paradigm." Whereas Popper rejected the use of ad hoc hypotheses as unscientific, Lakatos accepted their place in the development of new theories, as of course did Peirce.

The main contribution to forensic science made by Lakatos is his insistence on incorporating the concept of *ceteris paribus* into the evaluation of a hypothesis. Translated as "other things being equal," *ceteris paribus* provides a context in which the formulation of a hypothesis and its acceptance, rejection, or modification depend on the state of our knowledge at the time. Its significance can be summarized as "We don't know what we don't know" and future discoveries – either new and unsuspected information like black swans or apparent contradictions to a well-established theory like Newton's law – can alter the interpretation of observations.

In summary, a single exact definition of science has eluded philosophers of science, but whether it is defined in terms of the scientific method or the behaviors of scientists, one thing is certain: science is not absolute. At any given time the apparent facts revealed by observations and their interpretation are limited by many factors: the limitations imposed by the state of technology (the quality of telescopes, the ability to travel to the opposite side of the Earth); the influence of the paradigms within which scientists work (the behavioral norms of each sect within science); the danger of seeking to confirm rather than challenge; and the impossibility of dealing with things unknown – "we don't know what we don't know" but seldom qualify (or are allowed to qualify) judgment with an admission of *ceteris paribus*.

From Science to Knowledge to Evidence

Recognizing that science is not absolute is not the same as saying science is junk, to use a term currently in vogue among the detractors of forensic science. Bacon was a champion of science as a tool for the betterment of humankind, and our lives today are shaped by it, from the trivia of the entertainment industry, to the very practical and significant advances in medicine, and the seemingly arcane activities associated with increasing our understanding of the natural world, such as the Large Hadron Collider. In the criminal justice system, the value of science is that it provides evidence that is indisputably more reliable than eye-witness testimony and that provides significant information not otherwise available to the triers of fact. For example, the Innocence Project in the United States (Innocence Project 2011) has identified eye witness misidentification as a leading factor in 75 % of the wrongful conviction cases that it has successfully pursued.

The place of forensic science within the justice system is described by Hald, in the entry on “The Philosophical Basis of the Forensic Process,” and the section that follows is to some degree complementary to her paper, as it demonstrates how applying the scientific method to the investigation of crimes produces knowledge that is expressed as evidence for the legal tribunal.

An understanding of the scientific method, and to an extent, how it can produce the knowledge inherent in forensic science, permits an analysis of the purpose of the forensic process, evidence. Although dictionaries define “evidence” in a legal context as something submitted to a court to assist it in arriving at a decision, crime laboratories in the US also use “evidence” synonymously with “exhibit,” namely a physical entity recovered from a crime scene. This chapter avoids that usage, preferring the sequence that an exhibit is something presented before a court, the examination of which has produced evidence that may assist the tribunal in its determinations.

Types of Forensic Evidence

The legal process distinguishes between real and testimonial evidence, which are respectively

evidence resulting from the examination of exhibits and evidence resulting from the direct narrative of a witness as to what he or she has done or observed. Real evidence can be further classified according to the use to which it is put, namely inceptive, exclusionary, corroborative, or associative.

Inceptive evidence is evidence that indicates a crime has been committed. Straightforward examples include detection of accelerant traces at a fire scene and identification of cocaine in someone’s possession. Less straightforward is the finding of semen in samples taken from the vagina of a woman who states that she has been raped, as the evidence speaks to intercourse and not to absence of consent – perhaps a good instance around which to discuss the application of *ceteris paribus* to forensic science. Identifying semen in intimate samples from the alleged victim, and DNA profiling to link the semen to the alleged assailant is the first and core part of the forensic science investigation of a rape. But is it possible to have rape without semen – the use of a condom, or penetration without ejaculation for example, and possible to have semen identified as being from the suspect but without rape – if the court accepts that the intercourse was consensual or that the semen was from a different act of intercourse that occurred some time before the act of rape.

Exclusionary evidence is to identify in forensic science as falsification is to hypothesis testing in science in general. There are many techniques available that speak to individual identity: fingerprints; DNA profiles; facial recognition; acquired characteristics such as tattoos, restorative dentistry, and skeletal fractures; and voice. These may be ranked in ascending order of how well they approach the critical test of knowing the identity of someone: voice, facial recognition, acquired characteristics, DNA, and fingerprints. There is a continuum from transient and variable voice, through subjective facial recognition (although computer programs are changing this), to acquired characteristics which may not be individual and unique, to DNA and fingerprints. Even the DNA and fingerprinting may fall short of what could be accepted as

indisputable knowledge. The DNA profile maybe partial, and the latent print may be limited in its data content. Even if they are both complete there is considerable debate at present as to the uniqueness of any individual fingerprint, with little information on population frequencies, and in regard to even the extremely rare DNA profiles, some statisticians will argue that if there is a finite probability of something happening then sooner or later, it will (Gettinby 2012).

And of course “we don’t know what we don’t know” – how were the fingerprint or biological fluid rendering the DNA profile deposited? It therefore may be more scientifically sound to follow Popper and to say that if there are differences in DNA profiles or fingerprints, then can safely be concluded that this excludes a given individual as being the source, and if none are observed, then the conclusion is that the target person or object cannot be excluded as being the source.

Some of the most compelling examples of exclusionary evidence have come from the Innocence Project (2011) in the United States. As of March 2011, there have been 267 post-conviction exonerations in the US arising from the application of DNA testing, the first being in 1989. The true suspects or perpetrators were identified in 117 of the cases.

Corroborative evidence is what remains when there is no exclusion. Corroborative evidence is very similar to *associative evidence*, which seeks to establish the relationship between things – people with other people, people with places, and objects with other objects. Hairs, fibers, paint glass, soil, some examples of impression evidence, are all examples of areas of forensic science that have traditionally contributed to corroborative evidence. They are also areas that are being encountered less and less frequently because of the difficulty of attaching a weight to them, due to the absence of databases to provide reliable frequencies of the different kinds of corroborative evidence in the population.

The difference between exclusionary evidence and corroborative evidence is mirrored in another way of looking at some forensic evidence, namely identification and class characteristics.

Everyday language uses identity and its derivative “identification” to mean “uniqueness” and the act of “assigning a unique identity” to something or someone. This usage does not correspond however to that found in forensic science, which distinguishes between “identification” and “individualization” thus: identification means that items share a common source or possess the same properties, and individualization means coming from a unique source.

The Latin roots of the two clarify the difference: the root of identify (identification) is *idem*, which means “the same” whereas the root of individualize is *individus* meaning “not divisible.” However, the distinction is not absolute and just when “same” becomes “unique” can be unclear. For example the mechanisms used to fire a bullet and to extract the spent cartridge case from a gun impart markings on the case; imperfections left on the lands of the gun barrel during manufacture and the relationship of the lands to the grooves can be used to associate spent bullet cases and recovered projectiles with the weapon that fired the round. Some of these such as the number of lands and grooves and their twist are characteristic of all examples of a particular model of gun. Others such as the imperfections created by tools when machining the barrel will be specific to an individual gun of that model. However, wear and tear and further imperfections created by dirt in the barrel mean that the features specific to a particular gun will change in the course of time.

It may be of value to refer back to some of the classic scientific activities of the nineteenth and early twentieth centuries, where some of the advances in science were related to classifications within the natural world, that is, with taxonomy. The taxonomy of forensic science can be reduced to two levels: What? and Which or Who? The question “what is this” together with the question of “who is this” (which one of all of the higher taxonomic level) can be simplified into just one – identity. The answer to the WHAT question will be the correct identification of something as being *one of a specified group* of things and the answer to the WHO (or which) question will be the correct identification of a *unique member* within that group.

This differentiation can be illustrated by returning to Popper and his swans. If there is sufficient agreed information to define what a swan is – the highest forensic taxonomic level – then it is possible to test the hypothesis of whether or not something is a swan.

Popper's example uses one property of "swan-ness," color. Note that Popper does not argue that all white birds are swans, but rather that one of the properties of being a swan is to possess white plumage. If the proposition is extended to be "a large waterbird with a long flexible neck, short legs, webbed feet, a broad bill, and all-white plumage" the finding of a bird with all but one of these would have three possible outcomes: the bird would be classified as other than a swan; the bird could be classified as a new species within the *Cygnus* (swan) genus; or the variant would be dismissed as an anomaly.

The second taxonomic level is to identify which swan – no two will be identical (birds don't have identical twins), and to do this requires information beyond the core properties. The conclusions are contextual and depend on the question being posed. If the question is "WHAT" is this, then the object is a swan if it possesses the agreed core properties. If the question is akin to "WHO" is this, which swan, then something other than the core properties is required to confirm its individual identity. How much more is one of the major issues in forensic science.

This approach makes it apparent that the question of identity is a matter of graduation on a continuum between identification of class to unique individualization. Between these two points there are an infinite number of degrees of identification, and many opportunities for these to vary with time.

As an example, a tire - print is located in the snow at a scene in northern Denmark. A cast is made and the size and general brand of the tire determined. The tread pattern shows uneven wear and an area of damage, probably a cut. Some months later investigators identify a suspect whose car has tires of the same make, and one of which has similar wear and damage patterns. However, because of the amount of driving in the time interval and the problems of capturing fine

detail in marks made in snow, there will be differences, at the same time as areas of correspondence between the tire and the print. It is therefore not always possible to conclude that the suspect car is the vehicle involved in the incident, but neither can it be concluded that it is not. This is an example of corroborative evidence that is positioned at an intermediate point in the continuum from what to who.

The above discussion leads to the conclusion that the characteristics of good forensic evidence are:

- It should be stable and not change with time.
- It should have well defined core attributes and known population statistics.
- The basis of variations in non-core attributes should be known and scientifically justifiable.

The above discusses the examination of exhibits in the context of the goal of identifying, usually as "WHAT" or "WHO." There is another framework that is widely used in practice, namely the Locard Exchange Principle and the use of comparison testing. The Locard exchange principle, summarized "as every contact leaves a trace", is the foundation of traditional associated evidence based on microscopy. The concept is simple and can be illustrated as follows: Mr. A met Miss B in a bar one evening; they knew each other slightly because Mr. A was a friend of a distant relative of hers. The conversation was amicable and Miss B had no problem with accepting Mr. A's offer to escort her home when the bar closed. The next morning Miss B's body was found in a river that ran through the park that was on the way between the bar and her home. She had been strangled and raped. This happened some years before the introduction of DNA testing and although blood grouping on the semen from her vagina showed it to be the same as Mr. A's, it was a group O which is found in 50 % of the population. Mr. A was detained on suspicion and his clothing examined microscopically along with that of Miss B. Fibers found on her coat could have come from the jacket of Mr. A and fibers found on Mr. A's underpants could have come from the skirt worn by Miss B.

This two-way exchange of fibers is exactly what is entailed by the Locard exchange

principle. In considering the weight to be attached to the findings the exchange from the outer clothing is of very little significance given the known and admitted social contact in the bar. However the exchange between the skirt and the underpants is an entirely different matter.

This simple and to some extent simplistic example is taken from a real case in which the author was involved. The complexities of the Locard exchange principle and the reason why it is now not so significant in most laboratory investigations, relate to the dynamics of fiber exchange, loss, and redistribution, to the difficulty of attaching an objective meaning to the relatively few possibly transferred fibers recovered compared to the many orders of magnitude more present from the background population, and to the lack of sound databases regarding the frequencies of the types of fibers encountered.

The question was posed earlier of how much information is required in regard to the move of identity from “WHAT,” or class, to “WHO,” or individualization. The same issue of how much is enough applies to comparison testing. However, the fundamental aspect of individualization is that the core properties that define “WHAT” are known and if they are not present then the person or thing is not a member of the class of interest. Comparison testing does not hinge around these core properties, and professional judgment has to be employed to decide if two items could have come from the same source. This is a major shortcoming in comparative identity.

Beyond What and Who: The Science Behind the Forensic Process

The European Network of Forensic Science Institutes (ENFSI 2002) has defined the Forensic Process by means of a set of standards covering activities from the time of the initial actions at the scene to reporting the results of the investigation and testifying in court. They are:

- (a) Undertaking initial actions at the scene of crime
- (b) Developing a scene of crime investigation strategy

- (c) Undertaking of scene of crime investigation
- (d) Assessment of scene of crime findings and considering further examination
- (e) Interpreting and reporting findings from the scene of crime
- (f) Laboratory examination, testing and presumptive testing
- (g) Interpretation of the result of examinations and tests
- (h) Reporting from examinations and tests including interpretation of results

The standards involve planning, identifying and preserving potential exhibits, conducting tests, and, at (e) and (g), evaluation and interpretation of what has been done. Considering whether or not these are scientific activities, requires consideration of parts of the section discussing what science is. As discussed by Hald, even though they followed the highly deductive and anthropocentric model of advancing knowledge based on a foundation of accepted societal beliefs, the ancient Greeks and Romans developed the skills of rhetoric to the point at which logic, evidence and argument became refined and effective tools for decision-making. The integral process, due to the Greek rhetorician Hermagoras but described in Latin by St Augustine, is building the argument round the questions *quis, quid, quando, ubi, cur, quem ad modum, quibus adminiculis* (who, what, when, where, why, how, in what way and by what means/which aids). These are questions that cannot be answered “Yes” or “No” but demand that explanations based on evidence be given.

Lying at the opposite end of the time frame from the ancients, Thomas Kuhn’s concept that science can be defined as something that scientists do provides a link between social behavior, facts, and rhetoric. Accepting that there is indeed an activity called “forensic science” and that it is conducted by crime scene examiners and chemists and biologists working in places called “forensic laboratories” (and the equivalents of these in the several other subgroups of forensic science described in the introduction), then what happens in forensic science can be measured against the requirements of the scientific method, and the evidence-based

objectivity of Hermagoras' questions, or, simply "The 6 Ws":

- WHAT has happened
- WHERE did it take place
- HOW did it happen
- WHO was involved
- WHEN did it happen – at what time and in what sequence
- WHY did it happen

The principle underlying the maxim is that each question should elicit a factual answer. What is important is that none of these questions can be answered with a simple "yes" or "no." They therefore present the incident as a problem to be solved. Solving the problem by providing factual answers to each dimension – each question – is the purpose of investigative inquiry.

The 6Ws define a process (of inquiry) to be used within a process (the ENFSI forensic process), and can be related back to the discussion on the types of evidence that forensic science can furnish. The examples below are from scene examination, because this is the place where Peirce's surprising event begins, and abduction is a necessary part of the process. However, the same principles apply in all stages of the forensic process.

Inceptive evidence was discussed in relation to information regarding specific crimes (arson and drug possession). The first of the 6Ws – What? – is a broader exposition of inceptive evidence. Was a crime committed or not, and if so, was it planned or accidental? In many cases this may appear to be a trivial matter – a body concealed in a shallow grave, with a gunshot wound to the temple. Or a husband calls, pretty agitated. His wife has fallen from the fourth floor window. Ambulance and police are dispatched to the scene and find her lying dead on the footpath, wearing underwear, socks, pants, sweater and rubber gloves of the type normally used for household cleaning activities. She had several injuries that could have been due the fall from the window.

Applying *ceteris paribus* to these cases suggests that the first demands to be treated as murder and the second as accident. But "we don't know what we don't know," and have to challenge the hypotheses. In the apparent murder, the questions of how did the body get there, who else was

involved, and why all had to be pursued and led to the discovery that it was a case of suicide and concealment of the body. The murder hypothesis was falsified by exclusionary evidence, when it was shown that the blood patterns and bullet trajectory at the scene where the shooting occurred – not where the body was found – could only have arisen from a self inflicted shot. The accident hypothesis in the other case was firstly challenged by finding blood inside of the gloves – a surprising event – that led to the discovery of widespread residues of blood in the shower in the house, providing corroborative evidence to the conclusion of murder and not accident. The woman had been bludgeoned to death by her husband.

The murder/suicide case also demonstrates one aspect of the second of the 6Ws – Where? Knowledge of where the incident took place is essential in any investigation. It establishes the boundaries within which to search for evidence and can lead to other layers of interest, for example to the primary scene in the first case above.

The 6 W question of "How did it happen" is both obvious and hidden in the murder case, obviously by a fatal gun shot but hidden that it was self-inflicted. The case of the fallen wife illustrates a different aspect of "How" since the injuries themselves could be explained by the fall, and it was the additional information from the blood inside in the glove – exclusionary or falsifying information in regard to the accident hypothesis – that led to the correct conclusion.

"Why" is the most intangible of the 6Ws. It is perhaps less valuable in recreating events than in bringing a sensible closure to a case. In the suicide/concealment case, the deceased shared a house with others who like him were illegal migrants. His friends found him dead when they returned from work and feared the consequences of reporting the death, hence the concealment of the body.

If It Is Science, Does It Matter?

The development of a sustainable description of what science is – the scientific method and its modern refinements – together with the rigorous

scrutiny of the steps in the forensic process using the ancient tool of rhetoric transcribed into the 6Ws, establishes forensic science as being a well-grounded scientific discipline. The question was not properly addressed in the NAS report but set aside in favor of a discourse on methodologies and statistical interpretations. Unlike the NAS committee, the US Supreme Court did try to come to terms with “science” in *Daubert –v– Merrel Dow*, as did others. However, both made the mistake of assuming that “scientific” somehow meant “reliable and absolute.”

This paper is about the scientific basis of forensic science. The first part shows how science is not absolute but is ever-evolving and contextual. Lakatos remains the thought leader in both areas. The limitations that *ceteris paribus* places on the formulation and evaluation of hypotheses have been described, but Lakatos was also a strong proponent that taking a broader view of Popper’s ideas of falsification is a vital part of the advancement of science. He created the concept of research programs to describe how exploration of real and apparent falsifications strengthen by providing a framework for the continuous iteration of ideas and refinement of knowledge. And so it should be with forensic science, the formulation of a hypothesis and its challenge, and the understanding that falsification of the initial hypothesis is not an end but a beginning, as new testing and a new examination of the existing information leads to a better hypothesis that perhaps approaches knowledge.

However, two vital parts of the work of the forensic scientist are the reliability of the testing methods and the validity of conclusions drawn from observations and tests. These have nothing to do with science any more than the quality of the prosecution or defense in a trial have. Quality assurance in chemical, biological and physical testing is a well-worked topic (Tilstone 2012), not covered here, but are dealt with in the NAS report somewhat (but not entirely) better than it dealt with science.

A more subtle issue is that of confirmation bias. It has long been a tenet of forensic science that findings based on observation, that is,

findings that are subjective, must be confirmed by an independent observer to add a degree of objectivity. The matter of objectivity has been dealt with as an element in quality assurance, by the definition of “Objective test” and a description of how to ensure objectivity in the guidance document for forensic accreditation (ILAC 2002), but recent work shows that the confirmation itself can be subjective, and systems approaches have been proposed to overcome the problem (NIST 2012). Aside from subjectivity and the breakdown (or lack of implementation) of quality assurance procedures, bias is an inevitable consequence of an inductive approach based on asking a question that can be answered “Yes” or “No” and a rigorous implementation of the 6Ws together with effective quality assurance and a more systems approach to observer-dependent testing should resolve the problem.

Related Entries

- ▶ [Automated and Manual Forensic Examinations](#)
- ▶ [Causes of Wrongful Convictions](#)
- ▶ [Cognitive Forensics: Human Cognition, Contextual Information, and Bias](#)
- ▶ [Crime Science](#)
- ▶ [Criminal Investigative Analysis](#)
- ▶ [Critical Report on Forensic Science](#)
- ▶ [DNA Profiling](#)
- ▶ [DNA Technology and Police Investigations](#)
- ▶ [Evidence-Based Policing](#)
- ▶ [Eyewitness Research](#)
- ▶ [Fingerprint Identification](#)
- ▶ [Forensic Anthropology](#)
- ▶ [Forensic Environmental Evidence](#)
- ▶ [Forensic Facial Analysis](#)
- ▶ [Forensic Linguistics](#)
- ▶ [Forensic Palynology](#)
- ▶ [Forensic Science and Criminal Investigation](#)
- ▶ [Forensic Science and Miscarriages of Justice](#)
- ▶ [Forensic Science and the Paradigm of Quality](#)
- ▶ [Forensic Science Culture](#)
- ▶ [Forensic Science Effectiveness](#)
- ▶ [History of Forensic Science in Policing](#)

- ▶ [Jury Decision Making and Eyewitness Testimony](#)
- ▶ [Legal Rules, Forensic Science and Wrongful Convictions](#)
- ▶ [Philosophical Basis of the Forensic Process](#)
- ▶ [Philosophy of Forensic Identification](#)
- ▶ [Probability and Inference in Forensic Science](#)
- ▶ [Scientific Evidence Before International Criminal Tribunals](#)
- ▶ [Scientific Evidence in Criminal Adjudication](#)
- ▶ [Scientific Evidence in Criminal Prosecutions](#)

Recommended Reading and References

- Daubert –v- Merrel Dow. *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 113 S.Ct. 2786, 125 L. Ed.2d 469 (1993)
- ENFSI (2002) Performance based standards for forensic science practitioners. <http://www.enfsi.eu/page.php?uid=108>. Accessed 22 Nov 2011
- Gettinby G (2012) Personal communication
- ILAC (2002) Guidelines for forensic science laboratories. International Laboratory Accreditation Cooperation, Silverwater
- Innocence Project (2011) http://www.innocenceproject.org/Content/Facts_on_PostConviction_DNA_Exonerations.php#. Accessed 10 Mar 2011
- Kuhn TS (1966) *The structure of scientific revolutions*, 3rd edn. University of Chicago Press, Chicago
- Lakatos I (1970) Falsification and the methodology of scientific research programs. In: Lakatos I, Musgrave A (eds) *Criticism and the growth of knowledge*. Cambridge University Press, Cambridge, pp 91–198
- NIST (2012) Expert working group on human factors in latent print analysis. *Latent print examination and human factors: improving the practice through a systems approach*. U.S. Department of Commerce, National Institute of Standards and Technology, Washington, DC
- Peirce CS (1934) In: Hartshorne C, Weiss P (eds) *Collected papers of Charles Sanders Peirce*, vol 5. Harvard University Press, Cambridge, MA
- Popper KR (1970) Normal science and its dangers. In: Lakatos I, Musgrave A (eds) *Criticism and the growth of knowledge*. Cambridge University Press, Cambridge
- Tilstone WJ (2010) Quality in the forensic science laboratory. In: Mozayani A, Noziglia C (eds) *The forensic laboratory handbook procedures and practice*, 2nd edn. Humana Press/Springer, New York
- Tilstone WJ, Hastrup M, Hald C (2012) *Fisher's techniques of crime scene investigation*, First International Edition. CRC Press, Boca Raton, 2012

Scientific Evidence Before International Criminal Tribunals

Melanie Klinkner

Business School, Bournemouth University,
Bournemouth, Dorset, UK

Overview

Over the past decades, a number of international tribunals and courts have been created to help bring international human rights and humanitarian law violators to justice, thus complementing the role of domestic courts. Substantive international criminal law comprises categories of crimes such as war crimes, crimes against humanity, genocide, the crime of aggression, torture, and terrorism (Zahar and Sluiter 2008). For an accused to be found guilty of any of these crimes, the alleged perpetrator must be found criminally liable either through having materially committed the crime or through their engagement in other forms of relevant criminal behavior.

International criminal prosecutions are often exceedingly complex, as evidenced by the overambitious prosecution approach taken in the case of former Serbian President Slobodan Milošević who was charged with over 7,500 crimes committed during three wars spanning 8 years. Expert witness testimonies can be paramount in providing specialist knowledge and contextual information. In international criminal trials, as in domestic English proceedings (x-ref Roberts), expert witnesses “are normally allowed to testify on issues about which the judges themselves, based on their personal knowledge and experience, cannot be expected to reach an opinion alone” (Schabas 2006: 480). As the extensive experience of the International Criminal Tribunal for the Former Yugoslavia (ICTY) demonstrates, scientific expertise, especially relating to mass grave evidence, has been used successfully to prosecute the categories of crimes falling under its jurisdiction, and it is fair to assume that for future international criminal

proceedings, the use of expert witnesses will continue to be imperative.

There are two aspects in particular that distinguish the presentation of expert evidence before international criminal tribunals from domestic proceedings. Firstly, investigations crucially rely on state cooperation, since international criminal institutions have no enforcement agency of their own (Del Ponte 2006). A delay in investigations can mean that much of the evidence, especially physical evidence, has vanished or deteriorated in condition (McGrath 2002). Witness testimony can be lost through death, intimidation, murder, or flight from an ongoing war zone, and access to war-torn countries can be both dangerous and difficult if international and domestic support is not forthcoming. Secondly, and more relevant to the discussion here, the mixed procedural model containing civil law and common law traditions produces a unique blend with few provisions relating to evidence presentation, creating a novel – and, to some, controversial (Murphy 2008) – environment for presenting and evaluating expert evidence. The following discussion provides an overview of the rules governing expert evidence presentation at both the ICTY and the International Criminal Court, before turning to consider how scientific expertise has assisted the ICTY in its fact-finding.

Expert Evidence Provisions at the ICTY and the ICC

Informed by the development of evidentiary rules, proceedings, and jurisprudence from previous international criminal tribunals, the ICC adopts a flexible approach to the admissibility of evidence. Article 69 of the Rome Statute states that “[t]he Court shall have the authority to request the submission of all evidence that it considers necessary for the determination of the truth” (Rome Statute, Art 69.3). The court adopts the affirmative, as opposed to exclusionary, method of evidence admission, taking into account in particular the probative value of the evidence and whether it is prejudicial to fair trial requirements (Rome Statute, Art 69.4).

Likewise, the ICTY’s procedural law does “not contain a detailed set of technical rules” (*Prosecutor v Delalić* 1998: [15]). Rule 89 of the ICTY’s Rules of Procedure and Evidence (RPE) states that evidence is admissible if it is relevant, of probative value, not to the detriment of a fair trial and not otherwise excluded as being obtained through methods that would be detrimental to its reliability or might damage the integrity of the proceedings (ICTY RPE, Rule 95). While the Trial Chambers may thus be inclined to admit scientific evidence, it does not follow that expert evidence will be afforded much, if any, weight in the court’s deliberations. In fact, at the ICTY it has been decided that initial admissibility rulings may be reversed at later stages in the proceedings, as and when further information relating to the validity or reliability of evidence becomes available (*Prosecutor v Orić* 2004).

Expert Witness Status

To the dismay of some (Derham and Derham 2010), neither the ICC nor the ICTY has specifically defined “expert witness.” The ICTY Trial Chamber in *Popović* described an expert witness as someone who possesses the relevant specific knowledge, experience, or skills to help the Trial Chamber come to a better understanding and conclusion on a technical issue (*Popović* 2007a). The qualifications of an expert, summarized in the expert’s curriculum vitae submitted to the court, authorize the expert – unlike an ordinary witness of fact – to state opinions, inferences, and conclusions on matters within the realm of her expertise.

At the ICC, pursuant to Regulation 44(1), the registry holds a list of experts whose qualifications have been verified and “have undertaken to uphold the interests of justice” (*Prosecutor v Dyilo* 2007a: [24]). Where possible, the parties are expected to rely on this list for expert instructions. To facilitate efficient trial management, joint instructions (including by victim representatives) of expert witnesses are preferred.

While qualifications are regarded as formal prerequisites for a witness to qualify as an expert,

objectivity and independence are not. Rather, “the questions of objectivity, impartiality and independence become relevant to assess the weight to be accorded to that opinion evidence” (*Prosecutor v Popović* 2007a: [26]). Affiliation with a party, which has been commonplace in relation to ICTY investigations, does not in itself constitute grounds for disqualification. Indeed, scientific experts were routinely employed by the ICTY’s Office of the Prosecutor, and without this arrangement, little scientific expertise or original physical evidence would have been available to the ICTY. The joint instruction procedure at the ICC should serve to ameliorate perceptions of witness bias, as the opposing parties will be bound to rely on the same expert. Neither the ICTY nor the ICC have Codes of Practice for expert witnesses, but one would expect any scientific expert to be independent and to act in good faith.

Pretrial Chamber and Expert Reports

Standard practice at the ICTY is to tender and admit expert reports through Rule 94 *bis* on the testimony of expert witnesses, which provides a timetable for disclosure and other preliminaries (*Prosecutor v Blagojević* 2003: [20]). At this stage, the opposing party is required to indicate whether it intends to accept the expert witness statement, desires to cross-examine the expert witness, disputes her qualifications, or challenges the relevance of the witness statement. Expert evidence can be denied admissibility on three grounds. It must be excluded firstly, if it has been “obtained by methods which cast substantial doubt on its reliability” (ICTY RPE, Rule 95); secondly, if it jeopardizes the fairness of the trial (ICTY RPE, Rule 89(D)); and thirdly, evidence may be excluded pursuant to Rule 89(C) because it lacks probative value. Of these three grounds for exclusion, Rule 95 has the greatest relevance for the work of scientific experts, as it directly addresses the expert’s methods of data collection and whether, in light of the way the scientific inquiry was conducted, its results are reliable.

Provided that no objection is made by the other side, a scientific report can be admitted

into evidence without hearing testimony from the expert, so long as the Trial Chamber is satisfied as to the evidence’s relevance and probative value (*Prosecutor v Blagojević* 2003). In *Popović*, the ICTY elaborated on the application of the general requirements of relevance and probative value to expert reports, in terms of

- (1) whether there is transparency in methods and sources used by the expert witness, including the established or assumed facts on which the expert witness relied; (2) whether the report is reliable; and (3) whether the contents of the report falls [sic] within the accepted expertise of the witness. (*Prosecutor v Popović* 2007a: [30])

Qualification as an expert does not automatically guarantee the admissibility of the expert’s report. The burden lies on the party tendering the evidence to convince the tribunal that it satisfies Rule 89(C).

At the ICC too, the Trial Chamber has an important role to play regarding expert witnesses. A chamber not only can instruct an expert *proprio motu* (ICC Regulations of the Court, Reg. 44.4.), but the chamber can also determine the subject of an expert report, number of experts in the case, the way experts are instructed, and how they are to present evidence and within what time limits (Reg. 44.5). In *Prosecutor v Jean-Pierre Bemba Gombo*, the ICC Pretrial Chamber issued detailed instructions for the disclosure of all evidence, to ensure that the defendant receives the evidence to be disclosed by the prosecutor (ICC RPE, Rules 76 and 77) and has adequate time and facilities to prepare a defense. Only evidence that is of true relevance to the case should be disclosed. In addition, the prosecutor must supply “sufficiently detailed legal analysis relating the alleged facts with the constituent elements corresponding to each crime charged” (*Prosecutor v Gombo* 2008: [66]). This general requirement naturally applies to scientific evidence. Pages and paragraphs of expert reports or testimonies as well as photographs, physical evidence, and maps must be analyzed in the same way, contributing to a summary table of evidence. The Trial Chamber ultimately rules on the admissibility and relevance of evidence, including scientific evidence.

Admissibility of Summary Reports and Transcript Testimony

The ICTY Rules of Procedure and Evidence contain provisions to facilitate expeditious presentation of complex scientific evidence. Rule 92 *bis* RPE authorized, for example, the presentation of summary reports by investigators in relation to mass grave excavations and examinations. These summary reports are compilations, derived from multiple sources, containing background evidence of the forensic examinations, contextualizing and reducing the complexity of the findings (e.g., Manning 2000). While summary reports can save precious trial time, they may be challenged as hearsay evidence, which is generally admissible in international criminal proceedings but may be accorded little probative value (*Prosecutor v Milošević* 2002: [2](i)–(ii)). That said, investigator Manning’s report on physical evidence recovered from Srebrenica execution points and mass graves used in *Krstić* was subsequently also found to be “highly relevant to the case and admissible under Rule 89” in *Prosecutor v Blagojević* (2003: [30]).

Similarly, Rule 92 *bis* (D) authorizes the admission of trial transcripts of evidence previously given by a witness, including expert witnesses, provided the evidence does not relate to the acts and conduct of the accused. In *Blagojević* statements and transcript testimony of numerous experts relating to mass grave, investigations were admitted in this way. The Trial Chamber was satisfied that the transcript testimonies submitted under Rule 92 *bis* (D), along with the expert reports received pursuant to Rule 94 *bis*, were relevant, probative, and together provided “a complete picture of the expert evidence” (*Prosecutor v Blagojević* 2003: [35]).

At the ICC, Article 69.2 of the Rome Statute and Rule 68 RPE govern the use of recorded testimony, including transcripts or other documented evidence as long as the measure is not “prejudicial to or inconsistent with the rights of the accused” (Rome Statute, Art 69.2). Where the witness is not present before the Trial Chamber, both the prosecutor and the defense must have had an earlier opportunity to examine the witness. Where the witness has consented for the

materials to be used and is present before the chamber, the prosecutor, the defense, and the chamber must have the opportunity to examine the witness during the proceedings if they so wish.

Expert Witness Proofing

At the ICTY, prior to giving evidence in court, experts may review their testimony with the party calling them. Lawyers and expert witnesses may discuss the relevance of the scientific evidence to eliminate any misunderstandings. Such “proofing” may assure the expert witness as to their role in court, help recollect important elements of the evidence, and assist the lawyer in using the evidence effectively. While “rehearsing, practicing, or coaching” is not permitted, it was held that “reviewing a witness’ evidence prior to testimony is a permissible practice under the law of the Tribunal and, moreover, does not per se prejudice the rights of the Accused” (*Prosecutor v Milutinović* 2006: [22]).

In contrast, the ICC’s Victims and Witnesses Unit (VWU) is considered the most appropriate organ of the court to help witnesses become familiar with the experience of giving evidence. The VWU is tasked, firstly, with reminding the witness to tell the truth; secondly, to provide the witness with a copy of their original statement as a memory prompt; and thirdly, to obtain copies of any other statements made by the witness and to ensure that witnesses do not discuss their evidence with the parties. In fact, the Trial Chamber in *Dyilo* emphasized that “the ICC Statute and Rules do not expressly provide for the possibility of parties preparing witnesses for testimony, and further finds no provision in the texts to justify the practice” (*Prosecutor v Dyilo* 2007b: [36]). This was in marked contrast to the ICTY’s more relaxed approach to witness proofing.

However, in a January 2013 decision on witness preparation, Trial Chamber V ruled differently, suggesting that the silence of the Statute does not automatically imply that witness preparation is forbidden. The judges found that witness preparation is likely to “enhance the efficiency, fairness and expeditiousness of the present trial” (*Prosecutor v Ruto and Sang* 2013: [35]).

Testing Scientific Evidence

Fairness demands that the opposing party in an adversarial proceeding must be given the opportunity to test the evidence presented to the fact finder. This is enshrined in Art 67.1(e) of the ICC's Rome Statute giving the accused the right "to examine, or have examined, the witnesses against him or her and to obtain the attendance and examination of witnesses on his or her behalf under the same conditions as witnesses against him or her." Within the general framework of victim representation at the ICC, legal representatives for victims may also question an expert witness subject to Trial Chambers' decisions on what questions may be put to the witness.

In general, testing the reliability and credibility of scientific evidence may involve some or all of the following related issues: (a) the expert's qualifications and status as an expert, (b) the scientific methods adopted, (c) norms of practice, (d) acceptance within the scientific community and validation of methods through publications and peer review, (e) whether and how the science is produced for litigation, and (f) the novelty of the scientific evidence presented (Edmond 2000). Evaluations of expert evidence in international criminal proceedings tend to focus on professional competency, methodologies, and the credibility of the findings in context (Schabas 2006: 480). Rule 140.2(b) of the ICC's RPE specifies that both the prosecution and the defense have the right to question an expert witness about the reliability of the testimony provided, the credibility of the witness, and "other relevant matters."

The following pattern is commonly observed when expert evidence is given in court. After the solemn declaration, examination in chief begins with a discussion of the expert's education and qualifications, employment record, and relevant experience, before substantial matters relating to the scientific report are queried to demonstrate the credibility of the witness and the reliability of her evidence. This is followed by cross-examination by the opposing party and, if necessary, reexamination by the party who initially called the expert. In light of the technical nature of the evidence, the accused can request the presence of her own expert during such testimony to

assist with cross-examination (*Prosecutor v Karadžić* 2011a). Judges at the ICTY may ask questions at any stage, while ICC judges may ask questions only before or after the witness is examined by a party. The defense has the right to be the last to examine the witness (ICC RPE, Rule 140.2(c) and (d)). Challenges to the credibility of experts have been frequent before the ICTY. Experts, such as Dr Haglund in *Prosecutor v Popović* (2007b), have been attacked on a personal level rather than being called upon to defend the substance of their evidence (Klinkner 2009). However, if the ICC's procedural mechanism for joint instruction of expert witnesses becomes well established, one would expect there to be little strategic incentive in trying to discredit an expert whose appointment has been agreed to by all the parties.

In terms of weighing the evidence, Sluiter and Zahar observe that "[c]ross-examination, giving evidence under oath, and the direct perception of the witness's demeanour are key factors in attaching weight to live testimony" (2008: 393). Fact finders will consider whether an expert witness appears honest, independent, and impartial, and must try to establish whether the scientific examinations were conducted according to appropriate scientific standards. Qualification as a bona fide "scientific expert" does not necessarily guarantee the quality of the work conducted in any particular case.

False Testimony and Misconduct

Like all witnesses, expert witnesses are under the obligation to tell the truth. False testimony under oath may result in an investigation, preparation of an indictment for willfully giving false testimony, and, ultimately, prosecution under Rule 91 of the ICTY's RPE or Article 70 of the Rome Statute. To convict an expert witness of false testimony or an offense against the administration of justice "requires the necessary mens rea and not a mere wrongful statement" (*Prosecutor v Akayesu* (1998: [140])). At both institutions, conviction may entail imprisonment or fines. Misconduct can also be charged under Article 71 of the Rome Statute, potentially resulting in interdiction by the court or a fine.

Needless to add, any of these sanctions would incidentally exact a huge reputational cost for the expert's professional standing.

Weighing Expert Evidence: Scientific Evidence from Mass Graves Before the ICTY

Experience from the ICTY demonstrates that scientific evidence, especially in relation to mass grave investigations, has been mostly uncontroversial and generally accepted. First and foremost, scientific evidence recovered from mass graves provides useful corroboration for eyewitness testimony. In Srebrenica, for example, evidence derived from execution points and graves matched accounts of the events by those lucky enough to escape. In the case of Dražen Erdemović, one of the first defendants at the ICTY, who pleaded guilty to murder as a crime against humanity, it was the accused himself who led the investigations to execution and burial sites which were not previously known to the Office of the Prosecutor (*Prosecutor v Erdemović* 1996: [135]). Excavations that took place at the Branjevo Military Farm revealed that there were 132 male victims in the grave, 130 of whom had died from gunshot wounds, and 83 ligatures were found in the grave. Furthermore, analysis showed that the Branjevo Military Farm mass grave had been disturbed. Individuals had been removed and placed in secondary graves (Manning 2000), indicating belated attempts at concealment.

Corroboration of evidence can assist with witness selection, especially as witnesses' memories may have faded or been affected by posttraumatic stress disorder and the passage of time. It is conceivable that some witnesses will suffer from memory loss or alteration by adding to their own painful memories details they heard from fellow sufferers, which poses challenges for the investigation. Furthermore, studies on eyewitness testimony have found that "recall of details from a violent incident was significantly worse than recall of a nonviolent incident" (Clifford and Scott 1978: 356). Contrary to the common sense assumption that most people would never forget the face of an individual who had physically confronted and threatened them, large numbers

of participants in one empirical study were unable to identify the person responsible for their ordeal. This research provides "robust evidence that eyewitness memory of persons encountered during events that are personally relevant, highly stressful, and realistic in nature may be subject to substantial error" (Morgan et al. 2004: 274).

Scientific evidence in general, and forensic science in particular, helps to clarify the context surrounding the crimes and contributes towards proving what crimes were committed and how they were perpetrated.

War Crimes

Grave breaches of the Geneva Conventions (ICTY Statute, Art. 2) fall within the subject matter jurisdiction of the ICTY, while Article 3 provides the tribunal with the power to prosecute "violations of the laws or customs of war," including cruel treatment, torture, and murder. Scientific evidence can assist in confirming these charges. Findings presented in *Prosecutor v Mrkšić* (2007) from forensic examinations of the bodies retrieved from the Ovčara mass grave in Croatia showed that 198 were male and two were female, with an age range from 16 to 72 years. The cause of death in 188 cases was attributable to single or multiple gunshot wounds. Seven individuals were believed to have died from trauma, while the cause of death is still unknown for the remaining five victims. Post-mortem examinations revealed that 86 individuals had suffered from injuries prior to their death on 20/21 November 1991. In 1997, it was possible to identify 192 of the victims buried at Ovčara. With the help of forensic science, the ICTY prosecutor had little difficulty in proving the crimes that had occurred at Vukovar. The tribunal was also satisfied that the victims who were taken from the Vukovar hospital on the morning of 20 November 1991 were at that time not taking part in hostilities and therefore could not be considered legitimate military targets.

Genocide

To constitute the crime of genocide (ICTY Statute, Art. 4), the accused must have deliberately

intended to destroy a protected group in whole or part (Cassese 2008). Where direct evidence of genocidal intent is absent, the requisite intent may be inferred from the factual circumstances of the crime. Perhaps the most interesting case to date where a defendant was indicted for genocide partly on the basis of scientific evidence is *Krstić*. Radislav Krstić stood accused for his actions as Deputy Commander of the Bosnian Serb Army during the Srebrenica massacre between 10 and 19 July 1995. During the trial the prosecution called six forensic experts and two ICTY investigators to give evidence in relation to the mass grave investigations. The Trial Chamber found that the forensic evidence corroborated “important aspects of the testimony of survivors from the execution sites” (*Prosecutor v Krstić* 2001: [71]) and was sufficiently credible and compelling to confirm the *actus reus* of genocide.

The judges concluded that “following the take-over of Srebrenica, thousands of Bosnian Muslims were summarily executed and consigned to mass graves” (*Prosecutor v Krstić* 2001: [73]). The investigations suggested that most of the deceased had not been killed in combat, leading the judges to infer that some 7,000 missing persons had been executed and buried in mass graves. The Trial Chamber reasoned that the disappearance of generations of men showed an intent to physically destroy Bosnian Muslims as an ethnic group. Further indication of the intent to destroy the group, as such, was provided by a “well-established pattern” (*Prosecutor v Krstić* 2001: [68]) of executions. Bodies were not only concealed in mass graves, but were at a later time excavated in an attempt to hide the crimes. Expert examinations of seven secondary graves found commingled and mutilated body parts rendering identification efforts, repatriation, and appropriate burials extremely difficult, causing further distress to the survivors. The fact that all located and examined gravesites associated with the Srebrenica massacre were within the Drina Corps area of responsibility contributed to the Trial Chamber’s overall belief that Krstić shared the intention to commit genocide (*Prosecutor v Krstić* 2001). The Trial Chamber was satisfied that Krstić had participated in the joint criminal

enterprise, sharing the genocidal intent to kill Bosnian Muslims, and duly convicted him of genocide. On appeal, however, this verdict was overturned as the Appeals Chamber felt that the necessary intent to commit genocide was not proven beyond reasonable doubt.

In this case, scientific evidence helped to determine that (a) a specific group was targeted; (b) the killings and burials were systematic; (c) many civilians were amongst the dead; (d) demonstrable attempts had been made to conceal the crimes; and (e) a high level of cooperation was required to undertake such executions and burials.

The *Krstić* judgments and the forensic evidence presented during that trial have since been relevant to other Srebrenica cases, especially *Blagojević*, *Popović* et al., and *Milošević*. In December 2011, three of the forensic experts that appeared in *Krstić* gave evidence before the *Karadžić* proceedings with more experts scheduled to testify in early 2012 (*Prosecutor v Karadžić* 2011b).

Crimes Against Humanity

Numerous defendants have been charged with crimes against humanity (ICTY Statute, Art 5), mostly in relation to attempted “ethnic cleansing” of particular regions. In the *Popović* trial, where five of the defendants stood accused of extermination as a crime against humanity, the defense was keen to clarify whether those found in mass graves had been killed legitimately in combat or whether they were identifiable as civilians whose murder would constitute a crime against humanity. An expert witness, for example, was asked whether some victims from mass graves could have died as a result of combat as opposed to execution and whether military clothing was found on the bodies (*Prosecutor v Popović* 2007c). According to the expert, the evidence suggested that the dead had not been killed in combat as (a) they were not wearing military clothing; (b) the deceased were of all ages, some with physical disabilities; (c) blindfolds and ligatures were found in some graves; (d) many victims had been killed from behind by a single shot to the head; and (e) there was little

indication of previous injuries consistent with combatant status. While it could not be fully excluded that some had been killed in combat, the majority of dead could not be accounted for in that way (also see *Prosecutor v Popović* 2007d). The Trial Chamber accepted these findings and was “satisfied with the reliability of the conclusions in relation to the cause of death reached in the Prosecution expert reports” (*Prosecutor v Popović* 2010: [619]).

In *Prosecutor v Milutinović* (2007, 2009), the accused were allegedly responsible for deportation, forcible transfer, murder (as a crime against humanity and a violation of the laws or customs of war), and persecution of Kosovo Albanians. Volume two of the judgment reviews the evidence relating to the alleged crimes, relying on much of the forensic evidence gathered from investigations conducted in Kosovo during 1999. In light of these findings, the Trial Chamber concluded that over 700 bodies originally buried throughout Kosovo during the NATO bombing campaign were secretly exhumed and transported to Serbia in an attempt to conceal them from citizens of the former Yugoslavia and from the international community (*Prosecutor v Milutinović* 2009). These clandestine operations led the Trial Chamber to believe “that the great majority of the corpses moved were victims of crime, as opposed to combatants or people who perished during legitimate combat activities” (*Prosecutor v Milutinović* 2009: [1357]). Forensic science evidence thus underpinned the Trial Chamber’s conclusion that some of the deceased (who included women and children) were victims of crimes against humanity.

The Future of Scientific Evidence in International Criminal Proceedings

Critics say that the amalgam of civil law and common law approaches to the rules and procedures of evidence creates an “uncertain, obscure, and unworkable body of law that does not expedite proceedings, but offers numerous possibilities for parties to submit motions for the exclusion of evidence” (Zahar and Sluiter 2008: 394).

Others contend that judicial liberality results in the admission of dubious evidence which prolongs proceedings and complicates the tasks of adjudication (Murphy 2008). A flexible approach to the reception of scientific evidence does, however, mean that judges are not required to assess scientific validity when ruling on issues of admissibility.

This uncluttered approach to determining the admissibility of expert evidence at the ICC and the ICTY in no way detracts from the critical importance such evidence may play in international criminal proceedings. In such complex cases, the probative value of scientific evidence is highly contextual. Questions of scientific methodology, of reliability, and of experts’ credibility, objectivity, and impartiality are addressed on a case-by-case basis through the process of testimony in court and tested through cross-examination. It is then the judges’ role to weigh the evidence presented and to arbitrate between reliable and unreliable, as well as scientifically valid and invalid, evidence. The slender body of procedural rules governing expert witness testimony reflects an implicit faith that effective communication between scientific expertise and the law is possible and that traditional legal procedures will facilitate accurate fact-finding in international criminal proceedings.

Related Entries

- ▶ [Expert Evidence and Criminal Trial Procedure](#)
- ▶ [Scientific Evidence in Criminal Prosecutions](#)

Recommended Reading and References

- Clifford BR, Scott J (1978) Individual and situational factors in eyewitness testimony. *J Appl Psychol* 63:352–359
- DelPonte C (2006) Investigation and prosecution of large-scale crimes at the international level – the experience of the ICTY. *J Int Crim Justice* 4:539–558
- Derham R, Derham N (2010) From ad hoc to hybrid – the rules and regulations governing reception of expert evidence at the International Criminal Court. *Int J Evid Proof* 14(1):25–56
- Edmond G (2000) Judicial representations of scientific evidence. *Mod Law Rev* 63:216–251

- International Criminal Court (2004) Regulations of the court (adopted by the judges of the court on 26 May 2004 The Hague) ICC-BD/01-01-04
- International Criminal Court 'Rules of procedure and evidence' UN Doc ICC-ASP/1/3
- International Criminal Tribunal for the former Yugoslavia 'Rules of procedure and evidence' UN Doc IT/32/Rev 42
- Klinkner M (2009) Forensic science expertise for international criminal proceedings: an old problem; a new context; a pragmatic approach. *Int J Evid Proof* 13:102–129
- Manning D (2000) Srebrenica investigations. Summary of forensic evidence execution points and mass graves, available at http://www.domovina.net/archive/2000/20000516_manning.pdf. Accessed 21 February 2013
- McGrath R (2002) Problems of investigations into war crimes and crimes against humanity during and after ethnic conflict. In: Bassiouni MC (ed) *Post-conflict justice*. Transnational, Ardsley
- Morgan CA et al (2004) Accuracy of eyewitness memory for persons encountered during exposure of highly intense stress. *Int J Law Psychiatry* 27:265–279
- Murphy P (2008) Excluding justice or facilitating justice? international criminal law would benefit from rules of evidence. *Int J Evid Proof* 12:1–31
- Prosecutor v Akayesu* (Case No. ICTR-96-4-T) Judgment, 2 September 1998
- Prosecutor v Bemba Gombo* (Case No. ICC-01/05-01/08) Decision on the evidence disclosure system and setting a timetable for disclosure between the parties, 31 July 2008
- Prosecutor v Blagojević and Jokić* (Case No. IT-02-60-T) Decision on prosecution's motion for admission of expert statements, 7 November 2003
- Prosecutor v. Delalić* et al (Case No. IT-96-21-T) Decision on the motion of the prosecution for the admissibility of evidence, 19 January 1998
- Prosecutor v. Erdemović* (Case No. IT-96-22-T) Witness statement by ICTY investigator Jean-Rene Ruez, 19 November 1996
- Prosecutor v Karadžić* (Case No. IT-95-5/18-T) Decision on Accused's request for assistance of defence experts in the courtroom during testimony of seven prosecution expert witnesses, 27 October 2011a
- Prosecutor v Karadžić* (Case No. IT-95-5/18-T) Trial transcripts, 1., 2., and 8. December 2011b
- Prosecutor v. Krstić* (Case No. IT-98-33-T) Judgment, 2 August 2001
- Prosecutor v Lubanga Dyilo* (Case No. ICC-01/04-01/06) Decision on the procedures for instructing expert witnesses, 10 December 2007a
- Prosecutor v. Lubanga Dyilo* (Case No. ICC-01/04-01/06) Decision regarding the practices used to prepare and familiarise witnesses for giving testimony at trial, 30 November 2007b
- Prosecutor v Milošević* (Case No IT-02-54-AR73.2) Decision on admissibility of prosecution investigator's evidence, 30 September 2002
- Prosecutor v Milutinović* et al (Case No. IT-05-87-T) Decision on Ojdanić Motion to prohibit witness proofing, 12 December 2006
- Prosecutor v Milutinović* et al (Case No IT-05-87-T) Expert witness testimony by Dr Eric Baccard, 19–20 February 2007
- Prosecutor v Milutinović* et al (Case No IT-05-87-T) Judgment, 26 February 2009, vol 2
- Prosecutor v Mrkšić* et al (Case No. IT-95-13/1-T) Judgment, 27 September 2007
- Proesctor v Orić* (Case No. IT-03-68-T) Order concerning guidelines on evidence and the conduct of parties during trial proceedings, 21 October 2004
- Prosecutor v Popović* et al (Case No. IT-05-88-T) Decision on defence rule 94 bis notice regarding prosecution expert witness Richard Butler, 19 September 2007a
- Prosecutor v Popović* et al (Case No. IT-05-88-T) Expert witness testimony by Dr William Haglund, 15 March 2007b
- Prosecutor v Popović* et al (Case No. IT-05-88-T) Expert witness testimony by Dr John Clark, Transcript, 20 February 2007c
- Prosecutor v Popović* et al (Case No. IT-05-88-T) Expert witness testimony by Prof Richard Wright, 21 February 2007d
- Prosecutor v Popović* et al (Case No. IT-05-88-T) Judgment, 10 June 2010, vol 1
- Prosecutor v Ruto and Sang* (Case No. ICC-01/09-01/11) Decision on witness preparation, 2 January 2013
- Rome Statute of the International Criminal Court (adopted 17 July 1998, as amended 16 January 2002) UN Doc A/CONF.183/9
- Schabas WA (2006) *The UN International Criminal Tribunals: the former Yugoslavia, Rwanda and Sierra Leone*. Cambridge University Press, Cambridge
- Zahar A, Sluiter G (2008) *International criminal law. A critical introduction*. Oxford University Press, Oxford

Scientific Evidence in Criminal Adjudication

John D. Jackson
School of Law, University of Nottingham,
Nottingham, UK

Overview

The different functions of science and law have long led to tensions between methods of investigation deployed within the sciences

and the various methods of legal fact-finding that are deployed for determining legal disputes. These tensions become particularly acute in relation to scientific evidence in criminal adjudication. On the one hand, the courts have become increasingly reliant on scientific evidence; on the other hand, they cannot for the sake of their own legitimacy be seen to hand over the evaluation of such evidence entirely to scientists. The traditional “adversarial” and “inquisitorial” models of adjudication have each adopted different approaches towards the gathering and evaluation of scientific evidence which have come under critical scrutiny. The use of scientific evidence in “adversarial” systems would seem to have been particularly susceptible to miscarriages of justice while its use in “inquisitorial” systems has led to human rights violations. Against this background, it may be suggested that one way forward is to realign existing procedures in the direction of a participative model of criminal adjudication whereby *both parties and* adjudicators play a role in ensuring that scientific evidence is effectively prepared, presented, and tested.

Scientific Evidence in Criminal Adjudication

The growth in the use of expert or scientific evidence in legal disputes poses challenges to the traditional manner in which legal claims have been resolved. It is noteworthy how many issues of fact are increasingly coming within the domain of expert or scientific evidence (Redmayne 2001). This “creeping scientization of factual inquiry” (Damaška 1997: 143) has arisen as a result of science’s ability to generate knowledge by means of sophisticated technical instruments well beyond the knowledge of laypersons. Equally significant has been the growth in the sciences of the mind which has led to the recognition of new forms of psychological evidence such as post-traumatic stress disorder, battered woman’s syndrome, and recovered memory syndrome. Even the

task of assessing the credibility of witnesses which has long been considered best left to the common sense of the lay judge or jury has been challenged by the growth of experts willing to testify on witness credibility and by such “scientific” techniques as polygraph evidence.

This increasing intrusion of scientific evidence into both civil and criminal adjudication can conjure up a somewhat apocalyptic vision of legal fact-finding processes being replaced altogether by scientists and experts. There are certainly instances where challenges have been made. Challenges to the traditional common law methods of eliciting testimony by means of examination and cross-examination were to be seen in the work of early twentieth century experimental psychologists (see Muensterberg 1908). Although, more recently, legal systems have refined their approaches towards certain kinds of witness testimony, such as the evidence of children and other vulnerable witnesses, behavioral science challenges have not prompted any fundamental review of legal procedures (Greer 1971). Indeed the need for testimony to be tested orally by means of cross-examination would seem to have gained rather than diminished in importance in criminal adjudication, not only in common law processes but in civil law processes as well – partly as a result of the human rights fairness requirement that everyone charged with a criminal offense has the right “to examine or have examined witnesses against him” (see Art. 14(2)(f) of the International Covenant on Civil and Political Rights, Art. 6(3)(d) of the European Convention on Human Rights).

This illustrates that scientific procedures or “scientific method” can never dictate how fact-finding should be conducted in the course of criminal adjudication, even though the same processes of inductive reasoning may be used to reach conclusions in both domains (Twining 1982; Jackson 1988). However dependent legal processes become on scientific evidence, fundamental differences between the scientific enterprise (the pursuit of truth) and the legal enterprise (the resolution of disputes through

justice) mean that different methods are required to produce the desired results in each field, scientific discoveries, or legal verdicts. It is in the nature of the scientific endeavor constantly to seek better evidence, to be prepared to revise even the most entrenched claim in the face of unfavorable evidence (Haack 2004; and for the different values within science and law see Schuck 1993 and Goldberg 1994). This spirit of inquiry is incompatible with the constraints of legal inquiry where decisions have to be reached in the interest of finality (Redmayne 1997; Jackson 2004). As a former English Law Lord, Lord Scarman, once said, Justice cannot wait upon the truth; as the famous legal maxim goes, “justice delayed is justice denied.”

This means that legal procedures cannot simply replicate scientific modes of inquiry. Irrespective of the accuracy of legal procedures, errors will inevitably occur and political decisions which may be highly contestable need to be reached as to how the risk of errors should be distributed. In criminal adjudication there is a consensus across the common law and civil law traditions that proof of guilt must be established beyond reasonable doubt (*in dubio pro reo*, as the principle is known in continental jurisdictions) which reflects a bias in favor of false negatives (acquittals of the guilty) over false positives (convictions of the innocent), although there are differences across the traditions as to how formidable the evidentiary barriers should be for prosecutors seeking to establish guilt (Damaška 1973). The need for finality, which can have profound consequences for those who are found guilty, correspondingly engenders expectations of fair and transparent procedures allowing defendants to contest the charges against them in an open forum and requiring that decisions are fully justified by an independent and impartial tribunal. Because of the “tragic choice” inherent in legal proceedings, whereby decisions need to be sufficiently appeal-proof even when they may be wrong in order to withstand continuing uncertainty (Nobles and Schiff 2000), the process of reaching decisions becomes a vital legitimizing tool in determining their social acceptability.

Although different models of criminal adjudication have developed across legal systems expressing varying societal conceptions of adequate proof, there is a consensus across western countries that they must not only provide for accurate procedures but must also secure public acceptability through fair and open procedures.

The ambivalent attitudes that continue to be displayed towards science and technology suggest that it will be some time before “men and women in white coats” are trusted to replace lay decision-makers in the making of final decisions (Roberts 2008: 324). Scientific evidence needs to be carefully regulated and evaluated so that its probative benefits can be harnessed to assist the lay decision-makers who make final determinations in individual cases.

Adversarial and Inquisitorial Models of Adjudication

Comparative literature on criminal procedure and evidence has been dominated by a tendency to describe different criminal justice systems as either “adversarial” or “inquisitorial,” with the term “accusatorial” often being used synonymously with “adversarial.” Although these labels have enjoyed long currency, they have tended to be used as caricatures which often obscure rather than enlighten how different legal processes might be best understood. It is misleading, for example, to use the terms “accusatorial” and “adversarial” interchangeably to describe Anglo-American systems of criminal justice as the old accusatory forms of justice (characterized by spitting off the prosecutorial function from the final adjudication) have very little in common with contemporary “adversarial” trials, where accused persons are properly equipped through legal representation to contest the charges against them in an effective manner (Vogler 2005). Similarly, the term “inquisitorial” with its overtones of secret procedures dominated by authoritarian judges is hardly an apt description of the modern systems of justice to be seen on the European continent which replicate many of the “adversarial” features of Anglo-American trials.

Rather than use the terms as abbreviated descriptions of actual procedures, they have more helpfully been used as models to idealize opposing features to be found within the generality of Anglo-American and continental procedures. Although there is some difficulty in determining exactly which features should fit within the respective models, the essence of the contrast between adversarial and inquisitorial models of proof seems to boil down to this. Adversarial proceedings are organized around the notion of a dispute or contest between two sides – prosecution and defense – in a position of theoretical equality before a court which must determine the outcome, whereas inquisitorial procedures adopt the notion of an official and thorough inquiry driven by court officials. Thus, in the “contest” model, the prosecution prepares the case, brings the charge, and is responsible for presenting the evidence and proving the offense. The defendant either pleads guilty or attempts to rebut the charge by presenting evidence and arguments against the prosecution. The proceedings are presided over by a neutral adjudicator whose function is to see that the parties play by the rules of the contest but not to take an active part in the presentation of the evidence. In order to ensure that the parties play fair, there is a complex body of rules ensuring that relevant evidence is disclosed, particularly by the prosecution to the defense, in advance of the trial and that evidence is only admissible at the trial if it has sufficient probative value. In the “inquest” model court officials take center stage in the handling of the evidence. The prosecution may first decide the charge but officials of the court then have the responsibility for gathering, testing, and evaluating the evidence. A case file is built up containing all the documents relating to the proceedings and all the evidence that is gathered. This file can be consulted by the parties, but any role the prosecution and defense play in the proof process is incidental and subordinate to the court’s primary function of finding the truth. The trial that follows the proof-gathering phase is largely based on material already in the file and this phase of the proceedings is also dominated by the court rather than the parties.

Although these models are considered to be a valuable heuristic tool for theorizing about different features of Anglo-American and continental processes of proof, more and more “real life” systems have been moving away from the models’ classic features and (on some accounts) converging towards a “middle position” mixing features of both “adversarial” and “inquisitorial” traditions (Markesinis 1994: 30). For example, in the Swiss system, which is commonly perceived to fall within the inquest model, judicial authorities no longer take center stage in the handling of evidence. In Switzerland the investigating judge is now the prosecutor and, in common with a number of other continental systems, the position of the investigating judge has been abolished altogether. In the pre-trial process the prosecutor arranges confrontation hearings and is responsible for conducting them. There is no impartial third party present – just the prosecution and defense – and the defense has the chance to question the witnesses. Another novel feature in many “inquisitorial” systems is that the prosecution and defense may now reach an agreement over the outcome of the case, with the result that the court no longer plays any meaningful role in fact-finding at all (Weigend 2008). One distinguishing feature of common law systems, by contrast, is that for all the possibilities for plea bargaining, the trial is still generally the only forum where evidence can be effectively challenged by the defense. Within civil law systems, on the other hand, the defense has the opportunity and is expected to challenge evidence before trial in the various pre-trial phases, even though in many instances there may not be a judge present to oversee fairness in the process. This suggests that the contrast between criminal procedure systems is better found at the stage of the process at which “adversarial” evidence-testing takes place, rather than in any fundamental or wholesale difference between “adversarial” and “inquisitorial” proceedings.

For all this, the two models appear to retain some explanatory force if they are broadened out to include different political structures of authority and different legal cultures, for

they can then be viewed as two different procedural traditions that have dominated the common law and civil law worlds, respectively, encompassing different normative expressions of how legal proceedings *should* be organized (Langer 2004; Field 2009). Jurists and practitioners within common law and civil law systems tend to identify strongly with idealized procedural models expressing different values. For example, there is a long-standing tradition of anti-inquisitorialism in American law which has acted as a kind of “negative polestar” for American criminal procedure (Sklansky 2009), while “anti-adversarialism” has tended to dominate debates surrounding many criminal justice reforms in France, such as those strengthening defense rights, the movement towards plea bargaining, and the role and function of the *juge d’instruction* (Hodgson 2005).

The Role of Scientific Evidence Within the Adversarial and Inquisitorial Procedural Systems

With the caveat that the adversarial and inquisitorial models may be more useful in reflecting beliefs about how procedural systems *should* be organized than as descriptions of how they actually *are* organized, we can begin to detail how scientific evidence is treated within each of the dominant procedural families (see generally Champod and Vuille 2011). Within the adversarial model, experts are employed by the parties in the pre-trial and trial phase of the criminal process and the court does not become involved in the appointment and use of experts except during the sentencing phase when it does make use of experts whom it appoints to assist in determining risk factors relevant to imposing penalties on convicted defendants. Before this phase, the parties employ experts to gather scientific evidence and write reports. If experts give evidence at trial, they are called as party-instructed witnesses.

The principle of party presentation by which parties call “their” own witnesses at trial extends to expert as well as lay witnesses,

although no party has any exclusive property right in the evidence or any right to prevent another party from calling any witness (*Harmony Shipping Co SA v Saudi Europe Line Limited* [1979] 1 WLR 1380). Scientific evidence is tested in the same manner as other evidence is tested, through cross-examination at trial by counsel. Experts are treated like any other witness in this respect. However, expert witnesses are entitled to somewhat more leeway than ordinary witnesses in being allowed to state their opinion on matters within their expertise, and also to give hearsay evidence of scientific knowledge. Expert reports are often admitted as well by way of exception to the hearsay rule.

Although it is open to both sides to appoint experts, in practice defendants who are legally aided have to make a case to the legal aid authorities for scientific assistance. Even when this is forthcoming, defense experts have to overcome the structural and practical disadvantages of arriving late on the scene after prosecution experts have received items for analysis and conducted their tests and they may have to rely heavily on prosecution experts for information (Roberts 2002). Increasingly, prosecution experts are required to disclose all relevant information to the defense, while the defense is usually only obliged to disclose the scientific evidence it intends to adduce at trial.

In contrast with adversarial process, scientific experts in inquisitorial systems are in theory appointed by, and work under the supervision of, an examining judge or the court. They occupy a unique position in the criminal process, as they are neither witnesses nor judges. Increasingly, however, scientific expertise is called into the case by the police or by the prosecutor in cases where investigating judges are not involved. In reality, experts have often reached their conclusions in conjunction with the police investigation before the case reaches a judge (Nijboer 1993). The experts are usually appointed from lists or from accredited laboratories and have a status superior to that of witnesses. When they have completed their

research and inquiries, experts submit a written report to the court and may be called to give oral evidence to explain their findings under questioning, which is mainly conducted by the court rather than by the parties, although counsel may be given permission to put questions directly to experts. Although the court must freely evaluate the expert evidence for itself, in practice expert opinions are afforded considerable deference.

The defense has traditionally had a limited input into the forensic process in inquisitorial systems, but the defense role has been expanding in more recent years. Defense lawyers sometimes request that particular questions are put to the expert or that certain tests are carried out (see e.g., Art 165 CPP France, Art 184 CPP Switzerland). They may comment on the expert's report once it has become part of the case file and they may request the appointment of another expert for a second opinion, although the courts can deny such petitions (see e.g., Art 167 CPP France, Art 244(4) CPP Germany). In certain systems, such as Italy (see Art 230 CPP), the defense may be able to appoint their own experts to work with the official expert but these party experts usually have inferior status to the official expert and permission may have to be sought before they are instructed (Champod and Vuille 2011: 21–22).

Miscarriages of Justice in Adversarial Systems

Although the use of scientific evidence has the potential to provide a more reliable evidential basis for verdicts than many traditional categories of evidence such as witness testimony and confessions, a litany of miscarriages of justice caused by faulty scientific evidence across a range of different forensic fields has come to light during the last 30 years, especially within the common law world. These have prompted criticism, not only of scientific evidence itself but also of systematic weaknesses in the “adversarial” model of proof. Various cognitive errors can arise when scientific evidence becomes so

dependent on the parties (see generally Roberts 2002, Champod and Vuille 2011). First of all, there is the question whether it is wise to permit the parties to choose which experts to bring into the case. As one judgment over a century ago reflected: “A man may go and does sometimes to half a dozen experts . . . He takes their honest opinions, he finds three in his favour and three against him; he says to the three in his favour will you be kind enough to give evidence? And he pays the three against him their fees and leaves them alone . . . the result is that the court does not get the assistance from the experts which if they were unbiased and fairly chosen, it would have a right to expect” (*Thorn v Worthington Skating Ring Co.* LR 6 ChD 414, 416, 1876). Within the criminal process, the parties may not have the luxury of such a wide choice of experts, but this suggests a further problem. The parties need proper resources to be able to employ good experts which may not be available, in particular, to the defense relying on over-stretched legal aid budgets to secure expert evidence.

The traditional way in which adversarial systems have sought to protect themselves from partisan sources of evidence is by developing strict admissibility tests to screen out unreliable evidence and to give each party an opportunity to challenge the sources of evidence presented by their opponent. These approaches, however, have only had limited success in respect of expert evidence. Adversarial systems have developed a number of rules to restrict the admission of expert evidence. English law has traditionally taken a fairly deferential approach towards experts by allowing them to express expert opinions provided the subject matter of the opinion is beyond the knowledge, skill, or experience of the tribunal of fact and the expert has sufficient expertise in the field (Redmayne 2001). Although the position may now be changing, there is no established requirement that any specific threshold be met to ensure the reliability of expert evidence (Dennis 2010: 895). In practice this has meant that if an expert is accredited or has the necessary qualifications, and the subject matter of the

expert's evidence is relevant to the case before the court, his testimony will be admitted (Allridge 1999).

In 1923 in *Frye v US* (1923) 293 F.1013 a US federal district court appeared to impose stricter gate-keeping responsibilities on judges when it agreed that the trial court had been correct to exclude a polygraph technique because it had not achieved general acceptance in the relevant scientific community. This seemed to put a reliability threshold on the evidence by deferring to scientific consensus in the relevant field. A majority of courts in the US applied this standard and it appeared to creep into other jurisdictions as well (see, e.g., in Australia *R v Bonython* (1984) 38 SASR 45, *R v Parenzee* [2007] SASC 143). The difficulty with this approach was that it led to uneven application in practice. Courts applied varying degrees of rigor to the standard, with little indication as to which facts of the expert testimony had to be generally accepted and how widely or narrowly the relevant scientific community was to be defined (Redmayne 2001: 113; Ligertwood and Edmond 2010: 7.51–2). In *Daubert v Merrell Dow Pharmaceuticals, Inc.*, 509 US 579 (1993), the US Supreme Court considered that the *Frye* test had been superseded by the Federal Rules of Evidence in 1975 which made no mention of the *Frye* test but implied that a judge had to ensure that any and all scientific testimony or evidence admitted was not only relevant but reliable. *Daubert* is commonly considered to have replaced the *Frye* test in the US federal and many state jurisdictions with a reliability test involving consideration of five factors: (1) whether a theory or technique can be (and has been) tested; (2) whether the theory or technique has been subjected to peer review and publication; (3) the known or potential error rate; (4) the existence and maintenance of standards controlling a technique's operation; and (5) general acceptance in the scientific community.

The scope and application of the *Daubert* criteria have been clarified in later Supreme Court decisions. A sixth criterion was added in *General Electric Co v Joiner* 522 US 136

(1997) which entails examining the extent to which general findings are applicable to the instant case. In *Kumho Tire Co. v Carmichael*, 526 US 137 (1999) the Supreme Court broadened the applicability of these criteria to all expert evidence including that which is not strictly, or at all, "scientific." But *Daubert* remains the leading case and is now applied in federal courts in the US and has been cited with approval in certain other jurisdictions as well. In the UK, for example, the England and Wales Law Commission (2009) has recommended that *Daubert*-style criteria be introduced into the English courts.

The difficulty with requiring judges to take on an effective gate-keeping role in relation to expert or scientific evidence is that this depends for its success on judges having a sufficient understanding of the science involved. Although the test would seem to put explicit responsibilities on judges to test the validity of scientific evidence, it has been claimed that within the realm of forensic science many courts have continued to admit bite mark analysis, microscopic hair analysis, voiceprint evidence, and handwriting analysis despite the absence of any demonstrable theoretical basis, population databases, standardized methodology, or empirical data on error rates for this kind of evidence (Beecher-Monas 2007). A comprehensive report into the state of forensic science in the US concluded from reported opinions in criminal cases that, despite *Daubert*, trial judges rarely excluded or restricted expert testimony offered by prosecutors and that appellate courts routinely deny appeals contesting trial court decisions admitting scientific evidence against criminal defendants (National Research Council 2009). Judges are not helped in this endeavor by criminal defense lawyers who rarely have any background in relevant scientific expertise, whilst prosecutors have little responsibility in an adversary system to vouch for the reliability of the expertise they place before the court.

The other means of ensuring reliability of evidence in the adversarial system is to permit

opposing parties to make an effective challenge to the evidence adduced. But this is only possible where the parties are properly resourced and there is sufficient disclosure of all the relevant evidence in the case. In certain notorious miscarriages of justice in England and Wales arising out of IRA terrorism in the 1970s, scientists failed to disclose important scientific information even to the prosecuting authority. The English Court of Appeal responded by requiring scientists advising the prosecuting authority to disclose material known to them which may have some bearing on the case (see *R v Maguire* (1992) 94 Cr App R 133). The difficulty is that, in the absence of effective oversight, investigators and forensic scientists are left to discharge these duties themselves and there may be little incentive for them to do so. Even if scientists are sufficiently objective in their approach, they may have received instructions which are incomplete and may not be alerted to all the relevant evidence in the case. When it comes to communicating their results and opinions, the requirement that they submit to examination and cross-examination by the parties means that questioning is focused on one direction or the other and expert witnesses may not be able to give as complete a picture to the court as they would like. The fact that experts are treated as witnesses also means that they are prevented from entering into open dialogue with decision-makers in the case, potentially frustrating the important educative role that experts have to explain their findings to the court (Allen and Miller 1993).

Human Rights Challenges in Inquisitorial Systems

Inquisitorial systems may be thought better able to withstand some of the cognitive weaknesses of adversarial criminal procedure. Under the inquisitorial trial model, scientific investigations are carried out under the supervision of an impartial court thereby facilitating the production of more impartial and reliable reports.

Experts are also able to communicate the results of their findings to the court, whether in reports or by oral evidence, without being hampered by the artificial adversary process of examination and cross-examination. Problems of disclosure between the parties are pre-empted by having a case file made freely available to both the prosecution and the defense. Experts have full opportunity to respond to any issues raised by the parties or to call for a second opinion. Such a procedural model, however, is heavily dependent on the integrity and competence of “impartial” forensic experts. Scientific experts are perceived to be impartial because they act for the court rather than being instructed by adversarial parties, but in practice many experts work exclusively with the prosecuting authorities and this may put the defense at a considerable disadvantage. In some legal systems, such as that of the Netherlands, government forensic institutes are only permitted to investigate cases when requested to do so by the police or justice authorities (Jakobs and Sprangers 2000). If the defense wishes to commission expert evidence, it must usually pay for it and may encounter difficulties recruiting willing and able experts, particularly if it intends to challenge the findings of the experts appointed by the judicial authorities. Even if the defense is able to commission its own expertise, there is no guarantee that the court will afford the same weight to this evidence as to that produced by the official forensic scientists.

In a string of cases arising from inquisitorial systems, the European Court of Human Rights (ECtHR) has ruled that there was a violation of the right to a fair trial on the grounds that the defense had been disadvantaged under the principle of the “equality of arms” which requires that the prosecution and defense are put on an equal procedural footing. In *Bönisch v Austria* (1987) 9 EHRR 191 the applicant was prosecuted for preparing meat which contained an excessive concentration of benzopyrene on the basis of a report prepared by the Director of the Federal Food Control Institute who was appointed later as an expert

by the court. The ECtHR considered that as it was his report that had led to the charges being brought, the Director of the Institute was more likely a witness against the accused than an independent court expert. The principle of equality of arms inherent in the concept of a fair trial required equal treatment as between the hearing of the Director and the hearing of persons who were or could be called in whatever capacity by the defense. Yet since the Director had been heard as an “expert” his statements must have carried greater weight than those of an “expert witness” called by the accused. As a formally recognized expert, the Director had enjoyed a privileged position in being allowed to attend throughout the hearings, put questions to the accused and witnesses with the leave of the court, and comment on their evidence. It followed that there had been a violation of the right to a fair trial (see also *Eggertsdottir v Iceland*, App. no. 31930/04, 5 July 2007; *Stoimenov v Former Yugoslav Republic of Macedonia*, App. no. 17995/02, 5 April 2007).

Bönisch is often contrasted with *Brandstetter v Austria* (1993) 15 EHRR 378, where the expert whom the court appointed was not the person who filed the report leading to the applicant’s prosecution but was employed by the same Institute as the expert who had filed the report. The ECtHR held that, in this case, doubts about the neutrality of the expert could not be objectively justified and the Austrian courts had not breached the principle of equality of arms in refusing the applicant’s request to appoint another expert. A court apparently does not have to appoint further experts at the request of the defense just because the opinion of the court-appointed expert supports the prosecution case. A second expert must be appointed to satisfy the requirements of the right to a fair trial only where there is some objective ground for suggesting that the expert was biased or, alternatively, some external factor suggesting that the expert’s testimony may have had a distorting impact on the tribunal of fact (*GB v France* ECtHR 2 October 2001).

Apart from falling foul of the requirement of equality of arms, inquisitorial systems have also run up against the European principle of “adversarial procedure.” This does not imply that European systems adopt fully fledged adversarial modes of proof (Jackson 2005). Instead the principle requires that the parties are given the opportunity to have knowledge and comment on the observations filed and the evidence adduced by the other party (*Brandstetter v Austria* (1993) 15 EHRR 378, *Rowe and Davis v UK* (2000) 30 EHRR 1, [60]). In *Mantovanelli v France* (1997) 24 EHRR 370 the ECtHR considered that the principle was violated because the applicants were not informed in advance of the dates on which an expert appointed to report on the circumstances of the death of the applicants’ daughter interviewed witnesses and examined documents. The Court considered that there was no general right to be present during an expert’s investigative activities. What was essential, however, was that the parties should be able to participate properly in the proceedings before the tribunal. Although the applicants could have made submissions to the court on the contents and findings of the report after having received it, the ECtHR was not convinced that the arrangement “afforded them a real opportunity to comment effectively on it.” The question the expert was instructed to determine was identical with the one the court had to determine and pertained to a technical field that was not within the court’s knowledge. Thus, the expert’s report was likely to have a preponderant influence on the assessment of the facts by the court. In these circumstances the applicants could only have expressed their views effectively before the expert report was lodged. Since there would have been no practical difficulty in allowing them to do this, failure to provide advance notification constituted a violation of their ECHR Article 6 right to a fair trial.

Strasbourg case law has established a “two-pronged test” to ensure that criminal proceedings involving expert evidence are fairly conducted (Van Kampen 2000: 201). First of all, national courts must ensure that both parties

are able to have knowledge of, and comment on, the evidence adduced by the opposing party and that they are able to question and challenge any court-appointed expert. According to one of the latest cases, *Mirilashvili v Russia* (2008) App no. 6293/04, 11 December 2008, [190], “if the court decides that an expert examination is needed . . ., the defence should have an opportunity to formulate questions to the experts, to challenge them and to examine them directly at the trial.” This still falls short of any requirement that the defense be allowed to appoint its own expert to mount an effective challenge. If, however, objectively justified fears exist that the expert appointed by the court is not in fact acting with the presumed impartiality and neutrality, a second requirement is that national courts need to ensure that the defendant is provided with the opportunity to secure the attendance and examination of experts and/or witnesses on their behalf under the same conditions as the experts against them. This may require that the court appoints a counter-expert to ensure that the statements of the experts for the prosecution and defense are treated as probative equals.

Future Directions: A Participative Model

A recent report commissioned by the Council of Europe on the state of forensic science both globally and in the European Union has concluded that accusatorial and inquisitorial models alike do not lend themselves to the sound management of scientific evidence (Champod and Vuille 2011: 31). In those systems which give the parties prime responsibility for finding evidence, defendants who are already disadvantaged in the criminal process are penalized still further. Defense lawyers are ill-equipped to scrutinize prosecution expert reports. They may attempt to secure experts themselves but they frequently face difficulty in finding experts working in equivalent conditions to prosecution experts and even when defense experts can be found, they become involved at a comparatively late stage when

critical samples have already been taken and they have to rely on information provided by the prosecution experts. The adversarial trial then requires judges and juries to assess scientific evidence in the “worst conditions imaginable” where the principle of “contradiction for contradiction’s sake” holds sway (Champod and Vuille 2011: 31). Inquisitorial systems, by contrast, place too much trust in court-appointed experts who are complacently supposed to be both neutral and competent and given the misplaced status of a “judge in white” (Volk 1993: 45). The expert’s work is rarely called into question and even when it is, the evidence is not assessed on its scientific merits but on the basis of exogenous factors, such as the expert’s reputation, employment, or official status.

Against the failure of legal systems across the common law and civil law traditions to provide a reliable system for gathering and assessing scientific evidence, it is hardly surprising that much of the focus of recent debates has been on improving standards within forensic science itself (see, e.g., National Research Council 2009; Champod and Vuille 2011). If we extend our gaze beyond the constraints of the adversarial and inquisitorial models of proof, however, improvements might also be made to the legal process within which scientific evidence must be gathered and assessed. The principles of equality of arms and adversarial procedure developed by the ECtHR provide the basis for realigning traditional proof systems, in accordance with a new model of proof better characterized as “participative” than as “adversarial” or “inquisitorial” (Jackson 2005). The participation model is predicated on the notion that although the main actors in the criminal process – prosecutors, defense lawyers, and adjudicators – each have separate defined roles in adjudication, processes of proof should be constructed to provide all these actors with as much relevant evidence as possible, as early as possible in the process so that it can be assessed, challenged, and evaluated in a spirit of positive participation. When it comes to scientific evidence, this would mean that the

parties should ensure that scientific experts are fully informed during the course of their investigations and that the experts themselves are encouraged to share and examine information between themselves. Opportunities for effective challenge and testing should be maximized. This may be better achieved by way of written exchanges before trial than in the full glare of an adversarial trial. Traditional methods of oral examination and cross-examination developed to test cognitive capacities or the good faith of ordinary witnesses are less suited to scrutinizing scientific methodology.

In common with other generalized models of adjudication, it may be that the positive participation model is most useful in expressing values that should be embedded in adjudication than as a detailed “one-size-fits-all” blueprint for conducting all cases. There are also procedural values at work that tend to detract from active participation. The defense has no obligation in Anglo-American systems to put forward any positive theory of its case and is entitled to play a passive role throughout the process. Here the defense may be more inclined to play a negative role of demolishing the prosecution case than a positive role in the generation of specific theories and hypotheses. One commentator has claimed that adversarial systems, in which two parties dance a “tango” with each other to settle for an “interpretive” truth, are ill-suited to the more interactive or collective enterprise of searching for substantive truth which resembles a “rumba” performed by coordinated troupes of dancers moving in time to a shared rhythm (Grande 2008). Where scientific evidence is involved, however, it may be argued that all the lay participants have to try to understand what the evidence means in order to assess its value. Since legal systems cannot rely on scientists themselves to validate scientific theories and techniques, lay participants have no choice but to make the necessary epistemological inquires for themselves and this suggests that, at least in exchanges between the parties and the scientists, a more positive participative exchange should be encouraged and facilitated (see Imwinkelried 2011).

Recommended Reading and References

- Alldrige P (1999) Scientific expertise and comparative criminal procedure. *Int J Evidence Proof* 3:141–164
- Allen RJ, Miller JS (1993) The common law theory of experts: deference or education? *Northwestern Univ Law Rev* 87:131
- Beecher-Monas E (2007) *Evaluating scientific evidence*. Cambridge University Press, Cambridge
- Chamod C, Vuille J (2011) Scientific evidence in Europe – admissibility, evaluation and equality of arms. *Int Commentary on Evidence* 9(1):1–68
- Damaška M (1973) Evidentiary barriers to conviction and two models of criminal procedure: a comparative study. *121 University of Pennsylvania Law Review* 506
- Damaška M (1997) *Evidence law adrift*. Yale University Press, New Haven
- Dennis I (2010) *The law of evidence*, 4th edn. Sweet & Maxwell, London
- Field S (2009) Fair trials and procedural tradition in Europe. *Oxf J Legal Stud* 29:365
- Goldberg S (1994) *Culture clash: law and science in America*. Harvard University Press, Cambridge
- Grande E (2008) Dances of criminal justice: thoughts on systemic differences and the search for truth. In: Jackson J, Langer M, Tillers P (eds) *Crime, procedure and evidence in a comparative and international context: essays in honour of professor Mirjan Damaška*. Hart, Oxford
- Greer D (1971) Anything but the truth? The reliability of testimony in criminal trials. *Brit J Criminol* 11:131
- Haack S (2004) Truth and justice, inquiry and advocacy, science and the law. *Ratio Juris* 17:15–26
- Hodgson J (2005) *French criminal justice*. Hart, Oxford
- Imwinkelried EJ (2011) The end of the era of proxies. In: *Proceedings of the third international conference on evidence law and forensic science*, Beijing, 8
- Jackson JD (1988) Two methods of proof in criminal procedure. *Modern Law Rev* 51:549
- Jackson J (2004) Managing uncertainty and finality: The function of the criminal trial in legal inquiry. In: Duff A, Farmer L, Marshall S, Tadros V (eds) *The trial on trial volume one*. Hart, Oxford
- Jackson JD (2005) The effect of human rights on criminal evidentiary processes: towards convergence, divergence or realignment. *Modern Law Rev* 68:737
- Jakobs LEMP, Sprangers WJMM (2000) A European view on forensic expertise and counter expertise. In: Nijboer JF, Sprangers WJMM (eds) *Harmonisation in forensic expertise*. Thela Thesis, Leiden
- Langer M (2004) From legal transplants to legal translations: the globalisation of plea bargaining and the Americanization thesis in criminal processes. *Harv Int Law J* 45:1

- Law Commission (2011) Expert evidence in criminal proceedings in England and Wales. HMSO: London. Report 325
- Ligertwood A, Edmond G (2010) Australian evidence, 5th edn. LexisNexis, Chatswood
- Markesinis, BS (1994) The gradual convergence. Clarendon Press, Oxford
- Muensterberg H (1908) On the witness stand: essays on psychology and crime. Doubleday, New York
- National Research Council (2009) Strengthening forensic science in the United States: a path forward. The National Academies Press, Washington, DC
- Nijboer J (1993) The law of evidence in Dutch criminal cases in a nutshell: the role of the expert. In: Nijboer JF, Callen CR, Kwak N (eds) Forensic expertise and the law of evidence. Royal Netherlands Academy of Arts and Sciences, Amsterdam
- Nobles R, Schiff D (2000) Understanding miscarriages of justice: law, the media and the inevitability of crisis. Clarendon, Oxford
- Redmayne M (1997) Expert evidence and scientific disagreement. *UC Davis Law Rev* 30:1027–1080
- Redmayne M (2001) Expert evidence and criminal justice. Clarendon, Oxford
- Roberts P (2002) Science, experts and criminal justice. In: McConville M, Wilson G (eds) The handbook of the criminal justice process. Oxford University Press, Oxford
- Roberts P (2008) Faces of justice adrift? Damaška's comparative method and the future of common law evidence. In: Jackson J, Langer M, Tillers P (eds) Crime, procedure and evidence in a comparative and international context: essays in honour of professor Mirjan Damaška. Hart, Oxford
- Schuck PH (1993) Multi-culturalism redux: science, law and politics. *Yale Law Policy Rev* 11:1–46
- Sklansky D (2009) Anti-inquisitorialism. *Harv Law Rev* 122:1634
- Twining W (1982) The rationalist tradition of evidence scholarship. In: Campbell E, Waller L (eds) Well and truly tried: essays on evidence in honour of Sir Richard Eggleston. Law Book, Sydney
- Van Kampen P (2000) Confronting expert evidence under the European convention. In: Nijboer JF, Sprangers WJJM (eds) Harmonisation in forensic expertise. Thela Thesis, Leiden
- Vogler R (2005) A world view of criminal justice. Ashgate, Aldershot
- Volk K (1993) Forensic expertise and the law of evidence in Germany (criminal cases). In: Nijboer JF, Callen CR, Kwak N (eds) Forensic expertise and the law of evidence. Royal Netherlands Academy of Arts and Sciences, Amsterdam
- Weigend, T (2008) The decay of the inquisitorial ideal: plea bargaining invades German criminal procedure. In: Jackson J, Langer M, Tillers P (eds) Crime, procedure and evidence in a comparative and international context: essays in honour of professor Mirjan Damaška. Hart, Oxford

Scientific Evidence in Criminal Prosecutions

Paul Roberts¹ and William O'Brian²

¹School of Law, University of Nottingham, Nottingham, UK

²School of Law, University of Warwick, Coventry, UK

Synonyms

[Expert evidence](#); [Expert witnesses](#); [Forensic evidence](#); [Forensic science evidence](#)

Overview

There is no doubt that forensic science today makes an enormous contribution to the detection and proof of crime in modern legal systems. Scientific evidence is often useful, and sometimes vital, in proving offenses and bringing perpetrators to justice who might otherwise evade detection or conviction. Strong scientific evidence implicating the accused often induces guilty pleas without the need for a contested trial. Scientific evidence simultaneously plays an important role in excluding innocent suspects from further official inquiries or surveillance. Yet scientific evidence can also be a potent source of injustice when errors are made.

This entry explores the role of scientific evidence in criminal prosecutions, particularly with regard to common law jurisdictions such as those in the UK and the USA. It highlights challenges for criminal process actors and indicates issues of on-going controversy and concern. Prosecutors play a pivotal role in the use of scientific techniques of detection and in providing scientific evidence to the court. Common law prosecutors are instrumental in instructing expert witnesses and, in consultation with trial counsel, in adducing scientific reports and testimony in cases proceeding to trial. While the entry

focuses primarily on prosecutors' professional responsibilities within an institutional legal framework informed by particular ideals of justice, the final section briefly addresses a more socio-legal issue: the impact on prosecutions of the "CSI Effect."

Scientific Evidence as Proof of Crime

The use of scientific techniques and expertise in the investigation and prosecution of crime is not a new phenomenon. The appliance of cutting-edge technology in the administration of criminal justice goes hand-in-hand with the Enlightenment and modernity's scientific revolution. Indeed, English law cases addressing issues surrounding expert evidence, such as *Buckley v Rice Thomas* (1554) 1 Plowden 118, CB, (which is still occasionally cited today), date back to the Renaissance. However, the twentieth century is properly regarded as the century of forensic science. This is when the early practical experimentation of the nineteenth century, immortalized by Sir Arthur Conan Doyle in the figure of Sherlock Holmes and his avowedly scientific methods of criminal detection, truly bore fruit. First pathology and fingerprinting, then blood-typing (serology) and a raft of comparison "sciences" (handwriting, ballistics, toolmarks, glass, hairs, fibers, footwear marks, dentition, facial-mapping, voice analysis, etc.), and most recently CCTV cameras, mobile phones, digital forensics, and – above all – DNA profiling have transformed criminal investigation, prosecutions, and trials around the globe.

Forensic science has become an integral part of modern criminal justice systems. Scientific evidence routinely features in high-profile serious crimes of murder, robbery, and sexual assault, but also plays an important role in relation to more mundane offending lower down the criminal calendar. Empirical research indicates that relying on DNA evidence in prosecutions of high volume crimes such as burglary yields many more identifications of perpetrators than more traditional methods of detection,

including fingerprinting, at apparently reasonable cost (Roman et al. 2008: but cf. Williams and Johnson 2008: Chap. 6). It is not surprising that governments have invested heavily in the development of forensic science technologies, notably including the UK's DNA expansion program and the creation of a large National DNA Database (NDNAD) (Williams and Johnson 2008: Chap. 5) – although the Coalition Government's highly controversial recent decision to close the British Forensic Science Service (FSS) is regarded by many as a retrograde step (see House of Commons Science and Technology Committee 2011).

Acknowledging that the development and proliferation of forensic science over the course of the twentieth century has been a marvelous boon for the administration of criminal justice is not to say that it comes without any risks or drawbacks, much less to encourage complacency about it. Like any powerful tool, forensic science gets the job done efficiently and with impressive results but it is equally capable of doing appalling damage if it is not handled carefully and treated with respect. Precisely because scientific evidence has a – generally speaking, well-merited – reputation for objectivity and reliability, it is prone to be highly misleading in cases where it is either inherently unsound or put to improper uses. This is why forensic science is often implicated in miscarriages of justice. In a National Registry of Exonerations recently launched by Michigan and Northwestern law schools in the USA, "bad forensic science" features in 24 % of 891 confirmed false convictions since 1989 (Gross and Shaffer 2012: 63–65). It is often suggested that, whatever defects scientific evidence might suffer from, it has to be more reliable than other forms of evidence traditionally relied on in criminal adjudication, such as eyewitness identifications and confessions by suspects, which have often produced false convictions in the past and have been repeatedly torn to shreds by behavioral science researchers. While it might well be true that scientific evidence is comparatively more reliable than other forms of criminal proof, this is no reason not to strive to improve its performance and to make

every effort to safeguard against errors, to the extent that this might be possible.

Prosecutors play a pivotal role in the use of scientific techniques of detection and in providing scientific evidence to the court. In many continental jurisdictions prosecutors direct criminal investigations, including recourse to scientific testing, and judges seek further technical advice, as required, from their own court-appointed experts. In common law jurisdictions, criminal investigations have traditionally been run by police detectives with substantial operational autonomy and comparatively little input from prosecutors. But this is changing. For example, many state prosecutors in the USA today adopt a proactive approach to participating in major criminal investigations, and even in England and Wales – where strict separation between the police and the Crown Prosecution Service (CPS) used to be an article of faith – there is far greater emphasis on interagency dialogue and cooperation between what is now, revealingly, described in official documents as “the prosecution team” (Moreno and Hughes 2008). Prosecutors are instrumental in instructing expert witnesses and, in consultation with trial counsel, in adducing scientific reports and testimony in those cases that proceed to trial.

Scientific Evidence in Adversarial Criminal Process

Reliance on scientific evidence in particular criminal prosecutions is the outcome of a deliberate, self-conscious process, comprised of a series of reasonably distinguishable, though often overlapping, temporal phases. Forensic science evidence is just as much a product of police case-building, structured by a hypothesized “theory of the case,” as any other type of evidence. Research conducted for the Royal Commission on Criminal Justice identified nine key phases in the production of scientific evidence (Roberts and Willmore 1993; Roberts 1994), suggesting the following, somewhat stylized, model of “the typical case” to which real proceedings conform to a greater or lesser extent.

(1) Investigators must first of all decide to utilize scientific expertise. The initial decision is made by police and prosecutors and therefore answers to investigative, rather than strictly scientific, imperatives. This creates an observable tension between investigators’ needs and expectations, and the ability of science to satisfy them. In recent times attempts have been made to educate investigators about the possibilities and limitations of scientific evidence, harnessing technological innovation to financial imperatives driving “cost-effective” policing (Lawless and Williams 2010). Police officers have been encouraged to turn to science in the investigation of “routine” volume crimes such as burglary and theft, where the potential for scientific assistance has often been overlooked in the past, but progress is uneven across the regions. It remains the case that scientific evidence is produced only when investigators think they need it, which is not necessarily when science might in fact be of most – or indeed, of any – assistance to the prosecution.

(2) Having elected to employ scientific assistance, the first task for police or prosecutors is to locate an appropriate expert. In the UK, until very recently, the FSS has been on hand to provide what was generally acknowledged to be a world-class service in a range of forensic specialisms, including DNA profiling. With the closure of the FSS, reliance must now be placed on market provision by a range of commercial suppliers, some of which – like LGC Forensics, the old Lab of the Government Chemist – were formerly state-run but subsequently privatized. Effective regulation, validation of techniques and processes, and accreditation of laboratories is vital to the integrity of this market-based system. In England and Wales, much of this responsibility has been invested in the newly created post of Forensic Science Regulator, with the assistance of the Forensic Science Advisory Council. For other types of expertise, including clinical, medical, and the more esoteric forensic sciences, investigators have to look to the hospitals, universities, research institutes, and private consultancies to locate an appropriate expert. Although these arrangements seem to

work out in the majority of cases, the process of hiring experts is surprisingly informal, and sometimes fails to produce the best evidence. It remains to be seen whether a post-FSS world of free-market provision will continue to supply an adequate range of high-quality forensic science expertise to the administration of criminal justice, with appropriate investment in quality control and research and development (Lawless 2011; Roberts 1996).

(3) Once an appropriate expert has been engaged, the next step is to supply the expert with relevant crime scene material (or other raw data) for analysis. Crime stain samples or material recovered from suspects or complainants must be identified, preserved, and transmitted to the laboratory free from contamination and protected from (further) degradation. Police and prosecutors must ensure that chain of custody is properly documented, since physical evidence is worthless unless the court can be confident about its provenance and integrity. For example, samples should be collected in “temper-evident” packaging (Lynch et al. 2008: Chap. 4). These simple administrative measures make an essential contribution to ensuring that justice is not only done, but also manifestly seen to be done.

(4) Another crucial aspect of the process of generating forensic science evidence concerns the nature of the instructions received by the expert. Scientists are inevitably influenced by the type and extent of background information provided to them by investigators, and even possibly by the form and wording of the police request for assistance. Researchers have drawn attention to the risks of unconscious biases creeping into scientific judgments through suggestive contextual information, leading forensic experts to “see,” and report to prosecutors and courts, what the information they were given has led them to expect (Risinger et al. 2002). One way to respond to the risk of “contamination” by such extraneous influences would be to insist that items must be sent to the laboratory without any accompanying background information and with a request for assistance in scrupulously neutral terms. At one time the FSS considered this best practice. Yet that approach, and the language

employed in its justification, betrays a fundamental misconception about the nature of forensic science. Following from the essentially applied nature of their discipline, most forensic scientists prefer to be told as much background information as possible, in order to be able to tailor their approach to the needs of the investigation. There is, most obviously, little point in a scientist wasting time and energy on matters that are not disputed in the proceedings. Background information is not so much, on this view, an external source of “contamination” or “bias,” but an essential part of the scientist’s data for analysis. Nonetheless, the continued risk of inappropriate, and possibly unconscious, influence suggests that – to the extent that it can reliably be identified – irrelevant and potentially prejudicial material should be filtered out of the information provided to forensic scientists, at least until they have conducted relevant tests and produced their preliminary findings.

(5) The scientist next proceeds to conduct whatever testing or examinations are judged appropriate. (6) The results of the scientific investigation are then written up into a report. Both testing and reporting reflect investigators’ instructions and expectations, within the broader framework of criminal proceedings and the conventional practices of forensic science. The rules of substantive criminal law structure police investigations, which in turn influence the questions scientific experts are asked to consider, the tests they undertake, and the nature and content of the reports they write. For example, scientists asked to produce a DNA profile from blood-stained clothing in order to establish identity would not routinely comment on a self-defense theory of the case, or any other conceivable hypothesis such as provocation, mistake, or accident. But if specifically invited to do so, the expert could consider how blood spatter patterns, DNA mixtures or secondary transfer, etc., might bear on potential defense arguments and report any relevant findings or advice to prosecutors. According to one police training manual, “forensic awareness... involves the prosecution team providing the forensic scientist with all essential information in the case from the outset, keeping

them continuously updated and asking them the right questions to progress the case and get the best results” (Moreno and Hughes 2008: 25). When producing reports detailing their results and conclusions, experienced forensic scientists are conscious of the instrumental role of science in criminal proceedings, and of a report’s intended audience. Investigators expect scientists to help them prove, or disprove, criminal charges against a suspect, and the style and language of expert reports is directed to that end. There are generalized pressures to work quickly and produce definite conclusions. Even if scientists were always paragons of the ideals of impartial and objective inquiry, which most of them espouse and the law demands, their work product would still remain a highly selective, constrained, stylized, and instrumentally orientated form of science. Unfortunately, forensic scientists have occasionally been seduced, or corrupted, by the institutional pressures inherent to an adversarial criminal process, allowing themselves to become partisan instruments of the prosecution rather than impartial purveyors of objective scientific facts and considered expert opinions to the courts (see Erzinçlioglu 1998; Giannelli 1997).

The production of a scientific report often effectively signals the end of particular criminal proceedings, either because the accused is induced to plead guilty in the face of compelling incriminating evidence, or because the prosecution is too weak to proceed in the absence of scientific corroboration. Scientific evidence plays an important – and possibly somewhat overlooked – role in exonerating innocent suspects from continued suspicion, in cases where DNA profiling indicates that the suspect could not, in fact, have been the perpetrator after all. In those minority of cases that do proceed to trial, the production of scientific evidence typically involves three further key stages.

(7) Defense lawyers sometimes appoint their own experts, occasionally to follow up exculpatory leads, but more often just to double-check the work already conducted by a prosecution expert. The appearance of a defense expert may present opportunities for communication, or even cooperation and the

exchange of ideas, between experts on opposite sides of the adversarial divide. Forensic scientists seldom share the adversarial culture of the police and lawyers who hire them, and some scientists treat their notional adversaries as colleagues. On the occasions when scientists meet in the laboratory to review test results, defense experts have been known to persuade scientists working for the prosecution to undertake further tests or to reinterpret their results in the light of a different perspective or new information (Roberts and Willmore 1993: 56–57). Scientific investigations undertaken by the defense, in conjunction with prosecution scientists or independently, do occasionally produce significant exculpatory evidence. In the majority of cases, however, the defense examiner simply confirms the prosecution expert’s data and conclusions.

(8) Expert witnesses may attend a pretrial conference with counsel. This is potentially an important meeting, at which counsel can review the expert’s evidence and ensure that the expert is prepared to stand by the conclusions expressed in the expert’s report. Empirical research conducted in England and Wales in the early 1990s found that pretrial conferences between experts and counsel were frequently short or nonexistent (Roberts and Willmore 1993: 57–61). This was partly due to time pressures, but also reflected a deeply held aversion on the part of some barristers to meeting any witnesses prior to trial, for fear of attracting accusations of “witness coaching.” That fear seems misconceived with regard to expert witnesses, not because experts are immune to pressure and undue influence, but because counsel are unlikely to lead scientific evidence successfully if they do not themselves understand the expert’s evidence. Lack of pretrial communication between experts and counsel was said on occasion to have adversely affected the presentation of scientific evidence at trial. By contrast, witness preparation is a standard feature of criminal litigation in the USA. No US prosecuting attorney would proceed to trial, in any serious matter, without first having reviewed the evidence of key prosecution witnesses, including scientific experts, in person.

Over the ensuing years, the novel idea of conducting pretrial conferences with expert witnesses has become more familiar and culturally acceptable to barristers in England and Wales, too. More significantly still, English courts have adopted a general policy of proactive judicial trial management within the framework of the Criminal Procedure Rules (CrimPR), which were introduced in 2005. The CrimPR make explicit provision for clarifying, and if at all possible resolving, scientific issues prior to trial through conferences of experts and joint reports. The Court of Appeal has repeatedly emphasized the importance of these provisions, and constantly encourages trial judges to enforce them on the parties. It has come to be widely appreciated that adversarial trials involving “battles of experts” tend to obscure genuine scientific disagreements and sometimes give the impression of scientific dispute or uncertainty where none truly exists. The prevailing philosophy among the senior judiciary is consequently to try to deal with scientific aspects of the case outside the courtroom, while still preserving the determination of genuinely contested issues for the jury at trial.

(9) Although only a very small percentage of criminal cases results in a contested trial in any adversarial jurisdiction where guilty pleas and bargains are the norm, these are disproportionately serious and important cases. Scientific evidence features in many of these contested trials, to a greater or lesser extent. Such evidence is often uncontroversial, and may be agreed – possibly as a result of successful pretrial discussions between the parties and their experts. When not disputed, the expert’s report can simply be read out as documentary evidence, and is likely to be accepted at face-value by jurors (albeit that the jury is always at liberty to form its own view, and must be so directed by the trial judge in England and Wales: see e.g. *R v Allen* [2005] EWCA Crim 1344; *R v Hookway and Noakes* [2011] EWCA Crim 1989). Where the defense or, more infrequently, the prosecution wants to challenge scientific evidence, however, the expert will usually be called to court to testify in person by the party instructing the expert. There is

formal provision for court-appointed experts in many common law jurisdictions. However, judges tend not to utilize court experts, in deference to the adversarial precept that the parties run their own cases at trial, with the judge adopting a relatively passive role as neutral “umpire” to ensure “fair play” and litigants’ adherence to the rules of criminal procedure. (The “umpireal” model of judging in adversarial trials is an overstated and increasingly anachronistic simplification, but still serves to encapsulate a deep-seated cultural difference between “adversarial” and “inquisitorial” conceptions of the trial judge’s role.)

Scientific experts testify in the courtroom like any other witness, through a series of answers to questions put by counsel. The expert witness is first taken through examination-in-chief (in the USA, direct examination) by the advocate calling the witness, and then undergoes cross-examination by counsel for the other party or parties in a multi-handed trial. If thought necessary to clear up any matter raised in cross-examination, a third phase of questioning (“reexamination” in the UK, “redirect” in the USA) may be conducted by the side calling the expert. Needless to say, this is a highly artificial way of presenting scientific evidence to the court. Its success depends in large part on the skill and scientific understanding of counsel, which cannot always be relied upon, especially if pretrial preparation has been inadequate. It might be preferable, from the point of view of communicating scientific evidence effectively, if experts could present their evidence in a more direct and less constrained fashion. As things stand, counsel may through incompetence or as a deliberate strategy distort the intended meaning of an expert’s evidence, and the effect may be compounded where experts called by the prosecution and defense disagree, or appear to disagree, with each other. The evident limitations in this context of traditional oral trial procedure, whatever its efficacy or cultural significance in relation to ordinary witnesses of fact, underscore the wisdom of making more extensive use of pretrial mechanisms to fully and fairly exploit the potential of forensic science in criminal proceedings.

The Common Law Prosecutor's Role and Responsibilities

In many continental legal systems prosecutors are formally part of the magistracy. As a matter of jurisprudential theory, they are unequivocally regarded as acting in the pursuit of justice and they are supposed to be objective and nonpartisan. In common law countries, too, prosecutors are required to be “ministers of justice” who select and pursue appropriate charges against those suspected, on the basis of pertinent and reliable evidence, of having committed particularized criminal offenses. “Prosecution counsel has to exercise independent judgment throughout, with the objective not of obtaining a conviction at any cost, but of ensuring that justice is done” (Buxton 2009: 427). English prosecutors, it is said, prosecute but do not persecute. The special responsibility that prosecutors owe to justice may be rather less self-evident and unequivocal in an adversarial procedural system, in which criminal trials can sometimes take on the appearance of a gladiatorial contest. But it is no less fundamental to the legal system's claims to justice and legitimacy. There is a world of difference between a tough-minded lawyer who prosecutes firmly and fairly, and an ethically cavalier showboater who tries to win cases and secure convictions at almost any cost.

Another important difference between civilian and common law jurisprudential theory is that, whereas many continental systems adopt some version of the “principle of legality” requiring compulsory prosecution in every serious case in which there is solid evidence of criminality, common law prosecutions are fundamentally discretionary. Of course, a suspect cannot be prosecuted in the absence of evidence capable of demonstrating his guilt. But incriminating evidence alone is not automatically enough to justify, let alone mandate, a prosecution in common law jurisdictions. Prosecution in the instant case must also be judged to be in the public interest more broadly conceived. This principle is encapsulated in the two-step test for prosecution specified by the Code for Crown Prosecutors in

England and Wales (CPS 2010). The first part of the test poses the question of evidential sufficiency: is there a “reasonable prospect of conviction” of specified charges on the available evidence? This is generally understood to mean that the reviewing prosecutor must judge that conviction is more likely than not, on the balance of probabilities (ibid: para. 4.6). Otherwise, more evidence must be sought by the police; and the prosecution may ultimately have to be abandoned if better evidence is not forthcoming. Only if the evidential sufficiency test is satisfied should the Crown Prosecutor proceed to consider the second limb of the two-part Code test: is prosecution also justified in the public interest? The Code for Crown Prosecutors contains lists of public interest factors supporting or detracting from the case for prosecution, albeit that offense seriousness is always a weighty consideration and will often be dispositive in practice. Crown Prosecutors in England and Wales are under a duty of continuous review to ensure that the Code Test is satisfied in any case proceeding to trial. Since 2004, Crown Prosecutors have also been responsible for framing the initial charges in serious cases – a task formerly allocated to the police, on the traditional theory that common law prosecutions are initiated by private parties. Selection of appropriate charges likewise requires an (earlier) assessment of evidential sufficiency. Charging and evidential review are among the prosecutor's prime responsibilities in all common law jurisdictions, though detailed doctrinal arrangements and the institutional micro-dynamics of police-prosecutor relationships will naturally differ from one legal system to another.

As forensic science and other kinds of technical expertise have become increasingly prominent features of modern criminal investigations, they have also inevitably presented prosecutors with new challenges and opportunities. Some cases are effectively impossible to prosecute without scientific or medical evidence; in others, expert testimony plays a vital supporting role. Prosecutors must grasp scientific fundamentals in order to make informed assessments of evidential sufficiency. For example, if an expert

DNA report states that the accused's DNA profile matches a crime stain partial profile with a random match probability of 1 in 30,000, the prosecutor needs to be able to make sense of this jargon in order to assess the probative value of the DNA evidence, within the evidential context of the case as a whole. Or again: if a medical expert report states that the pattern of a child's injuries is "consistent with nonaccidental trauma," does this constitute strong evidence of child abuse? Or does it simply fail to rule out abuse as one among myriad other possibilities that could be "consistent with" the injuries observed? In short, widespread use of scientific evidence in criminal prosecutions demands at least a basic level of scientific literacy from prosecutors (as well as from other relevant criminal justice professionals, including defense lawyers and judges). Ideally, prosecutors should also be equipped to spot the potential for developing lines of scientific inquiry that the police might have initially overlooked or regarded as unnecessary. This implies that prosecutors cultivate reasonable familiarity with the range of scientific specialisms and technologies potentially available to support criminal investigations, and should know how to identify and instruct suitably qualified experts, in consultation with police investigators and trial counsel where appropriate.

There are various lists of forensic practitioners produced by professional associations such as the UK Forensic Science Society, the British Psychological Society, and the Society of Expert Witnesses. National prosecution services like the CPS and state District Attorneys offices in the USA might further assist individual prosecutors by building up local intelligence on the range and quality of forensic scientific assistance available for consultation. These largely invisible, office-based processes have rarely been studied by empirical researchers, but what little we do know tends to suggest that identifying and instructing scientific experts is a rather *ad hoc* affair, heavily influenced by the motivation and personal experience of individual prosecutors. As well as failing to capitalize on opportunities to develop better evidence

in the instant case, failure to share experiences more systematically may allow poorly performing experts to remain in circulation long after their consulting forensic practices should have been closed down. It is difficult to believe that this curiously unscientific approach to drawing on scientific expertise is beyond practical improvement – though it has to be said that developments over the last several decades (including the short life and ignominious demise of the UK Council for the Registration of Forensic Practitioners) provide little grounds for optimism.

The CPS in England and Wales has adopted various topic-specific policies pertaining to scientific evidence. One significant interagency document provides guidance on charging in cases involving DNA evidence (The Prosecution Team 2004), including the important principle that "[a] suspect should not ordinarily be charged solely on the basis of a match between his own profile and a DNA profile found at the scene of the crime" (ibid 11.1). This effectively introduced an informal corroboration requirement for DNA profiles in relation to suspects first identified through a speculative search of the NDNAD. It attempts to forestall the embarrassment experienced in several early DNA cases in which suspects who, viewed objectively, could not possibly have committed the crime in question were nonetheless charged purely on the basis of an adventitiously matching DNA profile. This was an object lesson in the dangerous fallacy of treating DNA evidence as though it supplied infallible proof of guilt and an invaluable reminder that, just occasionally, DNA matches can occur purely by chance (if not through contamination, analytical discrepancy, or interpretational error). CPS charging policy in England and Wales now sensibly proceeds on the assumption that, if the suspect is truly guilty, there should be more evidence linking him to the offense than an unexpected "hit" on the DNA database. Yet unjustifiably late decisions to discontinue prosecutions involving DNA matches evidently still do occur (see, most recently, Dodd and Malik 2012).

CPS training and policy specifically draw Crown Prosecutors' attention to the potential for utilizing scientific or medical evidence in a number of priority areas. In relation to domestic violence prosecutions, physical evidence of assaults backed-up by medical testimony may be vital in continuing with a prosecution if the complainant later – for whatever reason – withdraws her complaint (Dempsey 2004), by all accounts a frustratingly common occurrence. Scientific evidence can be important in building up the prosecution's case in rape and sexual assault prosecutions, partly because it can help to deflect the perception that such cases often involve "one person's word against another's," inducing the jury to acquit because it cannot be sure that the accused is guilty, despite jurors' suspicions that he probably is. At the limit, expert testimony can sometimes effectively constitute compelling evidence of guilt or innocence by itself. However, prevailing scientific opinion can and does change over time, making it risky to base a prosecution primarily on expert evidence, at least where the relevant science is rapidly evolving. This dilemma was presented starkly in England and Wales by a recent series of criminal prosecutions of mothers for murdering their children, based in part on the alleged improbability of multiple innocent cot deaths in the same family.

R v Clark [2003] EWCA Crim 1020 was one notorious case in which the prosecution expert testified that the probability that both of the defendant's children could have died from Sudden Infant Death Syndrome (SIDS) was 1 in 73 million. This opinion was invalid when given, since it improperly assumed that serial cot deaths are independent events when in fact they could be caused by the same underlying pathology or genetic predisposition (with the further implication that multiple SIDS might run in families). In *R v Cannings* [2004] 2 Cr App R 7, a case in which three of the accused's infant children had died, the Court of Appeal observed:

Experts in many fields will acknowledge the possibility that later research may undermine the accepted wisdom of today.... That does not normally provide a basis for rejecting the expert

evidence. With unexplained infant deaths, however ... in many important respects we are still at the frontiers of knowledge.... In cases like the present, if the outcome of the trial depends exclusively or almost exclusively on a serious disagreement between distinguished and reputable experts, it will often be unwise, and therefore unsafe, to proceed. (ibid. 111)

Standing alone, this statement might have been interpreted as giving a strong hint to prosecutors that they should discontinue prosecutions in cases of suspicious infant death whenever there was disputed medical evidence. In *R v Kai-Whitewind* [2005] 2 Cr App R 3, however, the Court of Appeal said that it would be a "startling proposition" if "whenever there is a conflict between expert witnesses the case for the prosecution must fail unless the conviction is justified by evidence independent of the expert witnesses" (ibid. 480). The Court proceeded to elucidate an essential distinction:

In *Cannings* there was essentially no evidence beyond the inferences based on coincidence which the experts for the Crown were prepared to draw. Other reputable experts in the same specialist field took a different view about the inferences, if any, which could or should be drawn and hence the need for additional cogent evidence. With additional evidence, the jury would have been in a position to evaluate the respective arguments: without it, in cases like *Cannings*, they would not. (ibid.)

Coincidence, in other words, cannot prove murder beyond reasonable doubt, if expert opinion is divided on whether the circumstances justify an inference of foul play. But the mere fact of expert disagreement will not block a prosecution supported by other evidence.

The Court of Appeal reiterated in *R v Henderson* [2010] 2 Cr App R 24 that conflicts in expert evidence are in principle to be resolved by juries just like any other conflict of factual testimony. *Henderson* concerned three consolidated appeals involving allegedly "shaken babies," prosecutions which have provoked much controversy in both the UK (*R v Harris* [2006] 1 Cr App R 5) and the USA (Tuerkheimer 2011). Prosecution evidence in these cases often almost amounts to an uncorroborated medical diagnosis of murder.

Doctors have testified repeatedly that babies exhibiting a particular combination of symptoms (known as “the triad”) were definitely injured intentionally, and even sometimes claiming that the temporal pattern of symptoms pinpointed the culprit. However, recent developments in medical research and scientific thinking have cast serious doubt on such diagnoses. In *Henderson* the Court of Appeal stated unequivocally that “the triad” is not to be treated as conclusively diagnostic of nonaccidental head injury. This is something else that Crown Prosecutors must now take into account in exercising their charging and prosecutorial discretions in cases of alleged child abuse. Current CPS policy states that prosecutions based only on the triad are unlikely to be brought or continued, in the absence of “appropriate supporting evidence (which in certain circumstances can be found in the absence of certain factors)” (CPS 2011).

Sexual assault cases based on physical examinations of the alleged victims have also posed difficulties for prosecutors. In the past, medical experts have sometimes been willing to infer intercourse from physical marks or injuries, e.g., a ruptured hymen or anal tearing. However, a recent report by the Royal College of Paediatrics and Child Health (2008) called into question the extent to which such examinations could warrant positive findings of abuse, prompting the Court of Appeal to quash a conviction of child sexual assault in *R v PF* [2009] EWCA Crim 1086. Medical evidence led by the prosecution at trial was also criticized for being too dogmatic by the appeal court in *R v Martin T* [2008] EWCA Crim 3229, but on this occasion the conviction was upheld on the strength of the complainant’s testimony. In these delicate situations, involving disputed allegations of sexual abuse or assault, prosecutors must balance their duty to present as strong a case as possible today, against the possibility that changes in medical opinion might tomorrow undermine the safety of any conviction secured at trial. Recent experiences of dramatic reversals in prevailing medical opinion will predictably make prosecutors more reluctant to pursue

cases built upon the shifting sands of scientific knowledge (O’Brian 2011).

These observations and illustrations prompt a more general question: to what extent should prosecutors be responsible for assessing the quality and validity of scientific evidence (as opposed to assessing its probative value on the assumption that scientific evidence is valid)? Of course, if a prosecutor knew or strongly suspected that a particular expert witness was incompetent or a charlatan it would constitute a gross breach of professional responsibility to present that person to the court as a witness of truth. Such faulty evidence could not legitimately contribute towards satisfying the evidential sufficiency test for prosecutions in England and Wales. But this is presumably – to put it no higher – a very unusual scenario. Prosecutors are not themselves trained scientists or forensic experts. When any indications to the contrary are absent, why shouldn’t the prosecutor simply take at face value the validity of scientific evidence produced by ostensibly well-qualified experts? After all, scientific experts are called upon to assist in criminal proceedings precisely because they contribute knowledge and skills otherwise unavailable to the justice system. Expert evidence comprises scientific findings, interpretative judgments, and expressions of opinion which prosecutors are presumptively ill-equipped to second-guess.

Two mutually aggravating considerations belie this superficially plausible rationalization of the legal status quo. First, recent years have witnessed a growing realization that many of the traditional forensic sciences routinely employed in criminal prosecutions and trials actually lack formal, rigorous scientific validation. A report published in 2009 by the highly respected National Research Council of the National Academy of Sciences in the USA concluded that “[i]n a number of forensic science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, and the courts have been utterly ineffective in addressing this problem” (National Research Council 2009: 1–14). Even the reliability of fingerprinting has been called into question by

high-profile controversies and inquiries in the USA (see *ibid*: 5–7 to 5–14) and the UK (The Fingerprint Inquiry: Scotland 2011). The point is not to induce blanket skepticism about scientific evidence and the thousands of convictions based on it, which would be a counterproductive and unwarranted overreaction to these revelations. A more measured approach demands continued vigilance and scrutiny of scientific evidence’s methodological and epistemic credentials, a task to which prosecutors should arguably contribute through searching pretrial evaluations of evidential sufficiency.

Less initiative might be expected of prosecutors in this regard if the validity of scientific evidence were sure to be thoroughly investigated at trial. But herein lies the second cause for concern. Common law courts have not traditionally imposed exacting admissibility requirements on scientific evidence. To the contrary, the general attitude has been that relevant scientific evidence should be admitted, leaving any (rare) challenges to validity to be resolved as questions of fact for the jury in the course of the trial. In recent decades, many common law jurisdictions have introduced somewhat more rigorous approaches to scientific validity, requiring trial judges to undertake some measure of “gatekeeping” scrutiny, at least in relation to novel techniques or applications. The best known of these tests is the admissibility standard elucidated by the US Supreme Court in *Daubert v Merrell Dow* 113 S Ct 2786 (1993) and subsequently adopted in many US state procedural codes. The success, or otherwise, of such interventions is still hotly debated. Commentators have noted a tendency (not necessarily restricted to US judges: Beecher-Monas 2007: 94–121) for *Daubert* to be applied quite stringently to disqualify claimants’ scientific evidence in civil tort suits, while a far more relaxed and indulgent standard is applied to admit scientific evidence adduced by the prosecution in criminal proceedings.

In the meantime, there is no equivalent admissibility test for scientific evidence in England and Wales, albeit that the Law Commission (2011) has proposed legislation to introduce

a locally adapted (and rather elaborate) version of the *Daubert* standard. One might argue that prosecutors’ professional responsibility to inquire into the scientific validity of expert evidence they propose to adduce at trial is all the greater in the knowledge that scrutiny in the courtroom is likely to be minimal. Lay jurors can hardly be relied upon to detect unacknowledged weaknesses in prosecution scientific evidence that lawyers and judges have failed, or disdained, to notice. On the other hand, it is probably unrealistic, in terms of institutional culture and personal psychology, to expect prosecutors in adversarial proceedings to adopt an overly critical attitude towards scientific evidence which, on the face of it, supports their case and would easily satisfy the law’s undemanding admissibility requirements.

Prosecution and Defense

In *R v Ward* (1993) 96 Cr App R 1, 51, the English Court of Appeal observed:

[A] forensic scientist conjures up the image of a man in a white coat working in a laboratory, approaching his task with cold neutrality, and dedicated only to the pursuit of scientific truth. It is a somber thought that the reality is sometimes different. Forensic scientists may become partisan Forensic scientists employed by the government may come to see their function as helping the police. They may lose their objectivity.

Consciously partisan scientific evidence is antithetical to the expert witness’s primary duty to the court. The judges in *Ward* were in no doubt that “the clear duty of government forensic scientists [is] to assist in a neutral and impartial way in criminal investigations. They must act in the cause of justice” (*ibid*. 52). This most basic of the expert witness’s duties has been reiterated many times (e.g. *R v Harris*, [2006] 1 Cr App R 5; *R v B (T)*[2006] 2 Cr App R 3), and is now embodied in rule 33.2 of the Criminal Procedure Rules, which states unequivocally that:

An expert must help the court . . . by giving objective, unbiased opinion on matters within his expertise. This duty overrides any obligation to the person from whom he receives instructions or by whom he is paid.

In the United States, Standard 3–3.3(a) of the *ABA Standards for Criminal Justice: the Prosecution Function* states to similar effect, but with the emphasis on the *prosecutor's* duty:

A prosecutor who engages an expert for an opinion should respect the independence of the expert and should not seek to dictate the formation of the expert's opinion on the subject. To the extent necessary, the prosecutor should explain to the expert his or her role in the trial as an impartial expert called to aid the fact finders . . .

Impartiality and objectivity are indubitably laudable ideals for scientific evidence and expert witness testimony. Whether it is entirely realistic to expect the parties to adversarial litigation to interpret their respective roles with a primary emphasis on achieving justice, viewed objectively and impartiality, is another question. The expert witness's abstract "duty to the court" and to justice provides useful, but very incomplete, guidance for ethical professional practice. Prosecutors and defense lawyers also owe duties to the court, but this hardly settles myriad difficult questions of legal ethics and strategic decision-making in adversarial trial proceedings.

It is accepted as axiomatic in all modern systems of criminal justice that the accused must have a fair opportunity to put his side of the case and adequate time and facilities to mount a defense. Justice is unlikely to be done, and certainly will not be seen to be done, if the accused is muzzled or unreasonably hampered in conducting his defense. However, active defense participation in the proceedings takes on additional functional and symbolic significance in adversarial systems. Adversary theory assumes that the trier of fact will best be able to discern the truth of contested events by hearing each side advance its best case at trial and adjudicating between them (Roberts and Zuckerman 2010: 46–65). There is no overriding duty on the court to discover the truth, as there would be in inquisitorial proceedings. It would consequently undermine not only the normative legitimacy, but also the epistemic efficacy of adversarial criminal trial procedure, if the trier of fact were allowed to hear only the prosecution's side of the story. In fact,

the common law jury is always presented with at least two versions of events in any contested criminal trial, even if the defense "story" is reduced to a blanket denial of the allegations advanced primarily through cross-examination of the prosecution's witnesses.

Adversarialism has broad ramifications for defense participation in criminal trials. Specifically in relation to scientific evidence, the most obvious implication is that the defense should have appropriate access to high-quality forensic science assistance, to retest and verify the prosecution's scientific evidence, and to pursue further scientific inquiries of its own (Giannelli 2004; Roberts and Willmore 1993: Chap. 3). This is partly a question of providing adequate resources, which in practice means legal aid funding for work that the uninitiated might regard as wasteful reduplication. Criminal legal aid is virtually always squeezed in times of austerity. In addition, there is the related question of adequate defense access to evidential material, data, test results, and government labs and technicians (to the extent that these public goods survive free-market provision; wherein commercial confidentiality becomes a further potential barrier to access). This brings us to the more general, and perennially controversial, topic of pretrial disclosure by the prosecution to the defense.

Contrary to what simplistic models of adversarial "contests" might lead one to expect, trial by ambush is largely a thing of the past and is deprecated by modern courts and judges. Most common law procedural systems today require extensive pretrial disclosure by the prosecution, and many jurisdictions also demand more circumscribed disclosure by the defense. (The presumption of innocence preempts fully reciprocal pretrial disclosure in criminal litigation.) Moreover, late or inadequate disclosure by the prosecution has been exposed as a potent cause of miscarriages of justice, not least in relation to scientific evidence.

The extent of the prosecution's common law duty to disclose scientific evidence was clarified in *Ward*, in which the Court of Appeal stated:

An incident of a defendant's right to a fair trial is a right to timely disclosure by the prosecution of all material matters which affect the scientific case relied on by the prosecution, that is, whether such matters strengthen or weaken the prosecution case or assist the defence case. This duty exists whether or not a specific request for disclosure of details of scientific evidence is made by the defence. Moreover, this duty is continuous: it applies not only in the pretrial period but also throughout the trial. The materiality of evidence on the scientific side of a case may sometimes be overlooked before a trial. If the significance of the evidence becomes clear during the trial there must be an immediate disclosure to the defence.

In *Ward* itself, negative test results seemingly contradicting the prosecution's case were withheld from the prosecutor, and consequently were neither disclosed to the defense nor adduced at trial, an outcome which particularly infuriated the Court of Appeal because the jury had been kept in the dark about potentially significant information:

The consequence is that in a criminal trial involving grave charges three senior government forensic scientists deliberately withheld material experimental data on the ground that it might damage the prosecution case. Moreover [in their testimony at trial, two of them] misled the court as to the state of their knowledge about the possibility of contamination occurring from the debris of an explosion. No doubt they judged that the records of the firing cell tests would forever remain confidential. They were wrong. But the records were only disclosed about 17 years after Miss Ward's conviction and imprisonment. (ibid. 49)

Pretrial disclosure in England and Wales is now governed by the Criminal Procedure and Investigations Act 1996, as amended, and the Criminal Procedure Rules. Although the principle of full and timely pretrial disclosure by the prosecution is accepted beyond question in English criminal proceedings, disputes can still arise over its precise requirements in individual cases, especially regarding "unused material" not forming part of the prosecution's affirmative case. Defense lawyers and academic commentators continue to express skepticism about levels of compliance in practice with the prosecution's disclosure requirements (see Lord Justice Gross 2011), possibly indicating deep cultural resistance to full disclosure in an adversarial system (Quirk 2006).

If it emerges at trial that the prosecution has failed to discharge the full extent of its disclosure duties in relation to scientific evidence, the proper course will normally be for the judge to consider granting an adjournment to allow the defense time to deal with new information. The defense would not necessarily require extra time in every case: it all depends on the precise nature of the evidence and the way in which the trial is being run by both sides. It is conceivable, however, that the evidence would have to be excluded altogether, in cases where the defense position has been irreparably damaged, for example, because the opportunity to conduct further testing on perishable samples has been lost. In the event that material nondisclosure of scientific evidence comes to light after the trial has been concluded, the Court of Appeal may determine that a conviction is no longer "safe," in the terminology of the Criminal Appeal Act 1968, and must be quashed. Retrial might be a practical option in some, but by no means all, such cases. Viewed in this light, conscientious compliance with the prosecutor's pretrial disclosure duties is calculated to safeguard the reliability and legitimacy of well-founded convictions of the guilty, at the same time as ensuring the defense has a fair opportunity to present its side of the story to the jury in an adversarial criminal trial.

A "CSI Effect"?

The "CSI Effect" has generated some anxious practitioner and academic discussion, predominantly in the USA (see Cole and Dioso-Villa 2009). The basic idea is that jurors weaned on a concentrated TV diet of the *CSI: Crime Scene Investigation* franchise, *Law & Order*, *Bones*, *Silent Witness*, *Waking the Dead* and similar top-rated serials may have come to entertain hopelessly inflated expectations of scientific evidence. The worry is that jurors in thrall to these cultural fantasies might effectively "punish" the prosecution by voting for an acquittal in any case entirely lacking in scientific proof, or in which the more banal reality of expert evidence

fails to live up to the glamorous fictional portrayals driving jurors' expectations. A handful of US prosecutors claims to have experienced the CSI Effect in cases in which the outcome has been, in their eyes, disappointing (e.g. Thomas 2005).

It is a perfectly plausible conjecture that fact-finding by lay jurors in criminal trials is influenced by a range of social and cultural factors, including expectations generated by crime fiction (Tyler 2006). But such cultural influences are doubtlessly diffuse and operate in concert with countless other complementary and confounding perceptions and motivations. The empirical basis for positing a genuine CSI Effect is exceedingly thin. Most academic commentators who have addressed the issue are highly skeptical, starting with the rather protean and ill-defined concept of "the CSI Effect" itself. In most of the shows in question police and forensic scientists are portrayed as civic heroes and forensic science itself invariably wins the day and puts the bad guys behind bars where they belong. One might equally posit that jurors immersed in these predictable narratives will be only too ready to accept the truth of the prosecution's case and, in particular, to defer to scientific proof of guilt whenever it is adduced (Godsey and Alao 2011).

Whatever the empirical truth of the matter, the theoretical possibility of a CSI Effect serves as a useful reminder that scientific evidence, for all its undoubted potency as a weapon in the armory of law enforcement, is neither infallible nor necessarily free from any conceivable drawbacks, downsides, or unanticipated side-effects. Part of the price of our increasing social investment in scientific proof may be eternal vigilance (not least on the part of prosecutors) to ensure that forensic science serves to promote, and not to thwart, the values and ideals of criminal justice.

Related Entries

- ▶ [Behavioral Science Evidence in Criminal Trials](#)
- ▶ [Causes of Wrongful Convictions](#)

- ▶ [Criminal Defense Profession](#)
- ▶ [Critical Report on Forensic Science](#)
- ▶ [Expert Evidence and Criminal Trial Procedure](#)
- ▶ [Expert Witnesses: Role, Ethics, and Accountability](#)
- ▶ [Fingerprint Identification](#)
- ▶ [Forensic Science and Criminal Investigation](#)
- ▶ [Forensic Science and Miscarriages of Justice](#)
- ▶ [Forensic Science Culture](#)
- ▶ [Forensic Science Professions](#)
- ▶ [History of Criminal Investigation](#)
- ▶ [Identification Technologies in Policing and Proof](#)
- ▶ [Organizing and Supplying Forensic Science Services](#)
- ▶ [Prosecution and Wrongful Convictions](#)
- ▶ [Prosecutorial Discretion](#)
- ▶ [Scientific Evidence in Criminal Adjudication](#)
- ▶ [Transnational Exchange of Forensic Evidence](#)

Recommended Reading and References

- Beecher-Monas E (2007) Evaluating scientific evidence: an interdisciplinary framework for intellectual due process. Cambridge University Press, Cambridge
- Buxton R (2009) The private prosecutor as a minister of justice. *Crim Law Rev* 427–432
- Cole SA, Dioso-Villa R (2009) Investigating the "CSI effect" effect: media and litigation crisis in criminal law. *Stanford Law Rev* 61:1335–1373
- Crown Prosecution Service (2010) The code for crown prosecutors. CPS Policy Directorate, London. www.cps.gov.uk/publications/code_for_crown_prosecutors/index.html. Accessed on 12 Feb 2013
- Crown Prosecution Service (2011) Legal guidance: non accidental head injury cases (NAHI, formerly referred to as shaken baby syndrome [SBS]) – prosecution approach. CPS, London. www.cps.gov.uk/legal/l_to_o/non_accidental_head_injury_cases/. Accessed on 12 Feb 2013
- Dempsey MM (2004) The use of expert witness testimony in the prosecution of domestic violence. CPS, London
- Dodd V, Malik S (2012) Forensics blunder "may endanger convictions". *The Guardian*, 9 March:1, 5
- Erzinçlioglu Z (1998) British forensic science in the dock. *Nature* 392:859–860 (30 April)
- Giannelli PC (1997) The abuse of scientific evidence in criminal cases: the need for independent crime laboratories. *Virginia J Social Policy Law* 4:439–478
- Giannelli PC (2004) *Ake v Oklahoma*: the right to expert assistance in a post-*Daubert*, post-DNA world. *Cornell Law Rev* 89:1305–1419

- Godsey MA, Alao M (2011) She blinded me with science: wrongful convictions and the “reverse CSI-effect”. *Texas Wesleyan Law Rev* 17:481–498
- Gross SR, Shaffer M (2012) Exonerations in the United States, 1989–2012. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2092195. Accessed on 12 Feb 2013
- House of Commons Science and Technology Committee (2011) The forensic science service. Seventh Report of Session 2010–12, HC 855, TSO, London
- Law Commission (2011) Expert evidence in criminal proceedings in England and Wales. Law Com No. 325. www.lawcom.gov.uk/publications/expert-evidence.htm. Accessed on 12 Feb 2013
- Lawless CJ (2011) Policing markets: the contested shaping of neo-liberal forensic science. *Brit J Criminol* 51:671–689
- Lawless C, Williams R (2010) Helping with inquiries or helping with profits? the trials and tribulations of a technology of forensic reasoning. *Soc Stud Sci* 40:731–755
- Lord Justice Gross (2011) Review of disclosure in criminal proceedings. Judiciary of England and Wales, London
- Lynch M, Cole SA, McNally R, Jordan K (2008) Truth machine: the contentious history of DNA fingerprinting. Chicago University Press, Chicago
- Moreno Y, Hughes P (2008) Effective prosecution: working in partnership with the CPS. Oxford University Press, Oxford
- National Research Council (2009) Strengthening forensic science in the United States: a path forward. National Academies Press, Washington, DC
- O’Brian WE (2011) Fresh expert evidence in CCRC cases. *King’s Law J* 22:1–26
- Quirk H (2006) The significance of culture in criminal procedure reform: why the revised disclosure scheme cannot work. *Int J Evidence Proof* 10:42–59
- Risinger DM, Saks MJ, Thompson WC (2002) The *Daubert/Kumho* implications of observer effects in forensic science: hidden problems of expectation and suggestion. *Calif Law Rev* 90:1–56
- Roberts P (1994) Science in the criminal process. *Oxf J Legal Stud* 14:469–506
- Roberts P (1996) What price a free market in forensic science services? The organization and regulation of science in the criminal process. *Brit J Criminol* 36:37–60
- Roberts P, Willmore C (1993) The role of forensic science evidence in criminal proceedings. RCCJ Research Study No 11, HMSO, London
- Roberts P, Zuckerman A (2010) Criminal evidence, 2nd edn. Oxford University Press, Oxford
- Roman JK, Reid S, Reid J, Chalfin A, Adams W, Knight C (2008) The DNA field experiment: cost-effectiveness analysis of the use of DNA in the investigation of high-volume crimes. Urban Institute, Washington, DC
- Royal College of Paediatrics and Child Health (2008) The physical signs of child sexual abuse: an evidence-based review and guidance for best practice. RCPCH, London
- The Fingerprint Inquiry: Scotland (2011) Report. Edinburgh. www.thefingerprintinquiryscotland.org.uk/inquiry/21.html. Accessed on 12 Feb 2013
- The Prosecution Team (2004) Guidance on DNA charging. ACPO/CPS, London
- Thomas AP (2005) The CSI effect and its real-life impact on justice: a study by the Maricopa county attorney’s office. *Prosecutor* 39(October):10–16
- Tuerkheimer D (2011) Science-dependent prosecution and the problem of epistemic contingency: a study of shaken baby syndrome. *Alabama Law Rev* 62:513–569
- Tyler TR (2006) Viewing CSI and the threshold of guilt: managing truth and justice in reality and fiction. *Yale Law J* 115:1050–1085
- Williams R, Johnson P (2008) Genetic policing: the use of DNA in criminal investigations. Willan, Cullompton, Devon

Selection Bias

- ▶ [Juvenile Justice System Responses to Minority Youth](#)

Selective Enforcement

- ▶ [Biased Policing](#)

Self-Control

- ▶ [Situational Action Theory](#)

Self-Control Theory

- ▶ [General Theory of Crime](#)

Self-Report Delinquency Survey

- ▶ [History of the Self-Report Delinquency Surveys](#)

Self-Reported Arrests

► [Self-Reported Offending: Reliability and Validity](#)

Self-Reported Offending

► [Self-Reported Offending: Reliability and Validity](#)

Self-Reported Offending: Reliability and Validity

Darrick Jolliffe¹ and David P. Farrington²
¹School of Law, University of Greenwich,
London, UK

²Institute of Criminology, University of
Cambridge, Cambridge, UK

Synonyms

[Concurrent validity](#); [Gender and race differences](#); [Predictive validity](#); [Reliability](#); [Self-reported offending](#); [Self-reported arrests](#)

Overview

Offending is commonly measured by asking people to admit whether they have committed each of a specified number of delinquent acts such as burglary, theft, robbery, assault, and vandalism. A key question is: How reliable and valid are these admissions? Given that people may conceal, exaggerate, or forget their offenses, how accurate are self-reports as a measure of actual delinquent behavior? This entry will focus on self-reported offending (SRO) surveys of large community samples (at least several hundred), to address this question. The extent to which these reports are differentially valid by gender and race will also be explored. There is

no space to review self-report surveys of prisoners or the literature on self-reported drug use.

Self-reports of offending have been used for a number of years, initially to uncover “hidden delinquency,” but more recently as a complementary measure to indexes of official offending (such as arrests or convictions). Self-reports of offending are reliable, since measures administered at different times produce similar results. The validity of self-reports is usually measured by comparing them with official arrests or convictions. Individuals who self-report offenses are more likely to have official records than those who do not report offenses, and individuals with officially recorded crimes usually admit these crimes. Crucially, self-reported offenses predict future arrests and convictions for the same crimes, even among people with no official record. Overall, the research suggests that self-reports provide a valid measure of offending for white males. Validity is sometimes lower for other demographic categories (e.g., females, Asians, and African-Americans), but these findings are far from consistent. More methodological research on self-reported offending is needed, especially on the differential validity of self-reported offending (SRO) by various demographic categories.

History of the Use of SRO

For many years, people in surveys have been asked to report offenses that they have committed that have not necessarily been detected by the police, in order to obtain information about “hidden delinquency.” For example, Burt (1929) in England and the Glueck and Glueck (1968) in the United States obtained this information from their interviewees and from other informants such as parents and teachers. However, the use of structured SRO questionnaires began in the 1940s. One of the most influential early studies was by Wallerstein and Wyle (1947), who distributed a 49-item mail questionnaire to nearly 1,700 adults in New York. The main finding of this research was that offending was very common: 99 % of adults admitted at least one offense,

and even ministers of the church admitted an average of 8 offenses each. Reflecting on this, however, Wallerstein and Wyle (1947) concluded that acts that were technically offenses in this study were often quite trivial.

The major influence of SRO research on criminology really began in the 1950s with the research of Short and Nye (1958). Perhaps the main reason why their SRO survey was influential was because they found no relationship between socioeconomic status (SES) and self-reported delinquency. However, when they compared training school boys with high school boys, they found that the institutionalized delinquents came from much lower SES backgrounds. The implication, therefore, was that official processing was biased against low SES people – a finding that fitted well with the prevailing theorizing at the time. This result is credited with triggering the “self-report revolution,” with criminologists enthusiastically embracing the self-report method, often using the same items as Short and Nye.

Gold (1970) carried out a very important SRO survey in Michigan. This study is noteworthy because of the care taken in addressing methodological issues such as random sampling, efforts to reduce attrition, and the comparison of self-reports with informant reports of offending. The major findings of this study were that three-quarters of offenses were committed with others (most commonly peers) and that getting caught by the police was followed by an increase in offending (in agreement with labeling theory). Hirschi (1969) was the first to use the self-report method to test a theory about the causes of delinquency, and this seminal and highly cited study inspired a number of similar cross-sectional surveys designed to investigate causes using self-reports.

National SRO surveys quickly followed in the United States (e.g., Williams and Gold 1972), and these provided an alternative measure of juvenile crime rates to the official arrest records. These surveys were in turn followed by the ambitious longitudinal US National Youth Survey (NYS), beginning in 1977 and continuing to the present day (Elliott et al. 1985). This survey is one of the

most important sources of self-report data on criminal careers and the causes of offending, but there are numerous other influential American prospective longitudinal studies which have used the self-report methodology, such as the Causes and Correlates projects in Pittsburgh, Denver, and Rochester; the Seattle Social Development Project; and the Oregon Youth Study (for details about all these surveys, see Farrington and Welsh 2007, Chap. 2). SRO surveys have been used in many other countries, with the Pan-European International Self-Report Delinquency Study (Junger-Tas 2010) being particularly noteworthy. The most comprehensive methodological work on these surveys has been conducted in the United States and England.

History of Research on Reliability and Validity

The first attempt to review the reliability and validity of SRO surveys was in a conference report by Hardt and Bodine (1965). They noted (p. 15) that little was known about these topics because sociologists preferred substantive research to methodological examination. The first systematic assessment of SRO surveys on standard psychometric criteria such as questionnaire construction, administration procedure, objective scoring, norms for various populations, internal consistency, retest stability, and concurrent and predictive validity was completed by Farrington (1973). He published the first demonstration that an SRO survey had predictive validity. The next important methodological and substantive assessment of SRO surveys was completed by Hindelang et al. (1981), who concluded that the reliability and validity of SRO surveys was quite good in comparison with other methods and did not vary much with the method of administration (questionnaire or interview, anonymous or not). Unfortunately, this influential and highly cited book caused a great decrease in methodological SRO research, because it was now assumed that the validity of SRO surveys had been established for all time and that methodological research was not needed. From then on,

the focus was on obtaining substantive results, using self-reports as the main outcome measure in criminology.

Reliability of SRO Surveys

Reliability refers to the consistency of a measure in providing similar results under similar conditions, but this concept is generally more relevant when attempting to assess attitudes or other difficult-to-observe social constructs, as opposed to the recall of actual behavior. However, there are two criteria by which the reliability of SRO surveys can be assessed: test-retest stability (the degree to which results are consistent from one questionnaire administration to the next) and internal consistency (the degree to which items on a questionnaire all measure the same underlying construct), with the former arguably being more important than the latter. This is because it would reflect poorly on SRO surveys if the same person reported a different profile of offending during a specified time period when asked on two separate occasions (low test-retest reliability). When evaluating the test-retest reliability of SRO surveys, the time lag between the two administrations is important. If this time period is too short (e.g., a few days), participants could simply be recalling what they had said previously, whereas if the time lag is too long (e.g., several weeks), participants may forget the offenses that they had committed or may report new offenses committed in the interim.

A number of studies have evaluated the test-retest reliability of SRO measures (with various time lags). Although the results varied depending on the number and types of offenses enquired about (and the methods of scoring), they suggest that these measures are at least as reliable as measures of attitudes (Thornberry and Krohn 2003). For example, Huizinga and Elliott (1986) reinterviewed a random selection of 177 participants from the NYS 4 weeks after the initial completion of the SRO survey. They found that crime index offenses (which are more serious and less frequent) had very high reliability, while less serious and more frequent offenses (e.g., public

disorder offenses) had lower but still acceptable reliability. They also did not find any consistent variation in test-retest reliability over gender, race, social class, or involvement in delinquency.

Validity of SRO Surveys

The key question for SRO surveys is validity: To what extent do self-reports produce an accurate estimate of the true number of offenses committed? And how accurately do self-reports measure the prevalence, frequency, and seriousness of offending? Setting aside the less important issues of content and construct validity, the validity of self-reports is usually assessed by comparing them with some external criterion of offending. The comparison can be concurrent (measure and criterion at the same time) or predictive (measure before criterion).

The main problem centers on what to use as an accurate external criterion of offending. Unlike drug use, for example (Harrison and Hughes 1997), there are no physical traces of burglary or shoplifting in hair, blood, or urine. Some researchers have compared SRO in the usual conditions and when people are told that their lying will be detected physiologically, and generally admissions increase in these physiological conditions. There is some evidence that admissions of problematic behavior are also greater in anonymous conditions, with audio computer-assisted interviews yielding the highest response rate (Tourangeau and Smith 1996).

Notwithstanding the fact that self-reports were intended to overcome some of the perceived deficiencies of official records, SRO results are usually validated against arrests or convictions. It is possible to compare self-reports with parent, teacher, or peer reports of offending and also to compare reports of offending with direct behavioral measures of actual offending, but there have been relatively few of these types of studies (see Farrington et al. 1980). Generally, the most important validity checks are in relation to official records.

There are two general ways in which the concurrent validity of SRO has been evaluated. The

first is to compare self-reported offending (e.g., reporting breaking into a house with the intention of stealing) with official records (e.g., an arrest or conviction for burglary). The second is to examine the fraction of those who are known to have an official record who self-report official offending or to compare self-reported arrests or convictions with official records of arrests or convictions.

The level of correspondence between SRO and official records is influenced by the severity and frequency of the delinquent behavior, with the expectation that more severe and more frequent behaviors would be more likely to result in an official record and more likely to be self-reported. For example, Maxfield et al. (2000) found that 59 % of those with one arrest self-reported that they had been arrested, compared with 73 % of those with 2–4 arrests and 85 % of those with five or more arrests (vs. 21 % of those with no arrests).

Concurrent Validity of SRO

Overall, the correlations between self-reports of offending and official records of arrests or convictions are generally substantial and statistically significant. However, correlations are not very good assessments of the validity of SRO surveys because most self-reported offenses will not result in an official record. The correspondence between self-reported offending and official records (both dichotomized) can be illustrated using a 2×2 table (Table 1).

In cell a (SR Yes, OR Yes) and cell d (SR No, OR No), the two measures of offending correspond, and this is treated as accurate responding. Cell b (SR No, OR Yes) is usually treated as concealment of offending by the individual and

as such is commonly referred to as underreporting. There are of course other reasons why someone might not report an offense that was officially recorded, including forgetting and being innocent of that particular offense (e.g., because of plea bargaining). Cell c (SR Yes, OR No) represents one of the key purposes of SRO surveys, to identify offenders and offenses that have not been officially recorded. A lack of correspondence here can therefore not necessarily be attributed to a lack of validity, although this cell count might be influenced by exaggerated reporting. However, when self-reports of arrests or convictions are compared to official records, cell c might indicate overreporting (or faulty records).

When examining the validity of SRO, it might be expected that the probability of an arrest or conviction would increase with the intensity (e.g., frequency or seriousness) of self-reported delinquency. For example, Jolliffe et al. (2003) investigated the validity of self-reports by comparing these to the court referrals of over 800 boys and girls (age 11–17) in the prospective longitudinal Seattle Social Development Project (SSDP). A total of 626 youths self-reported at least one offense during this age range, and 230 of these (37 %) had a court referral. In contrast, 101 reported that they had not committed any offense, and only 16 of these (16 %) had a court referral. This comparison was statistically significant and suggested that court referrals were much more common among those who self-reported an offense than among those who denied offending. This measure of concurrent validity was highest for those who self-reported burglary (a serious offense) or drug offenses (the most frequent). A number of other studies have also demonstrated the concordance between SRO and official records.

Some researchers have questioned, however, whether self-reports of offending are equally valid across demographic categories. Generally, male/female and black/white ratios in American SRO surveys are lower than in official records. One possible reason for this is that male and black youth are more likely to conceal their offenses than females and white youths, but there are also

Self-Reported Offending: Reliability and Validity,
Table 1 Self-reports versus official records

		Self-report	
		Yes	No
Official Record	Yes	a	b
	No	c	d

a number of other potential explanations. Since the prevalence of offending is usually much greater in self-reports than in official records, there may be a “ceiling effect”; for example, if 50 % of white youth self-reported offending, the maximum possible black/white ratio in self-reports would be 2:1, whereas if 20 % of white youth were convicted, the maximum possible black/white ratio in convictions would be 5:1. In addition, self-reported offenses may not be comparable to (may be less serious than) official offenses.

Methods of scoring may influence race and gender ratios. Elliott and Ageton (1980) found that these ratios were much greater when the scoring allowed for large numbers of admitted crimes, rather than the maximum category being “three or more,” for example. They concluded (p. 107) that race and gender differences were “more extreme at the high end of the frequency continuum, that part of the delinquency continuum where police contacts are more likely.” Another possibility is that official records may reflect police and court bias against male and black youth, and more recorded male and black youth might in fact be innocent. Alternatively, male and black youth may be disproportionately arrested because of their aggressive demeanor. In a validity test, Hindelang et al. (1981) investigated how many offenses known to the police were self-reported and found that white males failed to report 20 % of serious offenses, compared with 57 % for black males, 50 % for white females, and 59 % for black females.

Similar findings on differential validity of SRO have been reported in other studies, but the pattern is far from consistent. For example, using data from the Pittsburgh Youth Study, Farrington et al. (1996) discovered that, although the self-reports of both Caucasian and African-American boys were significantly associated with their court referrals, the relationship was stronger for Caucasians. However, in the SSDP, Jolliffe et al. (2003) found that concurrent validity was lowest for Asians and females and that African-American males had the highest concurrent validity.

The comparison between SRO and official records has also been studied for racial and ethnic minorities in other countries. In a longitudinal

study in New Zealand, Fergusson et al. (1993) found that children of Maori or Pacific Island descent were 2.9 times more likely to have a police record than children of European descent, but were only 1.7 times more likely to offend according to parent and self-reports. The researchers attributed these findings to ethnic biases in citizen reporting and police recording. In a national SRO survey in the Netherlands, Junger (1989) discovered that Moroccan and Turkish boys with police records were less likely to admit delinquency than indigenous Dutch boys or Surinamese boys. She pointed out that concurrent validity was not necessarily lower for ethnic minorities, because the Surinamese boys were black.

Another measure of the concurrent validity of SRO is the percentage of those with a court record for a particular offense who admit committing the same offense. For example, of those with a court referral for burglary, what percentage admit that they have committed a burglary? The association between self-reports and official records in this type of comparison is generally quite high. For example, in the SSDP, Jolliffe et al. (2003) found that, of the 246 youth referred to court, 94 % admitted committing at least one offense. The probability for specific offenses was lowest for robbery (38 %) and highest for marijuana use (100 %). Figures ranging from 54 % to 90 % have been discovered in similar studies, suggesting that most officially identified offenders self-report their particular offenses, but some differential validity by gender and race has been found. For example, in the study by Jolliffe et al. (2003), males were more likely to admit than females, and Asians were less likely to admit than either African-Americans or Caucasians. There were no differences in concurrent validity between African-Americans and Caucasians.

Concurrent Validity of Self-Reports of Arrests or Convictions

A more direct method of investigating validity is to compare self-reports of arrests or convictions

with official records of arrests or convictions. For example, West and Farrington (1977) found that, at age 18, 94 % of convicted boys admitted that they had been convicted, while only 2 % of unconvicted boys claimed to have been convicted. Of offenses leading to conviction, 53 % were reported accurately, 34 % were reported but minimized, 7 % were reported but exaggerated, and 6 % were not reported. The probability of arrested or convicted offenders admitting their crimes is also high in other studies, with those who have a greater number of convictions being generally more willing to self-report offenses (e.g., Brame et al. 2004).

Again, there is some indication of differential validity by gender and race in the self-reporting of official contacts. Hindelang et al. (1981) found that 76 % of white male official delinquents reported that they had been picked up by the police, compared with 52 % of white females, 50 % of black males, and only 30 % of black females. These results fostered the widespread belief that the self-reports of black youths (especially those who were the most delinquent) were less valid because of concealment. Other studies, however, have not found race differences in reporting arrests. Maxfield et al. (2000) discovered that 75 % of convicted white youths admitted being arrested, compared with 70 % of black youths (not significantly different), and 76 % of convicted males admitted being arrested, compared to 60 % of convicted females (significantly different). More recently, in a large sample of Dutch adolescents, Van Batenburg-Eddes et al. (2012) found that 62 % of those with official police contacts self-reported a police contact (interrogation in the police station), but this was somewhat higher for older (age 14–15) adolescents compared to younger (age 12–13) ones: 65 % compared to 57 % (significantly different). They also found that indigenous Dutch (65 %) and Surinamese (70 %) boys and girls were significantly more likely than Moroccan boys and girls (46 %) to report their police contacts.

The differential validity of the self-reporting of arrests was explored by Krohn et al. (2013) using data from the Rochester Youth Development Study. Overall, 61 % of arrests were

admitted. Arrested males (64 %) were more likely to admit their arrests than arrested females (46 %). Caucasian males were likely to overreport arrests, while African-American males were likely to underreport arrests. However, Krohn et al. (2013) also demonstrated that those who were arrested more frequently were least likely to report each arrest, regardless of race or gender. It could be that frequently arrested offenders had more difficulty in remembering all of their arrests.

There have been relatively few studies of how concurrent validity varies with age. However, Kirk (2006) compared age-crime curves based on official and self-reported arrests. The two curves were generally similar, but the self-report curve had a higher and earlier peak (at ages 18–19) than the official curve (at age 21). Kirk also found that the black/white prevalence ratio was higher for official arrests (2.3–1) than for self-reported arrests (1.3–1), but this was because whites overreported their arrests; 21 % self-reported an arrest compared to 13 % who had an official arrest.

Predictive Validity

Predictive validity is more impressive than concurrent validity, because being convicted may itself lead to an increase in the probability of admitting delinquent acts (Farrington 1977), perhaps because a person may assume that the researcher will know about convictions and therefore that concealment is futile. It is to be expected that current self-reported offending (of past acts) will predict future convictions, because many current self-reported offenders are also convicted and because future convictions are of course predicted by past convictions. A better test of the predictive validity of SRO surveys is to investigate the extent to which self-reports predict future convictions *among currently unconvicted people*.

There have only been four investigations (using three prospective longitudinal studies) of the predictive validity of SRO. In the Cambridge Study in Delinquent Development (a survey of

over 400 London boys), Farrington (1973) showed that, for unconvicted boys, a measure of self-reported variety of offending at age 14 significantly predicted their probability of conviction in the following 3 years. Farrington (1989) later repeated this analysis for specific types of offenses. For example, among boys not convicted of burglary up to age 18, significantly more of those who self-reported burglary up to age 18 were subsequently convicted of burglary up to age 32, in comparison with the remaining boys who denied committing burglary up to age 18 (20 % as opposed to 2 %). Similar results were obtained for vehicle theft, assault, vandalism, and drug use, but not for shoplifting.

In the first test of predictive validity using American data, Farrington et al. (1996) classified the self-reported offending of Pittsburgh Youth Study boys into four categories based on seriousness of offending and compared this with their later petitions to court. For both samples, the predictive validity of self-reports was high and statistically significant (except for drug offenses in the oldest sample). Of the boys who reported the most serious forms of delinquency, 29 % were petitioned to court for Index Violence, compared to only 7 % of the boys who reported no delinquency. This difference was statistically significant, again showing that SRO had predictive validity. Farrington et al. (1996) also examined the extent to which predictive validity varied according to race, but found that predictive validity was similar for Caucasians and African-Americans. Perhaps the most important finding with regard to the validity of SRO was that there was a substantial increase in predictive validity by combining the reports of the boys with those from other sources (mothers and teachers).

Jolliffe et al. (2003) examined the predictive validity of SRO for males and females of different racial backgrounds. In the SSDP, they found that predictive validity was high and statistically significant for most offenses (except robbery and vehicle theft). It was generally higher for males than for females and highest for Caucasians and lowest for Asians (predominantly Chinese and Filipinos). The self-reports of

Asian and African-American females had the lowest predictive validity.

Summary and Future Directions for Research

Self-reports of offending were initially developed to reveal the dark figure of crime, but they have evolved to be one of the most common outcome measures used in criminological research. On psychometric criteria, self-reports provide a reliable measure of offending. In addition, most research suggests that self-reports are concurrently valid, in that those who self-report offenses or arrests are more likely to have official records than those who do not self-report, and those who have been officially recorded tend to admit their offenses. Importantly, self-reports have predictive validity; among those with no official record, those who self-report offenses are more likely to be officially recorded in the future than those who do not self-report. This is good news for criminological researchers. The bad news is that the overall validity of self-reports conceals the fact that they appear to be more valid for some demographic categories than for others. Self-reports are clearly valid for white males, but they are less consistently valid for females, for Asians, and for African-Americans. Little is known about validity at older ages. That said, there are indications that other factors (e.g., number of arrests) might be more relevant to validity than demographic categories. More research is needed on the validity of self-reports of offending, and of self-reported arrests and convictions, to establish the optimal conditions for enquiring about these topics. Randomized experiments, such as that conducted by Enzmann (2013), are needed to study the effects of different methods of administration on the reliability and validity of different demographic categories. It is also important to establish to what extent differential validity might reflect biases in official processing. Since most criminological research is based on self-reports, it is very important that their reliability and validity should be measured.

Related Entries

- ▶ [History of the Self-Report Delinquency Surveys](#)

Recommended Reading and References

- Brame R, Fagan J, Piquero AR, Schubert CA, Steinberg L (2004) Criminal careers of serious delinquents in two cities. *Youth Violence Juv Justice* 2:256–272
- Burt C (1929) *The young delinquent*. Appleton, New York
- Elliott DS, Ageton SS (1980) Reconciling race and class differences in self-reported and official delinquency. *Am Sociol Rev* 45:95–110
- Elliott DS, Huizinga D, Ageton SS (1985) *Explaining delinquency and drug use*. Sage, Beverly Hills
- Enzmann D (2013) The impact of questionnaire design on prevalence and incidence rates of self-reported delinquency: results of an experiment modifying the ISRD-2 questionnaire. *J Contemp Crim Justice* (in press)
- Farrington DP (1973) Self-reports of deviant behavior: predictive and stable? *J Crim Law Criminol* 64:99–110
- Farrington DP (1977) The effects of public labelling. *Br J Criminol* 17:112–125
- Farrington DP (1989) Self-reported and official offending from adolescence to adulthood. In: Klein MW (ed) *Cross-national research in self-reported crime and delinquency*. Kluwer, Dordrecht, pp 399–423
- Farrington DP, Welsh BC (2007) *Saving children from a life of crime: early risk factors and effective interventions*. Oxford University Press, Oxford, UK
- Farrington DP, Knapp WS, Erickson BE, Knight BJ (1980) Words and deeds in the study of stealing. *J Adolesc* 3:35–49
- Farrington DP, Loeber R, Stouthamer-Loeber M, Van Kammen WB, Schmidt L (1996) Self-reported delinquency and a combined delinquency seriousness scale based on boys, mothers and teachers: concurrent and predictive validity for African Americans and Caucasians. *Criminology* 34:493–517
- Fergusson DM, Horwood LJ, Lynskey MT (1993) Ethnicity and bias in police contact statistics. *Aust N Z J Criminol* 26:193–206
- Glueck S, Glueck ET (1968) *Delinquents and non-delinquents in perspective*. Harvard University Press, Cambridge, MA
- Gold M (1970) *Delinquent behavior in an American city*. Brooks/Cole, Belmont
- Hardt RH, Bodine GE (1965) *Development of self-report instruments in delinquency research*. Youth Development Center, Syracuse University, Syracuse
- Harrison L, Hughes A (eds) (1997) *The validity of self-reported drug use*. US National Institute on Drug Abuse, Rockville
- Hindelang MJ, Hirschi T, Weis JG (1981) *Measuring delinquency*. Sage, Beverly Hills
- Hirschi T (1969) *Causes of delinquency*. University of California Press, Berkeley
- Huizinga D, Elliott DS (1986) Reassessing the reliability and validity of self-report measures. *J Quant Criminol* 2:293–327
- Jolliffe D, Farrington DP, Hawkins JD, Catalano RF, Hill KG, Kosterman R (2003) Predictive, concurrent, prospective and retrospective validity of self-reported delinquency. *Crim Behav Ment Health* 13:179–197
- Junger M (1989) Discrepancies between police and self-report data for Dutch racial minorities. *Br J Criminol* 29:273–284
- Junger-Tas J (2010) The significance of the international self-report delinquency study (ISRD). *Eur J Crim Policy Res* 16:71–87
- Junger-Tas J, Marshall IH (1999) The self-report methodology in crime research: strengths and weaknesses. In: Tonry M (ed) *Crime and justice: a review of research*, vol. 25. University of Chicago Press, Chicago, pp 291–367
- Kirk DS (2006) Examining the divergence across self-report and official data sources on inferences about the adolescent life-course of crime. *J Quant Criminol* 22:107–129
- Krohn MD, Thornberry TP, Gibson CL, Baldwin JM (2010) The development and impact of self-report measures of crime and delinquency. *J Quant Criminol* 26:509–525
- Krohn MD, Lizotte AJ, Phillips MD, Thornberry TP, Bell KA (2013) Explaining systematic bias in self-reported measures: factors that affect the under- and over-reporting of self-reported arrests. *Justice Quart* (in press)
- Maxfield MG, Weiler BL, Widom CS (2000) Comparing self-reports and official records of arrests. *J Quant Criminol* 16:87–110
- Short JF, Nye FI (1958) Extent of unrecorded juvenile delinquency: tentative conclusions. *J Crim Law Criminol Police Sci* 49:296–302
- Thornberry TP, Krohn MD (2003) Comparison of self-report and official data for measuring crime. In: Pepper JV, Petrie CV (eds) *Measurement problems in criminal justice research: workshop summary*. National Academy Press, Washington, DC, pp 43–94
- Tourangeau R, Smith TW (1996) Asking sensitive questions: the impact of data collection, mode, question format, and question context. *Public Opin Q* 60:275–304
- Van Batenburg-Eddes T, Butte D, Van De Looij-Jansen P, Schiethart W, Raat H, De Waart F, Jansen W (2012) Measuring juvenile delinquency: how do self-reports compare with official police statistics? *Eur J Criminol* 9:23–37
- Wallerstein JS, Wyle CJ (1947) Our law-abiding law-breakers. *Probation* 25:107–112
- West DJ, Farrington DP (1977) *The delinquent way of life*. Heinemann, London
- Williams JR, Gold M (1972) From delinquent behavior to official delinquency. *Soc Probl* 20:209–229

Semi-parametric Group-Based Method

▶ [Group-Based Trajectory Models](#)

Sensitive Questions

▶ [Randomized Response Methods](#)

Sentencing

▶ [Front Door and Backdoor Sentencing](#)

▶ [Punishment as Rehabilitation](#)

Sentencing as a Cultural Practice

Neil Hutton

Centre for Law, Crime & Justice, School of Law,
Strathclyde University, Glasgow, Scotland

Overview

Most of the scholarly literature on sentencing is written from a legal or philosophical perspective. Legal scholarship analyzes sentencing law. Philosophical work analyzes the normative debates about the aims of punishment in a liberal democratic society. This chapter examines sentencing as a cultural practice. Culture refers to sets of shared meanings or collective representations. To study culture is to examine the ways in which meanings are defined, enacted, mediated, communicated, and shared by a range of actors and audiences. A cultural analysis of sentencing is a study of how certain important meanings are represented. These include representations of moral boundaries, of justice, and of legitimate decision-making processes. David Garland argues that penal institutions have important

cultural dimensions and consequences which shape penal policies and practices.

Cultural categories, habits and sensibilities are embedded in and constitutive of our political and economic institutions. (Garland 2006)

A cultural understanding of sentencing seeks to understand sentencing as a collective practice which involves nonjudicial actors as well as judges. Garland distinguishes between different uses of the term “culture” in the sociology of punishment. On the one hand, “culture” describes a particular web of meanings which can be found empirically, for example, the local court culture of a particular jurisdiction which is shared and reproduced by the regular court actors. In this sense a culture is a more or less bounded set of customs, values, habits, and beliefs. On the other hand, “culture” can also be used analytically to describe ways of making meaning in a social setting, for example, how the sentencing process defines what it means to make just decisions about the allocation of punishment. In this sense cultural analysis is distinct from other forms of analysis, for example, political. A cultural analysis focuses on the creation of meaning, and a political analysis will focus on how a particular meaning becomes powerful and silences other potential meanings. This chapter focuses on the latter use of the concept; it provides a cultural analysis of sentencing rather than describing a particular sentencing culture (Yanow 1996).

There are three important ways in which a cultural approach provides a better understanding of sentencing.

First sentencing articulates the moral boundaries of society. In allocating different sorts of punishment to different sorts of offender, sentencing defines the boundaries between order and disorder, between the respectable and the disreputable, between the pure and the polluted, between good and evil, and between reason and emotion. Whatever particular forms punishment takes, it always performs this cultural task of ordering, separating the sacred from the profane (Smith 2008; Douglas 1966). Thus, sentencing helps to promote social solidarity.

Second, sentencing decision making performs and defends a particular definition of “justice.” A distinctive narrative of justice underpins sentencing decision making in most common law jurisdictions, even in those US states with sentencing guidelines. This narrative both purports to describe how judges reach their sentencing decisions and also provides a normative justification for these decisions. This narrative, which is commonly known as “individualized sentencing,” is described below.

Third, sentencing reproduces shared sets of meanings. It is a social practice and not just the action of an individual judge. Sentencing decision making, like other social action, is largely habitual, taken for granted, and unreflective. This does not mean that sentencers have anything less than a thoroughly professional, conscientious, and serious-minded approach to their work. But like all professionals, they work within a framework of meanings, perceptions, values, and motives that for most of the time are unquestioned. They are the taken-for-granted assumptions on which the challenging job of sentencing is based. Bourdieu calls this the “habitus” (Hutton 2006). Individualized sentencing forms the habitus for judges; it refers to the unquestioned, taken-for-granted cultural framework which defines both the way that decisions are made and also the way that these decisions are justified.

The chapter proceeds by looking at each of these three cultural tasks and in the final section analyzes the challenges to the dominant cultural approach to sentencing.

Boundaries of Moral Tolerance

From a broader Durkheimian perspective, sentencing decision making enacts deeper cultural meanings about moral boundaries which both bind society together and at the same time identify fractures and divisions (Smith 2008). Punishment expresses ideas about the sacred and the profane, about moral pollution, and about atonement and evil. As Philip Smith argues, “we can understand these basic, protean, cultural

categories running through and under what appear to be more rational modern scientific instrumental or bureaucratic tendencies.” Sentencing performs an “othering” function in all communities. There cannot be a community, “people like us,” without people who are not like us. Sentencing therefore both includes and excludes. For Smith, following Durkheim, all societies have crime and punishment but the cultural meanings of these vary. They are invoked and put into practice in a political context and are always contested and contestible.

In allocating punishment, sentencing performs the job of defining these moral boundaries. In a broad sense, this involves invoking binary classifications such as good/evil, sacred/profane, pure/polluted, safe /dangerous, and insider/outsider. Sentencing decision making involves drawing these black-and-white distinctions, but it also involves more subtle shading which blurs the apparently sharp binary division and produces distinctions which are not as clear-cut in practice as they may appear in cultural theory.

Criminalization and the decisions of police officers and prosecutors about whether to proceed further with a reported incident patrol the boundaries between criminal and noncriminal. Sentencing is really about the shape and gradation of the negative side of this binary divide and also sometimes about the potential for an offender to shift back across to the positive side. Sentencing is about separating the good guys from the bad guys. It is also about establishing just how bad the bad guys are, about the possibility or impossibility of bad guys transforming themselves into good guys, and about what opportunities may be offered to help them to change.

In sentencing, the distinction between prison and the community is symbolically crucial. In those jurisdictions where the death penalty has been abolished, prison is the most severe sanction now available. In both England and Wales and Scotland, legislation provides that custody should only be used as a last resort where no other sanction would be appropriate Padfield (2011), Tombs (2004). The removal of the individual from the community by the State signifies both

the power of the state and the subjugation of the body of the individual offender. The decision to imprison has therefore a qualitatively different significance from those other sanctions which deprive the offender of limited amounts of time, money, or association.

Aside from the custody/community tension, there are other important meanings being generated in sentencing. At the most serious end of the scale, there is a debate about how long serious offenders need to be imprisoned and the relativities both within offences (e.g., how to define different levels of seriousness of rape) and between offences (e.g., a rape and a serious assault). At the lower end, there are decisions about the boundary between fines and community sanctions (which have received less attention because of the symbolic and fiscal implications of custodial sentences) and the debates about whether fines are simply a form of economic regulation which carry little of the stigma of the other sanctions. The fine delivers pain while impacting minimally on the freedoms of movement, speech, association, and political participation that we call "liberty." The regulation of conduct virtually through automated bank transfers may be seen as a dystopian nightmare by some (Aas 2005), but it could also be the desirable freedom of a consumer society where we choose whether or not to conform and pay the price if we decide not to, literally the price of freedom (O'Malley 2009).

While it is true to say that the implementation of the criminal law performs the function of dividing conduct into acceptable and unacceptable, in practice the boundary is more accurately described as a sloping shelf, than a clearly defined wall. The boundary relates strictly speaking to actions, people can move from one side to the other. However, in practice we tend to think about criminals rather than criminal acts, and the label of criminal may persist independently of particular actions. Sentencing plays the important cultural function of defining the shape and texture of the boundary between "them and us" which turns out not to be a sharp binary distinction but a much more amorphous and liminal territory. The prison population is clearly visible

"out there," but there are crowds milling around the prison walls.

In this section we have seen that punishment is about drawing distinctions between insiders and outsiders and that sentencing is the performance of this cultural process of ordering. So sentencing is not about exclusion or inclusion but about both. Sentencing can support discourses of redemption, desistance, and rehabilitation, but it also needs to support discourses of punishment, pain, and exclusion. Sentencing therefore cannot choose between rational and emotional responses to offending; it has to be able to sustain both of these approaches.

The Discourse of Individualized Sentencing

The discourse of individualized sentencing may be summarized as follows. In reaching their decision, judges take into account all of the facts and circumstances of the individual case. Each case is composed of a very large number of relevant factors and is therefore held to be unique. No two cases are exactly the same. Judges reach their decision by an "instinctive synthesis" (R v Williscroft [1975]VR292 at 300) of these myriad factors. By defining each case as unique, this approach is able to remain silent about consistency, another important feature of liberal definitions of justice. So individualized sentencing performs a particular sort of "justice" which privileges the specificities of an individual case over the demands of consistency. Judges use the discourse of "individualized sentencing" to defend a particular approach to making just decisions. "Individualized sentencing" therefore performs a particular cultural logic which both produces "justice" in sentencing and defends this definition against its potential critics, primarily those who argue that it is possible to determine similarities between cases and that it is ethically important to treat similar cases in a similar way. "Instinctive synthesis" has a transcendent quality because it is not susceptible to further rational explanation. There is thus an element of the sacred in the discourse of

individualized sentencing. Intuitive synthesis presents sentencing decision making as being beyond the control of human agency, a function which can only be performed by those holding the office of judge.

Sentencing as a Social Practice

Sentencing is collective action made possible by shared cultural meanings and understandings. The actions of other criminal justice actors in the process play a part in shaping the judicial sentencing decision and are also shaped by this approach to decision making. Sentencing is a stage in the criminal justice process. Judges deal with cases which have been constructed by other actors. Each case proceeds through several processes of translation. The term “translation” (Latour 2005) is used here to make the point that what constitutes a “case” changes, as it passes through each stage of the process. Actors interpret information which is presented to them by others, act on this information in pursuit of their professional requirements, and pass this information on to the next set of actors in the next stage of the process. Cases are therefore constructed out of witness accounts, police reports, prosecutors’ professional practices, reports written for the court by social workers, medical professionals, forensic scientists, psychologists, psychiatrists, and defense pleas in mitigation (Castellano 2009; McNeill et al. 2009; Sudnow 1965; Tata 2007; Yngvesson 1989). When judges pass sentence, their understanding of the case comes from reading documents and sometimes listening to evidence and argument. They have no unmediated access to the event that gave rise to the case. Thus, by the time that the case has reached the sentencing stage, it is defined in a particular way which already limits the options open to the judicial decision maker. The case is a more or less familiar narrative which prompts a more or less familiar ending.

So sentencing is not simply about judges nor is it solely about individual cases nor about individual decisions. Sentencing is a social practice and also a cultural practice.

Sentencing is the social and cultural performance of a particular definition of justice. The term “performance” is chosen deliberately to emphasize the importance of action. Justice is not something which only exists in philosophical debates or in sentencing texts. Justice has to be enacted every day in the decision-making processes of actors working in the criminal justice system. One of the most important cultural tasks of the sentencing process is to persuade audiences that sentences passed by the court are just. Sentencing is therefore about the creation of meaning performed by an interpretive community. Sentencing is a means of communicating, persuading, convincing, justifying, and providing an account of decisions which have a profound impact on individual offenders and carry powerful messages both about the kind of society we live in and the kind of society we might want to create.

Philosophical Aims of Sentencing

From a philosophical perspective, the main issue related to sentencing is the moral question of how punishment can be justified and what should be the aims of punishment (Honderich 1989). The infringement of individual liberty and/or the imposition of pain by the state requires to be both lawful and morally justifiable. Debates range about whether punishment should be backward looking and impose punishment proportionate to the seriousness of the offence ((just deserts) or whether punishment should be forward looking and seek to have an impact on reducing crime by deterring individuals or the general population, by making reparation to the community or the victim, by rehabilitating the offender, or by protecting the public. Most jurisdictions adhere to these aims at a general systemic level, but in terms of sentencing decision making, it is not clear what impact, if any, the pursuit of aims plays in selecting the type and severity of sentence in individual cases. This “cafeteria” approach (Ashworth 2010) offers great flexibility as any sentence can be fitted into one or more of these justifications. It enables sentences to be justified on a case-by-case basis. However, the flexibility of this suite of aims is also a weakness.

A sentence which appears to be fair in comparison with sentences for similar cases may not be an appropriate sentence for protecting the public or for enabling the rehabilitation of the offender. The cafeteria approach does not encourage any systematic or general policy approach to sentencing which makes it difficult to provide a more general rationale for sentencing to the public. Decisions are made on a case-by-case basis and if necessary justified post hoc with an account which can include any philosophical justification or combination of justifications. From a cultural perspective, the definition of justice is controlled by the judiciary. The mode of accountability is the public trust in the office of the judge.

Thus, although sentencing is a collective social and cultural practice, the way in which decisions are justified and defended is located at the level of the individual judicial decision maker. The following section describes the cultural framework of sentencing. It shows how individualized sentencing, seen as an approach to making and justifying decisions about justice, promotes particular values, allocates the power to define what is to count as justice to particular agencies and processes, and adopts particular rhetorical devices to persuade audiences of its propriety and legitimacy.

Sentencing as the Performance of Justice

Sentencing decisions perform justice. They define what counts as a “just” decision. These decisions also need to be justified. The discourse of individualized sentencing needs to explain how decisions are just and also needs to demonstrate that the process by which these decisions were reached was legitimate and persuade audiences of the propriety and correctness of this decision-making process. In this sense sentencing shares much in common with administrative decision making. Recent work in the cultural sociology of administrative justice provides a useful framework with which to analyze the ways in which sentencing seeks to construct just decisions and defend these decisions.

Sentencing is a legal decision in so far as decisions are made by judges, have the authority of the court, and are therefore legitimate. However, the way in which sentencing decisions are made accountable and justifiable has as much in common with administrative decision making as it does with judicial decision making. Although sentencing decisions are made by judges, the procedural form and mode of justification through which these decisions are made and defended are, in important ways, quite unlike other judicial decisions. Kagan (2010) argues that the characteristic function of an administrative decision is to get the work of society done while the characteristic function of a legal decision is to establish the legal coordinates of a situation in light of preestablished legal rules. In terms of Kagan’s typology, sentencing decisions look more like administrative decisions than legal decisions.

Individualized sentencing is more about doing the job of making a “just” decision in an individual case than it is about applying rules to facts. “Rules are based on generalisations” (Kagan 2010, p. 11). A rule says that if A and B and C obtain, then the appropriate decision is X because X will advance the policy goals of the organization. In the choice of a type and amount of sentence, there are no rules which specify that if circumstances A, B, and C exist, then the judge must pass amount Y of sanction X (with the exception of mandatory life sentences in some jurisdictions). For example, there might be a well-known customary practice whereby offenders convicted of injury with a weapon which causes permanent disfigurement will receive a prison sentence, but this is a custom not a rule. The judicial decision about the appropriate type and severity of sanction is not rule governed, and thus the accounts provided to justify these decisions cannot be rule based.

Administrative decision making requires the decision maker to respond to complex and rapidly changing social circumstances which are not susceptible to the generalizations applied in fixed legal rules. Administrative decision making requires experience, expertise, and specialization to enable outcomes to be adjusted to meet the

peculiarities of each individual case. Judicial sentencing decision making is based on the exercise of discretion in individual cases to deal with complex and multiple facts and circumstances. It is in this sense that sentencing more resembles administrative decision making than legal decision making. This is not to argue that sentencing is an administrative decision but rather that the style of decision making employed and its modes of accountability have more in common with the processes commonly employed to defend “just” decisions in an administrative context. “Individualized sentencing” thus serves as an important means of justifying decisions but not necessarily as an accurate account of the empirical process of making decisions.

Mashaw has argued that administrative justice refers to

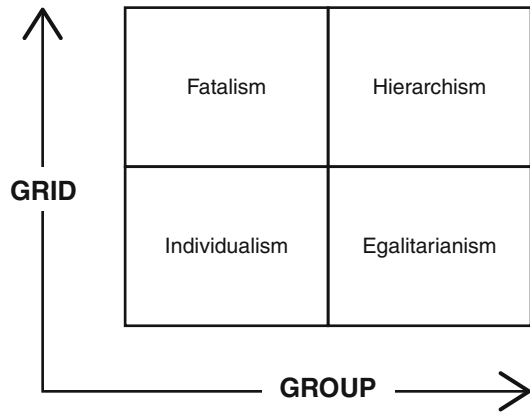
“the qualities of a decision process that provide arguments for the acceptability of its decisions”. (Mashaw 1983, p. 16)

Justice, in this context, is about the decision maker being able to provide an account of the decision process that enables the public to make a judgement about whether or not the decision is a fair and reasonable decision (Tata 2002). Members of the public may disagree with the substantive decision, but still be satisfied that it was reached by a fair and reasonable process. Individualized sentencing serves a similar purpose for sentencing. A “just” sentencing decision is reached by taking into account all of the facts and circumstances of the individual case and coming to a judgement about the type and level of sanction required.

Christopher Hood has produced a cultural analysis of administrative justice which is employed here to help understand how “individualized sentencing” defines and puts into practice a particular concept of “justice” and also to identify the challenges to this discourse based on an understanding of its weaknesses (Hood 1998).

Legitimacy and Accountability

Mary Douglas famously proposed that social relations between individual actors and the institutional arrangements for managing these



Sentencing as a Cultural Practice, Fig. 1 Grid-group analytical framework of cultural biases (From Halliday and Scott (2010))

individuals could be classified along two variables which she called group and grid (Douglas 1982 quoted in Halliday and Scott 2010) (Fig. 1).

In the context of decision-making processes, grid refers to the level of externally imposed prescriptions experienced by a decision maker. High grid means the decision makers are relatively strictly bound by rules/prescriptions over which they have little or no control, and low grid means the opposite: that decisions makers have wide-ranging discretion and are relatively unconstrained by rules. Group refers to the extent to which individuals see themselves as being incorporated into bounded units, as sharing a sense of the collective. In terms of justice, a high-grid approach believes that “justice” is most effectively delivered through a closed and controlled process and a low-grid approach implies greater openness and flexibility. A high-group approach attempts to deliver collective values, a low-group approach prioritizes individual self-interest. Group addresses questions about political legitimacy, that is about who is authorized to make decisions. Grid addresses questions about modes of accountability, about what processes make decisions transparent. So this refers to how individuals both perceive their social world and how they are able or not able to understand, challenge, and participate in decision making. It refers to the relationship between the state

and the individual, a fundamental concern of political theory and law.

Examining behavior across these two dimensions gives four ideal types which can be used to exhaustively classify decision-making processes: hierarchism (high grid high group), individualism (low group high grid), egalitarianism (high group low grid), and fatalism (low group low grid). These are normative ideal types which prioritize different values in decision making. Hierarchism emphasizes expertise and skills; egalitarianism emphasizes citizen participation; individualism prioritises competition; and fatalism favours stoicism and/or serendipity. As with all ideal type analysis, these are highly unlikely to exist in their pure forms in the observable world. Rather, particular decision processes will exhibit varying degrees of some or indeed all of the types, usually one being predominant.

Individualized sentencing exhibits high-group features. Authority over sentencing decision making is shared between the legislature and the judiciary. The legislature set the legal framework which judges implement. Judges are entrusted to act on behalf of the collective. In practice, the legislature leave the judiciary with wide discretion to make sentencing decisions. Individualized sentencing provides a persuasive account of how just sentences are made in order to promote the collective good.

The issue of grid is more complex. At first sight, the relative paucity of rules might suggest that sentencing is low grid, but a closer examination of the discourse of individualized sentencing suggests that it is more accurately classified as high grid. The discourse of individualized sentencing effectively restricts access to decision making to the judiciary. Individualized sentencing presents itself as the single correct way to make just sentencing decisions, an approach which can only be implemented by judges and experts by virtue of their education, knowledge, and experience and which is thus inaccessible to ordinary members of the public. Sentencing decisions are justified on a case-by-case basis. There is no set of rules which can be applied to provide a justification for the sentence chosen in each case. Judges are expected and entrusted to

exercise their skill and judgement for public benefit. Citizens are not expected to participate. Individualized sentencing may not have the transparency of a grid, but it has the rigidity and stability of a grid. Individualized sentencing is thus firmly a hierarchist form of decision making.

In an individualized sentencing regime, sentencing is controlled by the judiciary. Other agents and officials play a part in the decision-making process. Defense agents have the opportunity to make a plea in mitigation. Reports provided for the court by probation or social work staff provide an opportunity for these professional groups to contribute to the decision-making process. However, ultimately the decision is made by the judge. Sentencing is “owned” by the judiciary (Tata et al. 2008). Individualized sentencing is thus substantive rather than formal, irrational rather than rational, and hierarchical rather than participatory. Hierarchical, in so far as the judge controls the process and standards for decision making and yet informal in so far as the authority of the decision, rests with the power of the decision maker and not with detailed legal rules.

Challenges to Hierarchism in Sentencing Egalitarianism

In most late modern societies, confidence in authorities and experts is diminishing (Garland 2001). The judiciary are no exception. Judges seem to be losing public trust and may no longer be able to rely on their education, status, and tradition to sustain public confidence (Hough and Roberts 2004). The perceived monopoly of the judiciary as the “owners” of sentencing is being challenged. Kagan (2010) argues that political leaders will be comfortable with an expert judgement system so long as the electorate displays high levels of trust in the experts. However, where there is evidence that trust is in decline, leaders are likely to seek to impose a more formal and legalistic decision system with clearer accountability.

There is evidence of an increasing desire amongst citizens to have a greater input into sentencing decision making. Single-interest groups in many jurisdictions particularly those

representing victims of crime and their families seek to have a greater influence on sentencing decision making. Some US states have a strong populist tradition which enables more direct citizen involvement in sentencing (California's public initiatives). A variety of sentencing institutions, usually known as councils or commissions, have been introduced in many common law jurisdictions (Hutton 2008). Some of these have the power to establish sentencing guidelines; others provide public information, education, and advice. Most of these institutions offer citizens an opportunity to influence sentencing decisions, but the discourse of individualized sentencing has proved highly resistant to change because of the powerful cultural messages that it communicates.

The history of the sentencing guidelines movement (Tonry 1996) is the story of attempts to provide a more structured approach to sentencing which allocates greater significance to the pursuit of consistency in sentencing rather than the delivery of a just sentence in each individual case. This also makes sentencing decision making more transparent and accountable and therefore more susceptible to rational debate. Individualized sentencing invites challenges about who makes sentencing and policy and what the aims for such a policy should be.

Individualism

It is hard to imagine sentencing being operated by the market, although there is increasing private sector involvement in the administration of punishment. However, it is not so hard to imagine a market approach being used to assess sentencing. In most areas of government spending such as health or education, the executive is responsible and accountable for setting budget priorities for the expenditure of public funds to manage activities in these areas. Ministers are responsible for ensuring that their officials make their decisions in a just manner, ensuring, for example, that appropriate processes have been observed and that the decisions are fair and reasonable. In common law jurisdictions where there are no comprehensive sentencing guidelines, the aggregate of sentencing decisions is, *de facto*, the sentencing

policy for these jurisdictions. In this sense sentencing may be seen as a decision-making process which effectively allocates scarce penal resources in a particular way and provides a form of public justification for these decisions.

In an individualized sentencing approach, justice is defined with reference to an individual case and wider policy considerations about how sanctions might be allocated in a more cost-effective manner are not relevant. Decisions about justice are made by judges not policy makers although the allocation decisions of judges have significant public policy impact. Many US states are seeking to exert greater regulatory control over judicial decision making in sentencing to enable the government to exercise greater control over penal expenditure (National Conference of State Legislatures 2011). These pressures are being felt across Western jurisdictions. Politicians are keenly aware that justice clearly has a price and that for many jurisdictions, the current price is no longer affordable.

Fatalism

Fatalism is the belief that decision making is an unpredictable lottery. It can lead to cynicism and populist anger or it can lead to quietism and tolerance. There is also a strong element of fatalism in so far as citizens feel unable to influence sentencing policy and practice which is widely perceived to belong to the judiciary rather than the province of elected governments.

The strength of fatalism as a cultural bias should not be underestimated. It almost goes without saying that governments and judges need to see themselves as making positive and constructive decisions, as both being able to and having a duty to "make a difference." At the same time individuals judges, ministers, and policy makers will from time to time feel that their job is impossible. Crime will never go away and it is hard to find evidence of effective ways of dealing with crime. The media hold politicians responsible and it is not hard for them to find instances of failure. This is a cultural perspective, not an objective fact. The inevitability of crime and the apparent limitations of punishment to control crime need not lead to despair or cynicism, but

rather to different ways of defining the issue and different modes of action to develop higher-group identification.

From a cultural perspective, decisions about justice in sentencing are justified by a hierarchist approach which is based on public trust in the discretionary decision making of individual judges on a case-by-case basis. This approach prioritizes particular values, most significantly those of professional expertise and experience and intuitive moral judgement. The inherent weaknesses in this approach are challenged by approaches which want to give a higher priority to different values, for example, a more rational and transparent form of accountability and a greater involvement of citizens and interest groups in decision making.

Conclusion: Why Is the Cultural Analysis of Sentencing Decision Making Useful?

Cultural analysis is useful to examine both stability and change in sentencing. In cultural terms, individualized sentencing has proved exceedingly durable and resistant to criticism. The discourse allows judges to retain control over the ways in which sentencing decisions are made and, perhaps more significantly, over how decision making is justified. The idea of fitting a sentence to the particularities of each individual case is fundamental to common sense conceptions of justice, and individualized sentencing deploys classic rhetorical devices to persuade audiences that it is the only process which can guarantee the production of justice in sentencing. To that end, justice is held to depend on scrupulous consideration of the detailed facts and circumstances of each case. The decision is made by judges and members of a respected, principled, and almost sacred profession, located above the compromised, pragmatic, world of politics. Their method is archaic, mysterious, and sacerdotal. Concerns about consistency, accountability, and rationality are presented as misguided attempts to objectify something which is inherently subjective. This representation of sentencing chimes with our common sense understandings of crime

as a matter of individual responsibility, attributable to bad moral decisions made by individual offenders. It reflects the narratives of crime and punishment that form the basis of crime fiction, television crime drama, and Hollywood movies. At least part of the explanation for the resilience of individualized sentencing as an approach to sentencing lies in its insistence that “just” punishment is about the right punishment for an individual offender.

At the same time, like all justice discourses, the discourse of individualized sentencing contains the seeds of its own potential demise. As its strength is dependent on public trust in the judiciary, so it is vulnerable as this trust is perceived to be waning. Hood (1998) argues that cultural change always emerges from the perceived weaknesses of an existing state of affairs as opposed to being driven by utopian theoretical proposals premised on the possibility of starting with a blank sheet of paper. A cultural understanding tells us that any approach to decision making has both strengths and weaknesses. Change often results from the perceived failure of a process being attributed to its inherent weaknesses and solutions being sought in processes which exhibit strengths in these areas. Thus, the “problem” with individualized sentencing is its lack of transparency and limited capacity to hold judges accountable for their decisions, just as the problem with systems of sentencing guidelines is their perceived inability to take proper account of the distinctiveness of each individual case (Aas 2005).

A cultural analysis allows a better understanding of the traditional common law approach to sentencing. It shows that individualized sentencing is a political choice, not an inevitability, and that it is susceptible to challenge by approaches which offer a different version of what it means to produce justice in sentencing, for example, approaches which give a higher priority to consistency, accountability or effectiveness, and value for money. It shows that the approach to sentencing decision making in a jurisdiction arises out of political contest between politicians, administrators, judicial officers, third sector organizations, and other groups, all of which is

conducted in a particular time and place and which is represented in various forms of media. It shows that there is no single objective definition of justice but rather a range of definitions over which there is continual contest.

Recommended Reading and References

- Aas KF (2005) *Sentencing in the age of information: from Faust to MacIntosh*. Glasshouse Press, London
- Ashworth A (2010) *Sentencing and criminal justice*. Cambridge University Press, Cambridge
- Bourdieu P (1977) *Outline of a theory of practice*. Cambridge University Press, Cambridge
- Castellano U (2009) Beyond the courtroom workgroup: case workers as the new satellite of social control. *Law and Policy* 31(4):429–462
- Douglas M (1966) *Purity and danger*. Routledge and Kegan Paul, London
- Douglas M (1982) Introduction to group -grid analysis. In: Douglas M (ed) *Essays in the sociology of perception*. Routledge and Kegan Paul, London
- Garland D (2001) *The culture of control*. Oxford University Press, Oxford
- Garland D (2006) Concepts of culture in the sociology of punishment. *Theor Criminol* 10(4):419–447
- Halliday S, Scott C (2010) A cultural analysis of administrative justice. In: Adler M (ed) *Administrative justice in context*. Hart Publishing, Oxford
- Honderich T (1989) *Punishment: the supposed justifications*. Polity Press, Cambridge
- Hood C (1998) *The art of the state: culture, rhetoric and public management*. Oxford University Press, Oxford
- Hough M, Roberts JV (2004) *Public confidence in criminal justice: an international review*. King's College, Institute for Criminal Policy Research, London
- Hutton N (1995) Sentencing, rationality and computer technology. *J Law Soc* 22(4):549–570
- Hutton N (2006) Sentencing as a social practice. In: Armstrong S, McAra L (eds) *Perspectives on punishment: the contours of control*. Oxford University Press, Oxford
- Hutton N (2008) Institutional mechanisms for incorporating the public in the development of sentencing policy. In: Freiberg A, Gelb K (eds) *Penal populism: sentencing councils and sentencing policy*. Willan Publishing UK/Federation Press, Australia
- Kagan RA (2010) The organisation of administrative justice systems: the role of mistrust. In: Adler M (ed) *Administrative justice in context*. Hart Publishing, Oxford
- Latour B (2005) *Reassembling the social: an introduction to actor network theory*. Oxford University Press, Oxford
- Mashaw J (1983) *Bureaucratic justice: managing social security disability claims*. Yale University Press, Newhaven
- McNeill F, Burns N, Halliday S, Hutton N, Tata C (2009) Risk, responsibility and reconfiguration: penal adaptation and misadaptation. *Punishment Soc* 11(4):419–442
- National Conference of State Legislatures (2011) *Principles of effective state sentencing and corrections policy*. NCSL, Washington, DC. www.ncsl.org
- O'Malley P (2009) *The currency of justice: fines and damages in consumer societies*. Routledge Cavendish, London
- Padfield N (2011) Time to bury the custody “threshold”? *Crim Law Rev* 8:593–612
- Reitz K (2005) The enforceability of sentencing guidelines. *Stanford Law Rev* 58(1):155–173
- Smith P (2008) *Punishment and Culture*, University of Chicago Press, Chicago.
- Sudnow D (1965) Normal crimes: sociological features of the penal code in the public defenders office. *Soc Probl* 12(3):255–276
- Tata C (2002) Accountability for the sentencing decision process : towards a new understanding. In: Hutton N, Tata C (eds) *Sentencing and society*. Ashgate, Aldershot, pp 399–420
- Tata C (2007) Sentencing as craftwork and the binary epistemologies of the discretionary decision process. *Soc Leg Stud* 16(3):425–447
- Tata C, Burns N, Halliday S, Hutton N, McNeill F (2008) Assisting and advising the sentencing decision process: the pursuit of ‘quality’ in pre-sentence reports. *Br J Criminol* 48(6):835
- Tombs J (2004) A unique punishment? Scottish Consortium Crimed Criminal Justice. <http://www.scccj.org.uk/wp-content/uploads/2011/08/A-Unique-Punishment.pdf>
- Tonry M (1996) *Sentencing matters*. Oxford University Press, Oxford
- Yanow D (1996) How does a policy mean? Interpreting policy and organisational actions. Georgetown University Press, Georgetown
- Yngvesson B (1989) Inventing law in local settings: rethinking popular legal culture. *Yale Law J* 98(8):1689–1709, 6224 words August 2012

Sentencing Commissions

Michael M. O’Hear

Marquette University Law School, Milwaukee, WI, USA

Overview

Sentencing commissions and sentencing guidelines were the preeminent institutional

expressions of the sentencing reform movement that emerged in the United States in the 1970s and touched most American jurisdictions by the end of the twentieth century. This entry focuses on the history, design, and operation of American sentencing commissions. More than half of the American states have established sentencing commissions at one time or another, although several proved short-lived. At least a few commissions are believed to have played important roles in their states in advancing such ideals as uniformity, proportionality, and cost-effectiveness in sentencing policy and practice. Commissions are able to exercise influence through at least three distinct functions: collecting and analyzing sentencing-related data, advocating for or against new policy proposals in the legislature, and developing guidelines for use by sentencing judges.

The topics of sentencing commissions and sentencing guidelines are inevitably intertwined. Indeed, the core functions of many of the leading commissions have included developing and superintending guidelines. Any history or assessment of these commissions must therefore at least touch upon questions of guidelines design. However, a detailed treatment of such questions lies beyond the scope of this entry and is presented elsewhere.

Similarly beyond the scope of this entry is the history and function of sentencing commissions outside the United States, which include, for instance, commissions past or present in England, Canada, and Australia.

This entry's primary theme is the relationship between sentencing commissions and legislatures. Although sentencing commissions are predominantly legislative creations, commissions have often struggled to maintain their relevance in the face of ongoing legislative policymaking in the sentencing area, which frequently takes the form of harsh statutory responses to the "crime du jour." A secondary theme is the relationship between commissions and judges – another relationship that has sometimes proven quite challenging for the commissions to manage effectively.

Fundamentals

This section describes an early, influential articulation of the sentencing commission concept, presents case studies of two of the longest-functioning and most carefully studied commissions, and then more briefly surveys the history and design of commissions in other states.

Development of the Sentencing Commission Concept

United States District Judge Marvin E. Frankel is often credited with first proposing the idea of a sentencing commission in the early 1970s. Indeed, in his most well-known work, *Criminal Sentences: Law Without Order* (1973), Frankel himself characterized the sentencing commission as "the most important single suggestion in this book" (119). In particular, he saw a permanent sentencing commission as a way to mitigate systemic problems with legislative oversight of sentencing policy: "[T]he subject of sentencing is not steadily exhilarating to elected officials. There are no powerful lobbies of prisoners, jailers, or, indeed, judges, to goad and reward. Thus, accounting in good part for our plight, legislative action tends to be sporadic and impassioned, responding in haste to momentary crises, lapsing then into the accustomed state of inattention" (119).

Frankel envisioned three distinct roles for his commission. First, the commission would function as a sort of sentencing think tank, both collecting and synthesizing scholarly work in the field and designing and conducting its own studies. This function reflected Frankel's view that "the subject [of sentencing] will never be definitively 'closed,' that the process [of sentencing reform] is a continuous cycle of exploration and experimental change" (118–119). Second, putting its research and expertise to practical use, the commission would serve "as a kind of 'lobby'" with regard to sentencing issues (122). Frankel seemed to think that the commission would, in particular, represent the interests of prisoners and jailers. Indeed, he made a point of emphasizing that former or present prison inmates ought to serve as members of the

commission (along with social scientists, legal professionals, business people, and artists) (120). Third, and finally, the commission would itself have some power to enact rules. Frankel did not precisely delineate the commission's jurisdiction relative to the legislature's, but suggested that the commission "could prescribe in rules of general application the factors to be considered in individual sentences, the weight assignable to any specific factor, and details of sentencing and parole procedures" (123). Thus, it seems that Frankel contemplated that the commission would promulgate what we would now call sentencing guidelines.

Minnesota's Sentencing Commission

It did not take long for Frankel's idea to achieve practical realization. In 1978, the Minnesota legislature created a sentencing commission as part of a broader package of reforms that also included the abolition of discretionary parole and the authorization of sentencing guidelines.

In broad outline, Minnesota's commission has performed each of the three roles described by Frankel. First, as to the think tank role, the commission's enabling statute directs that it "shall serve as a clearinghouse and information center for the collection, preparation, analysis and dissemination of information on state and local sentencing practices, and shall conduct ongoing research regarding Sentencing Guidelines, use of imprisonment and alternatives to imprisonment, plea bargaining, and other matters relating to the improvement of the criminal justice system" (Minn. Stat. § 244.09(6)). In furtherance of this role, the commission has gathered a "huge inventory of sentencing and corrections information," which is said to be "one of the most comprehensive and detailed databases ever assembled by any state" (Frase 2005a, b). Of particular importance has been the commission's research on pre-guidelines sentencing practices, which informed the development of the guidelines; on the corrections impact of the guidelines and proposed amendments; and on changes in sentencing practices since the implementation of the guidelines (Knapp 1987).

Second, as to the lobbyist role, the commission has participated regularly and sometimes quite successfully in legislative processes since 1978. This is not to say, however, that the commission's positions have always prevailed. For instance, a commission initiative in the 1990s to reduce the impact of plea bargaining on sentences ran into legislative resistance and proved only partially successful (Frase 2005a, b). It should also be noted that, while the commission has added an expert, well-informed voice on sentencing policy to the legislative process, it has not been quite so broadly representative a body as Frankel envisioned, at least based on its formal composition. Under the terms of its original enabling statute, the commission was to be comprised of three judges, a public defender, a prosecutor, a corrections representative, a parole board representative, and two members of the public. The commission has since been expanded to include a crime victim and a police officer. Missing, though, are social scientists, business people, artists, and (perhaps most significantly from Frankel's perspective) former or present prisoners.

Finally, as to the lawmaking role, the commission was statutorily required to promulgate sentencing guidelines for felony cases and authorized to make later amendments as necessary, subject to legislative override. The legislature initially gave the commission wide latitude in deciding which purposes and factors to emphasize in the guidelines. Notable decisions made by the commission in developing the initial guidelines, which took effect in 1980, included decisions to adopt a "prescriptive" instead of a "descriptive" approach (i.e., to seek to selectively change, rather than merely perpetuate, existing sentencing practices); to treat prison capacity as a firm constraint on guidelines severity (which implied that increasing severity levels for one offense would require corresponding severity decreases for other offenses); and to emphasize desert and offense-based considerations as primary determinants of sentence length. (Because it also gave significant weight to criminal history, the Minnesota approach was sometimes referred to as "modified desert" (Frase 1997)).

Since 1980, although the commission has retained lawmaking authority, the legislature has periodically pared back the wide policymaking discretion originally given to the commission and asserted its own lawmaking authority in the sentencing area. Often, these legislative interventions were of the type that seemed of greatest concern to Frankel: “sporadic and impassioned, responding in haste to momentary crises.” For instance, in 1989, in response to a spike in violent crime, including a recent series of high-profile sexual assaults, the legislature adopted new mandatory minimum sentences for violent, sexual, and drug offenses, effectively overriding the commission’s more restrained response to the crime wave (Frase 2005a, b). The legislature also amended the commission’s enabling act to specify that “public safety” should be the “primary consideration” in establishing and modifying the sentencing guidelines – an implicit rebuff of the commission’s emphasis on desert and prison capacity. Similarly, in 1992, after the rape and murder of two female college students, the legislature doubled the guidelines’ presumptive sentences for sex offenses and imposed new mandatory minimums. In short, if an important objective of sentencing commissions is to preempt “sporadic and impassioned” penal legislation, the Minnesota commission has hardly been an unqualified success.

Yet, despite its periodic marginalization in the policymaking process, there are good reasons to regard the history of the Minnesota commission as a demonstration of the merit of Frankel’s original vision. Indeed, the bare fact that the commission has survived and remained institutionally relevant for 35 years – operating in a deeply politicized field that has been repeatedly buffeted by waves of public outrage and wild swings in intellectual fashion – must itself be counted a notable success. With a half-dozen professional staffers, the commission continues to produce a detailed annual statistical report for the legislature, as well as ad hoc reports on issues of particular concern. The commission also regularly adjusts the guidelines in light of new legislation and other considerations, reports on the projected fiscal and racial impact of legislative proposals,

and offers training and technical assistance for guidelines users (Minnesota Sentencing Guidelines Commission 2012). Through such activities, the commission has likely contributed to uniformity, rationality, and cost-effectiveness in Minnesota’s sentencing policies and practices (Frase 2005a, b).

The commission may thus deserve some credit for Minnesota’s remarkably low incarceration rate and success in avoiding the extreme prison overcrowding that has plagued so many other American states. On the other hand, Minnesota’s incarceration rate was already quite low at the time the commission was formed, which suggests that preexisting aspects of the state’s legal or political culture may have played an even more important role than the commission in restraining penal excess during the “get-tough” era of the 1980s and 1990s. Moreover, Minnesota’s prison population has grown dramatically since 1978, even if the rate of growth has been somewhat below American norms. The commission appears to have been especially ineffective in resisting upward severity pressure in the politically charged areas of sex and drug crimes; this ineffectiveness has arguably resulted in sentences in these areas that are markedly out of step with the “just deserts” philosophy that originally animated the Minnesota guidelines (Stuart and Sykora 2011). Additionally, it should be noted that Minnesota has one of America’s highest levels of racial disparity in its prison population. While these disparities seem to result largely from a complex interaction of social disadvantage, behavioral differences, and policing strategies, the commission’s decisions on calculating and weighing criminal history in the guidelines may also play an important, if unintentional, role in exacerbating the disparities (Frase 2009).

United States Sentencing Commission

Although the Minnesota sentencing commission may claim the title as America’s oldest, the federal commission has undoubtedly been the nation’s most prominent and intensely scrutinized. Enmeshed in a political culture that seems far more polarized than Minnesota’s, the federal commission has endured repeated,

vociferous attacks from the both the left and the right. Perhaps even more than the Minnesota experience, the federal experience casts considerable doubt on the capacity of sentencing commissions to cure the political pathologies that Frankel associated with penal lawmaking.

As with Minnesota, proposals for a federal commission and guidelines date to the 1970s, although enabling legislation was not enacted until 1984. Once created, the federal commission faced a statutory mandate very similar to Minnesota's: gather and analyze data, make recommendations to Congress, and promulgate (and as necessary amend) sentencing guidelines (28 U.S.C. § 994). The federal commission was to be more robe-heavy than Minnesota's (three of seven voting members must be judges), but the federal statute was otherwise less prescriptive as to composition, leaving the matter to the dynamics of Presidential nomination and Senate confirmation.

One might imagine that appointment by a single president would result in a cohesive initial group of commissioners. It turned out, however, that the first commissioners were deeply divided philosophically, and one of them even publicly dissented from the guidelines the commission eventually promulgated. These divisions seem to have had important consequences for the guidelines' structure and may have contributed to perceptions that the guidelines embodied a set of unprincipled compromises (Stith and Cabranes 1998). In fairness to the first federal commission, though, it must be noted that the federal enabling statute, unlike Minnesota's, included a long list of directives regarding the content of the guidelines, many of which seemed to point in different directions (O'Hear 2006). In any event, notable decisions made by the federal commission in developing the initial guidelines, which took effect in 1987, included decisions not to endorse an overarching purpose of sentencing (in contrast to the Minnesota commission's embrace of desert); to employ the descriptive approach (except in the several important areas in which the enabling statute expressly favored harsher sentences); and to measure offense severity based not merely on the formal offense of conviction, but on a plethora of "real-offense"

factors. The latter decision resulted in a set of guidelines marked by an extraordinary degree of complexity, which (along with their aggressive implementation of Congress's severity-enhancing directives) has been one of the chief sources of their unpopularity.

Since the guidelines' promulgation, the federal commission's path has paralleled that of the Minnesota commission: a strong record of data-collection and analysis has been accompanied by a more mixed record of resisting legislative interference with the integrity of the original guidelines system. With a staff of about 100, the federal commission's basic activities are similar to those of its Minnesota counterpart (publication of reports, amendment of guidelines, technical support for guidelines users, and so forth) but on a considerably larger scale.

Many commentators have effectively documented the troubled relationship between Congress and the commission, especially Congress's tendency to override the commission's policy choices in areas of particular public concern (e.g., Friedman and Supler 2008; Bowman 2004). One illustration may suffice for present purposes. In 1986, in response to a public outcry over crack cocaine and without waiting to see how the commission would deal with drugs in its pending guidelines, Congress enacted new mandatory minimum sentences for drug offenders. The law was especially tough on crack offenders: it only took 5 g of crack to trigger a 5-year minimum and 50 g for a 10-year minimum. Importantly, the corresponding weight thresholds for powder cocaine were 100 times greater, resulting in a wide disparity in the treatment of powder and crack defendants who were trafficking in similar volumes. Although not expressly required to follow suit, the commission nonetheless chose to incorporate the 100:1 ratio into its drug guidelines, thus extending the disparity across the full range of covered quantities, including those below the mandatory-minimum thresholds. Subsequent analysis by the commission, however, casts doubt on earlier assumptions that the crack form of cocaine was intrinsically more dangerous than the powder form, and also identified troubling racial disparities that arose from the 100:1 ratio

(crack defendants were disproportionately black, while powder defendants were disproportionately white). Accordingly, in 1995, the commission promulgated an amendment in order to equalize the guidelines' treatment of crack and powder. Congress, however, rejected the amendment. In a series of reports over the following dozen years, the commission continued to urge Congress to soften the 100:1 ratio but to no avail. Finally, in 2007, the commission was permitted to reduce the ratio in the guidelines without congressional interference, and Congress itself then followed suit by reducing the statutory ratio to 18:1 in 2010. Although the commission eventually had some success in achieving legal change, the protracted nature of the process hardly demonstrates a high level of congressional deference to the commission's expert judgments.

In a thorough assessment of the first 15 years of sentencing under the federal guidelines, the commission claimed credit for an increase in transparency and predictability in sentencing, as well as a decrease in inter-judge disparity (United States Sentencing Commission 2004). The commission observed that it had developed "huge" databases, representing the "richest sources of information that have ever been assembled on federal crimes, federal offenders, and sentences imposed," and that its prison impact model had proven reliable in projecting prison bed and supervision needs. And those needs had become considerable; the commission noted a large increase in sentencing severity over the guidelines era. Not only were a larger percentage of federal defendants receiving prison terms, but those who were sent to prison were facing terms that were twice as long, on average, as they had been before the guidelines. Some of the increase in severity was due to congressional policy choices, such as the 1986 mandatory minimums for drug offenders, but the commission's policy choices also played a role, as in the initial decision to integrate the statutory 100:1 ratio into the guidelines. Moreover, the increase in severity was borne disproportionately by black and Hispanic defendants; the gap in average sentences between white and minority defendants grew rapidly in the guidelines era. Much of the

black-white disparity was attributed to the then-prevailing 100:1 ratio.

Shortly after the 15-year report, the United States Supreme Court, in its 2005 decision in *United States v. Booker*, transformed the federal guidelines from mandatory to advisory. (Note that this decision affected only the federal guidelines, leaving intact the mandatory or presumptive character of the sentencing guidelines in Minnesota and a number of other states.) *Booker* fundamentally changed the role of the federal commission, which could no longer simply dictate policy to sentencing judges; rather, if the commission wished to maintain judicial compliance with the guidelines, it had to persuade judges that its policy choices actually merited deference. The Supreme Court made this clear in its 2007 decision in *Kimbrough v. United States*, in which the Court held that sentencing judges were no longer bound by the guidelines' 100:1 ratio. In so holding, the Court observed that the crack guidelines did not exemplify the commission's exercise of its "characteristic institutional role" because the commission relied on the statutory minimums and did not take account of "empirical data and national experience." The lesson seems to be that the commission's efficacy as a policymaking body in the future may depend in large part on its conformity to Frankel's vision of an independent, data-driven, expert agency.

Other State Sentencing Commissions

Although the Minnesota and federal sentencing commissions have been among the longest continually functioning and most carefully studied American sentencing commissions, they have hardly been alone. More than 30 other states have had sentencing commissions at one time or another (Barkow and O'Neill 2006). In some states, such as South Carolina and Alaska, the commissions were designed as temporary bodies. In other states, such as Wisconsin and Florida, the commissions were supposed to be permanent but were later abolished. In addition to Minnesota, some of the other states with well-regarded, long-established, still-functioning sentencing commissions include Kansas, North Carolina, Pennsylvania, and Virginia.

As of this writing, the National Association of Sentencing Commissions lists 22 active state sentencing commissions on its website (including a commission in the District of Columbia), although at least two of these have much broader mandates than just sentencing and at least one has recently been legislatively abolished. It is possible that states will be encouraged to develop a new wave of sentencing commissions in coming years by the American Law Institute's Model Penal Code: Sentencing project. The MPC: Sentencing "recommends to all American jurisdictions that they establish a permanent sentencing commission . . . as an essential agency of the criminal-justice system" (American Law 2007, 47).

As the Minnesota and federal case studies illustrate, sentencing commissions are often formed in connection with the abolition of parole and initially tasked with developing mandatory sentencing guidelines; in such jurisdictions, superintending the guidelines then typically becomes a core, ongoing function of the commission. A number of states, however, have not followed the Minnesota-federal model in all of these respects. In Florida and Michigan, for instance, the judiciary initially developed sentencing guidelines; commissions were only created later (Little Hoover Commission 2007). In some states, including Alaska and Tennessee, guidelines were maintained after the commission expired or was abolished. Also, in some commission states, such as Wisconsin and Virginia, the guidelines were advisory (as they now are in the federal system). In still other states, such as Louisiana and Massachusetts, commissions have operated without any guidelines at all (American Law 2007). Nor does having a commission necessarily imply the elimination of discretionary parole; both institutions existed for a time in Delaware, Pennsylvania, Virginia, and Wisconsin (Frase 2005b).

Although they vary widely in size, composition, and budget, state commissions typically have a larger and more diverse membership than the federal commission (Frase 2005b). The North Carolina commission, for instance, has 30 members, and the Virginia commission 17.

In states in which they were abolished, commissions failed for a variety of reasons. Often, the

difficulties stemmed from either the resistance of judges to efforts to control their sentencing discretion or from the legislature's marginalization of the commission in making sentencing policy (Little Hoover Commission 2007).

Key Issues and Controversies

Have Sentencing Commissions Succeeded in Lessening Legislative Tendencies to Excessive Harshness?

Judge Frankel proposed the sentencing commission as a response to the problem of "sporadic and impassioned" legislative action. In this era of punitive populism, such legislative action has typically been on the harsh side, contributing to an increase in the American incarceration rate from 93 per 100,000 in 1972 to 500 per 100,000 in 2010 (Bureau of Justice Statistics 1982, 2012). Many critics contend that this ballooning incarceration rate reflects policies that "are too severe, waste lives and money, and often produce unjust results" (Tonry 2004, 3).

As Frankel's analysis suggested, a sentencing commission might bring about greater rationality and restraint in a state's criminal-justice system in at least three ways: by producing data to help policymakers better understand the costs and benefits of different policy options, by advocating on behalf of sound policies, and by creating sentencing guidelines that embody such policies. And there is some anecdotal evidence supporting the hope that commissions do sometimes succeed in reining in legislative excess (Barkow 2012; Wright 2002). Indeed, at least one multistate study finds a statistically significant, inverse relationship between the presence of a commission and growth in corrections spending (Barkow and O'Neill 2006).

Yet, it is plain that commissions are not always successful; the Minnesota and federal case studies both provide illustrations of commission marginalization. Having a commission does not ensure either a low or a stable incarceration rate. Minnesota's prison population was the nation's second-fastest growing in the first decade of the twenty-first century, while the

federal system was tied for sixth-fastest (Bureau of Justice Statistics 2012). Moreover, two sentencing commission states, Alabama and Louisiana, are among the top five by incarceration rate. Commission states with low incarceration rates, such as Minnesota (notwithstanding its recent growth spurt) and Massachusetts, tend to be places that had relatively low rates even before their commissions were created; it is hard to say whether and to what extent the commissions deserve credit, as opposed to preexisting aspects of the states' political and legal culture.

It is even possible that some commissions may be counterproductive. The federal commission, for instance, is criticized for sometimes amplifying, rather than muting, ill-advised legislative policy choices, as with the 1986 drug mandatory minimums (Barkow 2012). Moreover, as Bowman argues, a complex guidelines system, like that of the federal commission, may invite, rather than discourage, legislative intervention; the federal commission, he observes, "created a mechanism that permits endless legislative tinkering in response to the crime du jour" (Bowman 2005, 250). Indeed, even a commission's data-collection and reporting activities present risks; as Wright notes, "A system that monitors sentencing practices carefully also keeps criminal justice closer to the top of the public agenda. Growth [in the prison population] may be built into this system" (Wright 2002, 90).

In the end, we cannot know for certain which commissions, if any, have truly made a long-term difference in reining in legislative excess. We cannot replay history with and without a commission, or randomly assign states to a commission or no-commission condition. States with well-financed sentencing commissions are a self-selected lot; we should not be surprised to see that these states also tend to pay greater attention to their commissions and to display greater penal restraint than other states. Whether and to what extent commissions serve to reinforce preexisting, positive tendencies in these states – although theoretically plausible and anecdotally supported – seems beyond conclusive determination.

Why Do Legislatures Create Sentencing Commissions?

In a notable series of entries, Barkow has considered the sentencing commission, not through the lens of penal policy (as it is usually considered), but through the lens of administrative law theory. One interesting question that she has posed, along with her coauthor Barkow and O'Neill (2006), is why legislatures create sentencing commissions in the first place. Although Frankel supplied good reasons to think that a commission would be in the public interest, public-choice theories suggest that legislatures are not likely to enact laws purely on the basis of public interest. In the standard explanatory model of legislative delegation of authority to an administrative agency, the legislature does so in order to avoid making a difficult decision that will inevitably offend one powerful interest group or another; legislators can then claim credit for addressing a matter of public concern without actually risking interest-group support. Sentencing, however, does not seem to fit the model: no powerful interest group opposes longer sentences, so legislatures seem free to adopt tougher sentencing laws without fear of reprisal.

Barkow and O'Neill empirically tested a number of possible explanations for commission-creation. They found statistically significant relationships between a state's likelihood of having a commission and such independent variables as a narrow partisan margin in the state, a high incarceration rate, and a high rate of corrections expenditures. In these findings, they see support for a cost-centered account of sentencing commissions: legislators support commissions in states in which corrections spending is sufficiently high so as to threaten other legislative priorities (such as keeping taxes low) and in which a close partisan divide makes it especially difficult for politicians to exercise restraint in sentencing policy and thereby risk appearing "soft on crime." In such states, delegating sentencing policy to a commission may indeed help legislators to avoid making politically risky choices.

What Actors and Activities Should Sentencing Commissions Regulate?

If, like Barkow, we think about sentencing commissions in relation to conventional government regulatory agencies, we might ask what actors and activities are regulated by sentencing commissions (Barkow 2005). Based on the example of the Minnesota commission and others that have followed in its footsteps, the answer seems to be that sentencing commissions regulate judges and, particularly, the decisions that judges make about whether and for how long to incarcerate convicted felons. However, a sentencing commission's regulatory mission might be conceived much more broadly along both the "who" and the "what" dimensions. Indeed, if the fundamental objectives of a commission are understood to be something like achieving uniformity and proportionality in punishment or allocating criminal-justice resources in the most cost-effective fashion, then the narrowly focused approach of the Minnesota commission may be less than optimal or perhaps even counterproductive.

Consider, for instance, what may be the most perilous "third rail" for sentencing commissions: the regulation of prosecutorial plea bargaining. Plea bargaining is without question an important source of unwarranted disparities in the criminal-justice system. In jurisdictions with unregulated judicial sentencing discretion, judges could, in principle, offset some of these disparities. Depending on the overlap of statutory sentencing ranges, for example, a defendant who was able to negotiate an unusually generous reduction in charges might nonetheless ultimately receive the same sentence as other similarly situated defendants who were less fortunate in their negotiations. Sentencing guidelines, however, can impede such corrective measures, thus potentially resulting in even more prosecutor-created disparity at the same time that judicial disparity is being controlled. Indeed, the federal commission adopted "real-offense" sentencing in its guidelines precisely in order to limit the significance of prosecutors' charging and plea-bargaining decisions. Yet, federal prosecutors have proven remarkably adept at manipulating the guidelines

so as to achieve desired sentencing results (Bowman and Heise 2002), which may or may not be consistent with such ideals as uniformity, proportionality, and efficiency. Indeed, in its 15-year report, the federal commission itself identified prosecutor-created disparity as an important, unresolved issue in the federal criminal-justice system. In any event, if real-offense sentencing has proven disappointing as an indirect means of regulating plea bargaining – and whatever its benefits, real-offense sentencing has undoubtedly created large complexity costs for the federal system – then more direct means might seem in order, such as the adoption of plea-bargaining guidelines. But this would put a sentencing commission at odds with a politically powerful interest group, prosecutors. Commissions have generally avoided such confrontations, and even modest commission efforts to address plea bargaining, such as the Minnesota commission's initiative in the 1990s (Fraser 2005a, b) have drawn strong resistance.

A thorough examination of all of the other things that commissions *might*, and in a few cases actually *do*, regulate lies beyond the scope of this entry. A few examples will serve to illustrate the range of possibilities. The North Carolina sentencing commission, for instance, has addressed misdemeanors and non-prison sentences in its guidelines (Wright 2002). Washington created a special commission to oversee juvenile sentencing in 1977, and then later transferred its responsibilities to the state's adult sentencing commission (Boerner and Lieb 2001). Hoping to better coordinate sentencing and release decisions, Chanenson (2005) has proposed a "Super Commission" that would promulgate both sentencing and parole release guidelines. A commission might also seek to regulate what sanctions are imposed for violations of the terms of probation or supervised release, as the federal commission does. Indeed, the federal commission has promulgated guidelines concerning a wide range of subjects that many state commissions have not tackled, including sentencing and plea-acceptance procedures, fines, restitution, and organizational sentencing.

Why don't commissions regulate more broadly? In part, this reflects limitations on their statutory authority. However, commissions have not always taken full advantage of what authority they do have. The Minnesota commission, for instance, could have, but did not, regulate the conditions of non-prison sentences (Fraser 2005a, b). Moreover, even when statutory authority is lacking, commissions could request amendments to their enabling acts. However, as suggested by the discussion of plea bargaining above, another important constraint has doubtless been a desire by commissions and legislatures alike to avoid interagency turf battles; commissions have had enough political difficulty when their regulatory targets have been limited to sentencing judges that it is understandable why they might be less than enthusiastic about tackling prosecutors, juvenile courts, parole boards, probation offices, and so forth. And, even when it comes to regulating adult sentencing judges more fully (e.g., addressing misdemeanors and non-prison sentences), commission restraint likely reflects some combination of concerns regarding judicial resistance and resource limitations. Then, too, a commission focus on prison sentences is perfectly consistent with the Barkow-O'Neill hypothesis that commissions are formed primarily to prevent corrections costs from interfering with other legislative fiscal priorities; given the high cost of imprisonment relative to community corrections, preventing overutilization of prison seems the most direct and effective way to put a lid on state corrections spending.

How Can the Integrity of Commission Policy Choices Be Preserved From Legislative Incursions?

Although legislators may create sentencing commissions in the hope of relieving political pressures to adopt costly new sentencing laws, legislators do not seem able to keep their hands entirely off sentencing policy afterwards, even when the commission has created a comprehensive guidelines system. Legislative incursions threaten the coherence of state sentencing policy and may severely impair

a commission's ability to achieve such objectives as proportionality and uniformity in punishment and stability in the size of the state prison population. Although ongoing legislative engagement with sentencing policy seems unavoidable – and, in light of democratic values, is probably even desirable at some level – a pressing question for commission design and operation is how to minimize the likelihood that the legislature will simply ride roughshod over commission policy choices; ideally, legislative engagement should be informed by commission data and expertise and should avoid large-scale disruptions to commission-created sentencing systems in the absence of some justification stronger than public outrage over the crime du jour.

Although no state seems to have found a foolproof formula, a number of devices seem potentially helpful. First, commission membership may help to strengthen the commission's relationship with the legislature or otherwise to enhance its political standing. Thus, the American Law (2007) recommends that sentencing commissions actually contain sitting members of the state legislature. Harkening back to Judge Frankel's ideas about commission membership, the ALI's model legislation also contemplates a broadly inclusive commission with representation from a range of stakeholder groups; such an approach may enhance the commission's political legitimacy. By contrast, the relatively small and judge-heavy composition of the federal commission may be one source of its political weakness.

Second, Barkow (2012) observes that reliable cost-forecasting has played a particularly important role in raising the political standing of some state commissions with their legislatures and in helping these commissions to head off costly legislative proposals with low crime-reduction benefits. This experience underscores the need for commissions to develop strong data-collection and analysis capabilities, as well as the value of statutes requiring a fiscal impact analysis of sentencing bills as part of the regular legislative process. At the same time, Wright (2002) offers an important cautionary note:

although commissions may find their greatest success by playing the role of corrections accountant, this orientation may come at the expense of promoting and contributing to public debate on some of the deeper ethical questions raised by sentencing policy.

Finally, Bowman (2005) suggests that a commission should cultivate a good relationship with judges; they are, he observes, “natural allies” in checking legislative excess in the sentencing area. He notes that some of the political weakness of the federal commission may be attributed to the adversarial relationship that developed between the commission and the judiciary in the first few years of the guidelines era. This poor relationship may seem counterintuitive, given the prevalence of judges among the commission’s members. However, the judge-members were not selected by the judicial branch itself and were not necessarily reflective of mainstream judicial perspectives; the original commissioners, for instance, were tilted toward the appellate courts and had little actual sentencing experience (Stith and Cabranes 1998). By contrast, for instance, one version of the ALI proposal provides for selection of judge-members by the chief justice of the state supreme court and requires that three of five judge-members be trial-court judges (American Law Institute 2007). In any event, whether strengthened or weakened by the federal commission’s membership, many judges quickly developed various negative impressions of the commission: that it was overly concerned with controlling sentencing judges and not sufficiently concerned with achieving just outcomes, that its guidelines were too harsh and too rigid, and that it was oblivious to courtroom realities and disinterested in feedback from the judges “in the trenches.” This federal experience suggests that a commission wishing to maintain a good relationship with its “natural allies” should try to preserve a substantial measure of flexibility and discretion in the sentencing system, engage in regular dialogue with the trial bench, and demonstrate a willingness to modify guidelines based on judicial feedback.

How Can Sentencing Commissions Remain Relevant in an Advisory Guidelines System?

With Minnesota establishing the model, many commission jurisdictions gave their commissions the authority to promulgate and/or amend sentencing guidelines that had some level of legally binding force. In such jurisdictions, the commission has an obvious and important policymaking role (assuming the legislature does not routinely override commission policy choices). However, some commission jurisdictions chose to adopt advisory, instead of mandatory guidelines, which raises the question of whether a commission can be relevant when judges are free to ignore the commission’s guidance. This question has taken on greater urgency in the wake of *United States v. Booker* (2005), which shifted the federal guidelines from mandatory to advisory, and a related line of constitutional decisions that led at least four other states to make the same transition (Pfaff 2009).

Of course, regardless of its policymaking authority, a commission may have an impact through the other two roles envisioned by Frankel, those of think-tank and lobbyist. But there are good reasons to think that a commission may also affect sentencing practices through advisory guidelines. One statistical study, for instance, found that the adoption of advisory guidelines in some states led to reduced variation in the length of sentences; although the effects were not as strong as those associated with the adoption of binding guidelines, they were nonetheless substantial (Pfaff 2006). Likewise, the *Booker* decision did not immediately render the federal guidelines irrelevant; since the Supreme Court’s decision, most federal sentences have continued to be imposed within the guidelines range. Even when they are not required to follow it, some judges apparently (and understandably) appreciate and rely on expert guidance in discharging what many regard to be the most challenging responsibility in the judicial portfolio.

This is not to say, however, that advisory guidelines will inevitably succeed. Wisconsin and Louisiana, for instance, both tried and abandoned experiments with advisory guidelines

(Pfaff 2009). Moreover, since *Booker*, there have been persistent calls for Congress to reinstitute mandatory federal guidelines, prompted in part by steadily declining rates of within-guidelines sentences.

Procedural rules may enhance the relevance of advisory guidelines (and by extension of the commissions that superintend them). Following *Booker*, for instance, the Supreme Court made clear that guidelines ranges (though advisory) must still be calculated in each case; that determining the guidelines range should be the *first* step in the sentencing analysis; and that sentences within the range may be treated as presumptively reasonable on appeal. When the sentencing judge invests considerable effort in determining a guidelines range, it seems likely that the range will tend to condition the sentence imposed, even if the range is not regarded as controlling – all the more so if the range constitutes a safe harbor of sorts from the threat of appellate reversal.

Apart from such procedural rules, the commission in an advisory system may enhance the impact of its guidelines to the extent that they are persuasively grounded in sound cost-benefit research and reflect an appreciation of judicial perspectives. The commission might think of its “lobbyist” role as not merely limited to legislative policymaking decisions, but also extending to judicial sentencing decisions. This way of framing the commission’s role in an advisory system underscores the importance of the commission’s standing with the judiciary and the strength of the justifications it offers in support of its guidelines. In this spirit, a number of the suggestions made in the previous section regarding commission design and operation (e.g., strong representation by *trial* judges on the commission) may also be pertinent to the goal of maximizing compliance with advisory guidelines.

Future Directions

This concluding section briefly describes three important challenges that some commissions are beginning to face, and that may fundamentally

alter the role of sentencing commissions in the coming years.

First, although many sentencing reformers hoped that sentencing commissions and guidelines would lessen the effects of racial discrimination in the criminal-justice system, wide racial disparities in incarceration have persisted or even grown worse in some commission jurisdictions, including Minnesota and the federal system. It does not seem that commissions have engaged in purposeful discrimination, but they have sometimes made policy choices with important racially disparate effects, such as the federal commission’s original approach to crack sentencing and the Minnesota commission’s treatment of criminal history. In any event, whatever the source of disparities, they seem to present an increasingly urgent threat to the legitimacy of the criminal-justice system. Commissions that concern themselves broadly with the system’s fairness and effectiveness ought to pay heed. Indeed, in a few states, the legislature has prodded the commission to do so, for instance, by requiring the commission to prepare racial impact reports on sentencing bills along with a fiscal analysis (Barkow 2012). In addition to such participation in the legislative process, commissions that superintend guidelines might also do well to undertake a systematic examination of the racial impact of their own policy choices and to consider amendments so as to minimize unjustified disparities. The federal commission was a pioneer of sorts; its persistent efforts to address the 100:1 ratio were largely motivated by concerns regarding racial impact.

Second, commissions have traditionally communicated information to sentencing judges in a general way and at periodic intervals, in the form of reports, manuals, guidelines, training seminars, and so forth. In recent years, though, there has been growing interest in systems that would provide judges with real-time data at the bench or in chambers to assist with individual sentencing decisions as they arise. Such a “sentencing information system” would allow a judge to enter certain basic data about an offense and an offender, and then receive information about sentences and/or outcomes in

similar cases (Miller and Wright 2005). This might include, for instance, a precisely calculated recidivism risk for an offender, or information about what probation conditions have been imposed on similar offenders in other cases. In jurisdictions in which a sentencing commission is already collecting a great deal of case-level data, the commission seems the natural agency to develop a SIS. However, putting a database created for one purpose –informing policy-level decisions by the commission and the legislature – to new uses by new users presents significant technical challenges at a minimum. The challenges would be even greater if, as some hope, different jurisdictions linked their systems and pooled their data. But the challenges may not be merely technical. For instance, one category of information that sentencing judges might like to have is cost data: what is the estimated expense to taxpayers of various sentencing options? The Missouri sentencing commission has, in fact, already attempted to make such information available (Flanders 2012). However, some critics object that cost data should not be shared with judges; this may invite a cold-blooded, cost-benefit decision that loses sight of the needs of victims or the demands of justice. Another challenge is presented by the calculation of recidivism risk. Assuming that risk forecasts may be made more accurate when data like race, sex, ethnicity, family relationships, education, income, mental health, and so forth are taken into account, difficult questions arise concerning the fairness and legality of basing sentencing decisions, even indirectly, on personal characteristics over which a defendant has little or no control, particularly when there is a history of invidious discrimination associated with those characteristics.

Finally, there is the challenge of effecting fundamental reform to an established guidelines system. Minnesota introduced the basic, widely imitated template for sentencing guidelines with the development of its two-dimensional grid more than 30 years ago. Sentencing knowledge has progressed considerably in the interim, and both public and expert beliefs seem to have shifted on a number of fundamental questions,

such as the feasibility of achieving, and the desirability of attempting, offender rehabilitation in the criminal-justice system. Although commissions routinely amend their guidelines, these amendments are usually narrowly focused and rarely involve anything like a fundamental reworking of the guidelines architecture. Once a system is in place, the normal pattern seems to be that the commission becomes invested in the system's basic approach and shows little interest in reopening discussion on the big questions that had to be answered at the outset in creating the guidelines. This is all perfectly understandable, but it may become increasingly problematic as the basic models of the 1970s and 1980s fall further behind current knowledge and values. Attempts to address racial disparities and to develop sentencing information systems may exacerbate the perceived obsolescence of traditional approaches. Shifts from mandatory to advisory, and vice versa, may also necessitate structural changes in order to maximize guidelines effectiveness. Yet, fundamental change requires institutional will, political capital, and other resources that may be in short supply on many commissions. In some guidelines jurisdictions, the current systems emerged from multiyear deliberative processes and painstakingly crafted compromises among antagonistic stakeholder groups. It is hard to imagine a successful replication of such efforts. Yet, it is almost as hard to imagine that sentencing commissions and guidelines will remain viable institutions for another 30 years if they do not prove capable of periodic fundamental reform.

Related Entries

- ▶ [History of Probation and Parole in the United States](#)
- ▶ [Plea Bargaining](#)
- ▶ [Proportionality](#)
- ▶ [Risk Assessment, Classification, and Prediction](#)
- ▶ [Sentencing Guidelines in the United States](#)
- ▶ [Sentencing Research](#)

Recommended Reading and References

- American Law Institute (2007) Model penal code: sentencing, tentative draft no. 1. Executive Office, American Law Institute, Philadelphia
- Barkow R (2005) Administering crime. *UCLA Law Rev* 52:715–814
- Barkow R (2012) The sentencing guidelines at the crossroads of politics and expertise. *Univ Pa Law Rev* 160:1599–1630
- Barkow R, O’Neill K (2006) Delegating punitive power: the political economy of sentencing commission and guideline formation. *Texas Law Rev* 84:1973–2022
- Boerner D, Lieb R (2001) Sentencing reform in the other Washington. *Crime Justice* 28:71–136
- Bowman F (2004) Pour encourager les autres?: the curious history and distressing implications of the criminal provisions of the sarbanes-oxley act and the sentencing guidelines amendments that followed. *Ohio State Crim Law J* 1:373–442
- Bowman F (2005) Mr. Madison meets a time machine: the political science of federal sentencing reform. *Stanford Law Rev* 58:235–265
- Bowman F, Heise M (2002) Quiet rebellion II: an empirical analysis of declining federal drug sentences including data from the District level. *Iowa Law Rev* 87:477–561
- Bureau of Justice Statistics (1982) Prisoner 1925–1981
- Bureau of Justice Statistics (2012) Prisoners in 2010
- Chanenson S (2005) The next era of sentencing reform. *Emory Law J* 54:377–460
- Little Hoover Commission (2007) Solving California’s corrections crisis: time is running out. Little Hoover Commission, Sacramento
- Flanders C (2012) Cost as a sentencing factor: Missouri’s experiment. *Mo Law Rev* 77:391–410
- Frankel M (1973) Criminal sentences: law without order. Hill and Wang, New York
- Frase R (1997) Sentencing principles in theory and practice. *Crime Justice* 22:363–433
- Frase R (2005a) Sentencing guidelines in Minnesota, 1978–2003. *Crime Justice* 32:131–219
- Frase R (2005b) State sentencing guidelines: diversity, consensus, and unresolved policy issues. *Columbia Law Rev* 105:1190–1232
- Frase R (2009) What explains persistent racial disproportionality in Minnesota’s prison and jail populations? *Crime Justice* 38:201–280
- Friedman I, Supler K (2008) Child pornography sentencing: the road here and the road ahead. *Federal Sentencing Reporter* 21:83–89
- Knapp K (1987) The sentencing commission’s empirical research. In: von Hirsch A et al (eds) *The sentencing commission and its guidelines*. Northeastern University Press, Boston
- Miller M, Wright R (2005) The wisdom we have lost: sentencing information and its uses. *Stanford Law Rev* 58:361–380
- Minnesota Sentencing Guidelines Commission (2012) MSGC report to the legislature
- O’Hear M (2006) The original intent of uniformity in federal sentencing. *Univ Cincinnati Law Rev* 74:749–818
- Pfaff J (2006) The continued vitality of structured sentencing following blakely: the effectiveness of voluntary guidelines. *UCLA Law Rev* 54:235–307
- Pfaff J (2009) The future of appellate sentencing review: booker in the States. *Marquette Law Rev* 93:683–715
- Stith K, Cabranes J (1998) Fear of judging: sentencing guidelines in the federal courts. University of Chicago Press, Chicago
- Stuart J, Sykora R (2011) Minnesota’s failed experience with sentencing guidelines and the future of evidence-based sentencing. *William Mitchell Law Rev* 37:426–468
- Tonry M (2004) Thinking about crime: sense and sensibility in American penal culture. Oxford University Press, New York
- United States Sentencing Commission (2004) Fifteen years of guidelines sentencing: an assessment of how well the federal criminal justice system is achieving the goals of sentencing reform. United States Sentencing Commission, Washington, DC
- von Hirsch A (1987) Structure and rationale: Minnesota’s critical choices. In: Von Hirsch A et al (eds) *The sentencing commission and its guidelines*. Northeastern University Press, Boston
- Wright R (2002) Counting the cost of sentencing in North Carolina, 1980–2000. *Crime Justice* 29:39–106

Sentencing Guidelines in the United States

Steven L. Chanenson
School of Law, Villanova University,
Villanova, PA, USA

Overview

Sentencing guidelines have been a central part of criminal law reform efforts in the United States since the late 1970s. Through sentencing guidelines, many American jurisdictions have attempted – with varying degrees of success – to improve their criminal justice systems by combating judicial disparity; increasing fairness, honesty, and transparency; and, in some cases, controlling costs. While a majority of states do not have them, the presence of guidelines in the federal system and a substantial percentage of high-profile states have provided sentencing guidelines with considerable visibility.

Sentencing is where the rubber meets the road in the criminal law. All of the niceties of a criminal trial, in the statistically unlikely event there was one, are over. The defendant is guilty. The system must now choose how to respond and which actor or actors should have what degree of discretion. Sentencing guidelines are a popular American tool for helping to inform and implement those choices. Depending on the method of counting, there are at least 53 full criminal justice systems in the United States, one in each state plus the federal government, the District of Columbia, and the military. Each system must decide, within the constraints of the United States Constitution (and its own Constitution if applicable), how to punish convicted offenders. Sentencing guidelines, as befitting a tool used by different sovereigns reflecting different political and legal cultures, vary quite a bit from jurisdiction to jurisdiction. There are, however, some common questions that many drafters of sentencing guidelines – whether the judiciary, the legislature, or a sentencing commission – face.

This entry will explore some of the fundamental aspects of sentencing guidelines, their structure, breadth, and challenges. This entry will not examine sentencing guidelines or their analogues outside of the United States, nor will it focus more than necessary on the sentencing commissions that often promulgate sentencing guidelines as those bodies have their own entry in this series. It will also not address capital sentencing.

This entry will, however, explore selected attributes and challenges of sentencing guidelines through the lens of some of the key sentencing guideline jurisdictions.

Fundamental Background of and Structures for Sentencing Guidelines

How Should Sentencing Systems Be Described?

Terminology describing sentencing systems is frequently confusing. A common language is thus essential. Sentencing systems are either determinate or indeterminate. “Indeterminate

systems use discretionary parole release while determinate systems do not. Determinate and indeterminate sentencing schemes can take various forms. Either sort of system may be discretionary or nondiscretionary [as to sentencing]. Discretionary systems – be they determinate or indeterminate – may be guided or unguided” (Chanenson 2005, pp. 382–383). Sentencing guidelines are a popular way to provide guidance to sentencing judges, especially in discretionary, determinate sentencing systems. As noted below, sentencing guidelines can have varying degrees of force and may be best conceived of as occupying a continuum of enforceability. Common, although necessarily imprecise, terms for the relative power of the guidelines include “presumptive” and “advisory.”

Why Do Some Jurisdictions Choose to Have Sentencing Guidelines?

Before the 1970s, the dominant approach to sentencing in the United States reflected a rehabilitative model and placed broad discretion in judges and parole boards. Judge Marvin Frankel and others questioned and criticized this highly discretionary approach as lawless and unfair. Academic commentators highlighted a “gross disparity in sentencing, with different sentences imposed upon similar offenders who have committed similar offenses by the same judge on different days, different judges on different days, different judges on the same day, and different judges in different jurisdictions” (Singer 1978, p. 402). US Supreme Court Justice O’Connor noted that unguided sentencing discretion “inevitably resulted in severe disparities in sentences received and served by defendants committing the same offense and having similar criminal histories. Indeed, rather than reflect legally relevant criteria, these disparities too often were correlated with constitutionally suspect variables such as race” (*Blakely v. Washington*, 542 U.S. 296, 315 (2004) (O’Connor, J., dissenting)).

Starting in the 1970s, as a result of these criticisms and other, sometimes local, concerns, various American jurisdictions reformed their approach to sentencing by implementing

different forms of guidance. “One way to describe sentencing reform over the past half century is that law came to sentencing” (Miller 2004, p. 121). Some jurisdictions – most notably Minnesota and the federal government – abolished discretionary parole release and implemented a determinate system with sentencing guidelines. Other jurisdictions – most notably Pennsylvania – retained an indeterminate system but added sentencing guidelines. As the Pennsylvania Supreme Court has noted, sentencing guidelines “were promulgated in order to structure the trial court’s exercise of its sentencing power and to address disparate sentencing” (*Commonwealth v. Mouzon*, 812 A.2d 617, 620 n.2 (Pa. 2002) (plurality)). Federal sentencing reform was the product of years of discussion, debate, and political compromise. Some scholars have asserted that the final product, the Sentencing Reform Act of 1984, reflected a “subtle transformation of sentencing reform legislation: conceived by liberal reformers as an anti-imprisonment and antidiscrimination measure, but finally born as part of a more conservative law-and-order crime control measure” (Stith and Koh 1993, p. 223).

Sentencing guidelines reflect the legal and policy choices of the jurisdiction. Thus, popular attitudes and opinion can play a significant role. Sentencing guidelines can spark intense debate in part because they are transparent. This law-based transparency “forces the resolution of issues of sentencing policy – and enforces the particular resolution – and thus makes the fact of resolution clear. This society is not of one mind on what values sentencing should embody and thus the resolution of these disputed issues means the values of some will prevail while those of others will not” (Boerner 1993, p. 174).

What Kinds of Structural Choices Do Legislatures Make When Starting This Process?

There are different ways to structure a sentencing system. Each jurisdiction must confront at least three common structural choices. First, should sentencing be determinate or indeterminate? Many commissions and guidelines – including

in the Minnesota and federal systems – emerged as part of a package of legislation that abolishes discretionary parole release, which transforms the system from an indeterminate one to a determinate one. Other guidelines – like Virginia – began within an indeterminate structure, but the legislature transitioned to a determinate approach later for reasons similar to those motivating the Minnesota and federal systems. Finally, there are a couple of jurisdictions – most notably Pennsylvania – that continue to have both an indeterminate structure (meaning discretionary parole release) and well-developed sentencing guidelines.

Second, which actor in the system should promulgate the guidelines? In part because of their popularity and specialized expertise, permanent sentencing commissions have captured the lay imagination concerning guideline creation and maintenance, but they are not the only approach. It is possible for the legislature or the judiciary to create guidelines either themselves or through temporary bodies established for that purpose. Even when it creates a sentencing commission, the legislature often reserves for itself the power to formally create the guidelines or at least a veto power over the guidelines promulgated by the commission. A popular approach, exemplified by the federal and Pennsylvania systems, involves the sentencing commission issuing guidelines which take effect after a predetermined amount of time unless the legislature and executive affirmatively block them. This power-sharing arrangement is often described in a positive light as allowing the commission to act on its expertise with less political pressure than elected officials might feel while still respecting principles of democratic accountability.

Third, how much binding force should the guidelines have? Not all guidelines have the same amount of binding force. As exemplified by the original Minnesota and the federal approaches, powerful guidelines were initially a popular approach. Many felt that anything less binding would not rein in judicial discretion. Before a new interpretation of the Sixth Amendment to the United States Constitution emerged in the early 2000s, the federal guidelines were

very binding (often called “presumptive”) and judges would frequently have difficulty successfully justifying sentences that diverged from the presumptive guideline range, a process called “departing.” In fact, the rigid nature of the federal guidelines – and the associated relative lack of judicial discretion – attracted severe and persistent criticism.

Terminology here can be challenging as labels like “presumptive” and “advisory” can mask many variations. Indeed, “there are an infinite number of stops between a purely advisory approach and a completely mandatory framework” (Reitz 2005, p. 157). However, a jurisdiction can influence how much power the guidelines (and, at times, functionally the prosecutor) have as opposed to the judge. One indication of a more voluntary guideline regime, exemplified by Virginia, is the absence of appellate review. The presence of appellate review generally reflects more binding guidelines, but the intensity of that review can vary widely. Releasing judge-specific sentencing information may also have an impact on guideline compliance as well as reflect a commitment to transparency (Bergstrom and Mistick 2003). It is interesting to note that Pennsylvania, a state in which judges run for election and stand for retention, releases judge-specific information while the federal government, in which judges are appointed for life, does not.

More binding or presumptive guidelines are arguably desirable because they have sufficient teeth to meaningfully limit judicial discretion. However, more advisory, and thus less binding, guidelines are arguably desirable because they provide sufficient judicial flexibility to promote fairness in a particular case. Just as there can be unwarranted disparity, there can also be unwarranted uniformity. Professor Doug Berman observed that “reformers believed that sentencing guidelines, by codifying standards which would direct judges’ sentencing decisions in most but not all cases, could reduce sentencing disparities and maintain sentencing flexibility, while promoting the development of principled sentencing law and policy” (Berman 2000, p. 35). Ultimately, each jurisdiction has to find its

own balance between individualization and uniformity of sentences.

How Do Sentencing Commissions Create Sentencing Guidelines?

Guideline drafters have a myriad of choices to make, and the American constitutional structure affords them substantial latitude within which to work. There are at least six decisions that can be characterized as critical:

1. What Philosophical Goals Should Animate the Guidelines?

The legislature and/or the sentencing commission should address what philosophical goals of sentencing should be reflected in the guidelines. The standard debates about purposes of punishment play out in this process. While the rehabilitative model had been dominant, people like criminologist and law school Dean Norval Morris successfully urged that punishments should be limited by principles of just deserts. Morris argued that desert is “an essential link between crime and punishment. Punishment in excess of what is seen by that society at that time as a deserved punishment is tyranny” (Morris 1974, p. 76). Most systems have embraced Morris’ idea of “limiting retributivism,” pursuant to which the offender’s desert defines the upper and lower limits of acceptable (not unjust) punishment and within that range other utilitarian goals can be pursued (Frase 2005, pp. 76–77).

Although most modern American sentencing systems are now largely retributive, guideline drafters often struggle with how to articulate their views and integrate competing ideas. The United States Sentencing Commission officially sidestepped the question at the macro level when it promulgated the first set of federal guidelines in 1987. It noted that, “[a] philosophical problem arose when the Commission attempted to reconcile the differing perceptions of the purposes of criminal punishment.... Adherents [to both “just deserts” and “crime control” approaches] urged the Commission to choose between them, to accord one primacy over the other....

A clear-cut Commission decision in favor of one of these approaches would diminish the chance that the guidelines would find the widespread acceptance they need for effective implementation. As a practical matter, in most sentencing decisions both philosophies may prove consistent with the same result” (U.S. Sentencing Commission 1987, Ch 1A 1.3–1.4). Nonetheless, some scholars of federal sentencing have argued that the federal guidelines reflect a “modified just desert” theory in which “the greatest weight in determining sentences is given to matching the severity of punishment to the seriousness of the present offense,” and the second most significant consideration is “the need to incapacitate for longer periods the more dangerous offenders” determined by criminal history (Hofer and Allenbaugh 2003, p. 24). This seems to be consistent with, if not a variant of, “limiting retributivism.”

Putting these pieces together is not easy. Professor Frase summed it up by noting that “early guidelines reforms attempted to narrow the focus of sentencing to strongly emphasize uniformity and ‘just deserts,’ and to promote more rational sentencing policy. The much broader range of contemporary sentencing goals demonstrates an important underlying truth, which early guidelines reforms (and some recent proposals) seem to have overlooked: sentencing policy is very complex, requiring compromise and careful balancing of numerous, often-competing goals” (Frase 2000, pp. 435–436).

2. Should the Guidelines Be Descriptive or Prescriptive?

Sentencing guidelines can be designed primarily to reflect past judicial practices which can be described as a “descriptive” approach. This has the benefit of more likely acceptance by the judiciary as it will seem familiar in most cases, but it may not always produce the most desirable results. When the Virginia guidelines were promulgated, they enjoyed strong judicial support in part because they were largely descriptive. On the other hand, sentencing guidelines can be

designed from the ground up to reflect what the sentencing commission believes is the best path even if it diverges from past judicial practice. This can be described as a “prescriptive” approach. Minnesota used a more prescriptive strategy in part to encourage incarceration for violent offenders and discourage it for property offenders. The prescriptive path may have the benefit of a more rational approach, but could encounter opposition from lawyers and judges who are accustomed to the old ways. As with so many things, the common responses can be plotted along a continuum with the bulk of guidelines avoiding the extremes.

3. Should the Guidelines Be Based More on a “Real” or a “Charge” Offense Model?

Which crime(s) should matter when sentencing a defendant? At first blush, it may seem like an odd question to ask. Should the defendant only be sentenced on the crime(s) for which he has been convicted? Does that mean that the judge must ignore compelling evidence of other criminal acts that, for whatever reason, were not the subject of the jury’s guilty verdict or the defendant’s guilty plea? On the other hand, should the defendant be sentenced based on every bad thing that the government alleges – and supports with proof at less than the beyond a reasonable doubt level – the defendant has ever done? Does that mean that the judge must factor in the amount of cocaine allegedly involved in a second drug charge on which the jury returned a verdict of not guilty? The restrictive approach is, at the extreme, described as “charge offense” sentencing while the inclusive approach is, at the extreme, described as “real offense” sentencing.

Charge offense sentencing runs the risk of the judge being prevented from sentencing based on everything known about the offender and the offense. It is also criticized for perhaps making it even easier for prosecutors to control the system by bringing more charges against some defendants and fewer charges against others. Real offense sentencing runs the risk of the defendant being punished on the basis of weak allegations that a jury

either rejected or was never even given the chance to consider. While the extremes of both approaches may strike many people as preposterous, jurisdictions do stake out positions leaning toward one side or the other. Pennsylvania, for example, is largely a charge-based system, although judges may consider information about other behavior in exercising sentencing discretion. The federal system, in contrast, enshrines a fairly aggressive real offense scheme in its guidelines. Judges must calculate the guidelines on the basis of “relevant conduct,” which encompasses acts similar to the offense of conviction. This relevant conduct must be included as long as it is proven by a preponderance of the evidence even if the government never charged those acts or the jury, employing the beyond a reasonable doubt standard, acquitted on them.

4. How, If at All, Should the Guidelines Address Multiple Convictions?

Guideline systems struggle to deal with multiple convictions. Many people would agree that multiple offenses usually deserve greater punishments than a single offense of the same nature. But how much more? Should the increase be geometric? Is a “volume discount” appropriate for the multiple burglar? Some guidelines require consecutive sentences for certain types of offenses, such as violent crimes, but mandate concurrent sentences for other offenses, such as property crimes. Other jurisdictions, like Pennsylvania, do not speak to multiple convictions at all and thus afford the judge nearly unfettered discretion to impose sentences concurrently or consecutively. Then-Judge (now US Supreme Court Justice) Breyer criticized that discretionary model by observing that “[a] moment’s thought suggests, however, that this approach leaves the prosecutor and the judge free to construct almost any sentence whatsoever” (Breyer 1988, p. 26). The federal system tries to avoid those extremes by creating a “system that treats additional counts as warranting additional punishment but in progressively diminishing amounts” (Breyer

1988, p. 27). While this federal approach may be more satisfying in theory, the practical result is a fairly byzantine process with nearly algebraic calculations that itself has been the subject of significant criticism.

5. How Should the Guidelines Deal with Mandatory Minimum Sentences?

Despite the existence of sentencing guidelines which are designed to take a nuanced approach, many legislatures continue to create mandatory minimum provisions. Mandatory minimum sentence laws are often criticized as being crude, inconsistently applied, and in tension with the idea of sentencing guidelines. US Supreme Court Justice Breyer, who was an influential member of the original US Sentencing Commission, noted that mandatory minimum sentences “prevent the Commission from carrying out its basic, congressionally mandated task: the development, in part through research, of a rational, coherent set of punishments.... Most seriously, they skew the entire set of criminal punishments, for Congress rarely considers more than the criminal behavior directly at issue when it writes these provisions...” (Breyer 1999, p. 184). Many sentencing commissions have to decide how to respond to such conflicting legislative directives. For example, Congress created both the federal sentencing guidelines system and a series of gun and drug mandatory sentences in close succession. In fact, the mandatory sentences were passed before the new commission could even promulgate the first set of guidelines.

There are two broad responses available to a sentencing commission. First, it could ignore the mandatory minimum sentences when crafting its guidelines. This allows the guidelines to reflect the views of the sentencing commission and simply be trumped by the mandatory minimum sentences when they apply. Doing so has the potential of creating a sentencing “cliff,” which is a point where a small change in offense behavior prompts a large change in the sentence. For example, the commission may set relatively modest weight-based guidelines for

a drug offense. If the amount of the drug involved reaches the level that triggers the mandatory minimum sentence, however, the defendant's sentence may increase precipitously. Sentencing cliffs are frequently criticized because of the perceived unfairness of the sharp increase in sentencing severity in response to the incremental increase in offending severity. Alternatively, the commission could accommodate the mandatory minimum scheme by integrating it into its guideline recommendations. This approach, largely followed in the federal system, avoids the "cliff" problem but requires the guidelines to be driven by the mandatory minimum sentences in a way that may diverge from what the sentencing commission would have otherwise done. The result may be greater severity for a wider swath of offenses and offenders. The Pennsylvania guidelines, providing another example of how broad labels can conceal variations, have largely declined to accommodate mandatory minimum sentences for its core recommendations. However, the guidelines have engaged with and attempted to provide alternatives to mandatory minimum sentences by providing such things as advisory sentence enhancements for conduct that might otherwise trigger mandatory minimums. It is interesting to note that Pennsylvania adopted a guideline system in part to "avoid mandatory minimum legislation that would severely restrict judicial discretion" (Kramer and Ulmer 2009, p. 15).

6. How Should the Sentencing Recommendations Be Communicated?

The most common way to reflect the jurisdiction's sentencing choices and recommendations is through a simple two-axis grid, although there are other approaches such as narratives. This two-axis grid commonly reflects the defendant's criminal history along the horizontal axis and the severity of the crime along the vertical axis. Where those two paths intersect is the box or cell containing the guidelines' recommended sentence or sentencing range, which is often expressed in months of incarceration (*see* Fig. 1).

Some jurisdictions, including Minnesota and Pennsylvania, have different grids for certain crimes (like sex offenses) or different ways in which crimes are committed (like with the possession or use of a deadly weapon). These separate grids are often explained as a way to reflect the particular seriousness of the targeted behavior without skewing the recommended sentences for the majority of offenses. This approach has been praised for limiting the impact of popular pushes for severity to the target offenses/offenders and has also been criticized for allowing that push for severity concerning those offenses/offenders to proceed more easily by bifurcating it from the bulk of cases.

The Mechanics of the Federal Sentencing Guidelines

Sentencing guidelines require the judge and lawyers to engage in a series of calculations that yield a particular recommendation, often as represented by a box or cell on the guideline grid. Guideline calculations vary in complexity depending on the jurisdiction. The federal system is widely viewed as the most complicated and is often criticized on that basis. Despite its distinctive federal complexities, a review of the kinds of calculations a federal judge needs to make before imposing a sentencing may afford a useful window on guideline sentencing generally.

There are 43 levels of what the federal guidelines call "offense seriousness" (US Sentencing Commission). In order to determine the ultimate offense level, the judge starts with the "base offense level" for the offense of conviction. Logically, offenses the Commission deems to be more serious are assigned a higher base offense level. The judge must then consider "specific offense characteristics," which are factors related to the offense such as the amount of money taken in a robbery or the quantity of drugs trafficked. The more the judge finds as taken or trafficked, the greater the increase in the offense level. At this point, the federal system's modified real offense feature can sweep in conduct for which there was no conviction. Next, the judge must determine whether

SENTENCING TABLE
(in months of imprisonment)

Offense Level	Criminal History Category (Criminal History Points)					
	I (0 or 1)	II (2 or 3)	III (4, 5, 6)	IV (7, 8, 9)	V (10, 11, 12)	VI (13 or more)
1	0-6	0-6	0-6	0-6	0-6	0-6
2	0-6	0-6	0-6	0-6	0-6	1-7
3	0-6	0-6	0-6	0-6	2-8	3-9
4	0-6	0-6	0-6	2-8	4-10	6-12
5	0-6	0-6	1-7	4-10	6-12	9-15
6	0-6	1-7	2-8	6-12	9-15	12-18
7	0-6	2-8	4-10	8-14	12-18	15-21
8	0-6	4-10	6-12	10-16	15-21	18-24
9	4-10	6-12	8-14	12-18	18-24	21-27
10	6-12	8-14	10-16	15-21	21-27	24-30
11	8-14	10-16	12-18	18-24	24-30	27-33
12	10-16	12-18	15-21	21-27	27-33	30-37
13	12-18	15-21	18-24	24-30	30-37	33-41
14	15-21	18-24	21-27	27-33	33-41	37-46
15	18-24	21-27	24-30	30-37	37-46	41-51
16	21-27	24-30	27-33	33-41	41-51	46-57
17	24-30	27-33	30-37	37-46	46-57	51-63
18	27-33	30-37	33-41	41-51	51-63	57-71
19	30-37	33-41	37-46	46-57	57-71	63-78
20	33-41	37-46	41-51	51-63	63-78	70-87
21	37-46	41-51	46-57	57-71	70-87	77-96
22	41-51	46-57	51-63	63-78	77-96	84-105
23	46-57	51-63	57-71	70-87	84-105	92-115
24	51-63	57-71	63-78	77-96	92-115	100-125
25	57-71	63-78	70-87	84-105	100-125	110-137
26	63-78	70-87	78-97	92-115	110-137	120-150
27	70-87	78-97	87-108	100-125	120-150	130-162
28	78-97	87-108	97-121	110-137	130-162	140-175
29	87-108	97-121	108-135	121-151	140-175	151-188
30	97-121	108-135	121-151	135-168	151-188	168-210
31	108-135	121-151	135-168	151-188	168-210	188-235
32	121-151	135-168	151-188	168-210	188-235	210-262
33	135-168	151-188	168-210	188-235	210-262	235-293
34	151-188	168-210	188-235	210-262	235-293	262-327
35	168-210	188-235	210-262	235-293	262-327	292-365
36	188-235	210-262	235-293	262-327	292-365	324-405
37	210-262	235-293	262-327	292-365	324-405	360-life
38	235-293	262-327	292-365	324-405	360-life	360-life
39	262-327	292-365	324-405	360-life	360-life	360-life
40	292-365	324-405	360-life	360-life	360-life	360-life
41	324-405	360-life	360-life	360-life	360-life	360-life
42	360-life	360-life	360-life	360-life	360-life	360-life
43	life	life	life	life	life	life

Sentencing Guidelines in the United States, Fig. 1 Sentencing table, United States sentencing guidelines

any “adjustments” apply. Adjustments are not offense-specific and include such matters as the defendant’s role in the offense (as a major or minor participant) and whether the defendant obstructed justice. Adjustments often increase a defendant’s offense level, but they can reduce it as well, as in the case of a minor participant. The judge then must consider whether there are multiple counts of conviction. As noted above, the federal system employs a complicated formula to give an incremental increase for multiple offenses. Next, the judge evaluates whether the defendant has “accepted responsibility” for his offense. This often correlates with a guilty plea. If the defendant has accepted responsibility, his offense level will be reduced by either two or three offense levels. The resulting number is the defendant’s offense level, which is represented on the vertical axis in the Sentencing Table (*see* Fig. 1). The horizontal axis reflects the defendant’s criminal history. The judge assigns criminal history points for previous convictions of various types and whether the defendant was under judicial supervision at the time of the offense. These points translate into a “Criminal History Category.” There are six Criminal History Categories. Where the offense level and the Criminal History Category intersect is the defendant’s guideline range.

The US Sentencing Commission recaps the process this way. “The final offense level is determined by taking the base offense level and then adding or subtracting from it any specific offense characteristics and adjustments that apply. The point at which the final offense level and the criminal history category intersect on the Commission’s sentencing table determines the defendant’s sentencing guideline range” (US Sentencing Commission, p. 3). An offense level of 19 and a Criminal History Category of I, which is where a first-offender would be classified, yields a guideline range of 30–37 months. If the defendant was in Criminal History Category VI, which is the highest possibility, the guideline range would be 63–78 months.

Despite having calculated the guideline range, the judge is not ready to impose a sentence. Next,

the judge must evaluate whether this is an unusual case warranting a different sentence under the guidelines. “[I]f an atypical aggravating or mitigating circumstance exists, the court may ‘depart’ from the guideline range. That is, the judge may sentence the offender above or below the range” (US Sentencing Commission, p. 3). For years, the availability of departures under the federal guidelines was a topic of frequent litigation and vigorous debate. Many critics argued that the appellate courts and the federal sentencing commission took an inappropriately crabbed view of these kinds of departures and that the fairness of the sentences often suffered. Yet judges did depart at times and could do so either up or down the offense level axis or left or right along the Criminal History Category axis. Since the Supreme Court of the United States made the federal guidelines “effectively advisory” (*see* below), judges must also consider whether the guidelines are appropriate in this case in light of the authorizing statute which invokes the traditional purposes of punishment as well as a parsimony provision. The judge may impose a completely different sentence if the judge believes that the guideline range does not meet the dictates of the statute, although the sentence is subject to appellate review for “unreasonableness.”

The Sixth Amendment’s Shifting Sands

Until approximately 2004, several guideline systems – especially the federal system – relied on facts that judges found at sentencing by a preponderance of the evidence in order to determine the presumptive sentencing range. “The top of the presumptive range was below the traditional statutory maximum for the offense of conviction. The actual sentence imposed might be higher or lower than the presumptive range, in part because of judicially found aggravating or mitigating facts” (Chanenson 2005, p. 378). In a series of decisions (including *Blakely v. Washington* and *United States v. Booker*) interpreting the Sixth Amendment to the United States Constitution, which guarantees the right to a trial by jury, the Supreme Court of the United States functionally invalidated crucial aspects of

these presumptive guideline systems. The Court held these schemes unconstitutional because judges were allowed “to impose sentences higher than the presumptive guideline range based on facts found by the judge, using the preponderance of the evidence standard, instead of by the jury, using the beyond a reasonable doubt standard” (Chanenson 2005, p. 378). The key, according to the Court, was that a judge could not impose a sentence that exceeds the statutory maximum which, for these purposes, it defined as the “maximum sentence a judge may impose *solely on the basis of the facts reflected in the jury verdict or admitted by the defendant*” (*Blakely v. Washington*, 542 U.S. at 303). The Court emphasized that “the relevant ‘statutory maximum’ is not the maximum sentence a judge may impose after finding additional facts, but the maximum he may impose *without* any additional findings” (*Blakely v. Washington*, 542 U.S. at 303).

This could have been a mortal blow to federal sentencing guidelines, which was a heavily presumptive system and relied on significant judicial fact finding. However, the Supreme Court provided an escape hatch. These presumptive systems had two primary options. First, juries could be asked to find all of the facts beyond a reasonable doubt that judges had been finding by a preponderance of the evidence. This would then authorize the judge to impose a sentence in the same guideline range as in the past. While some jurisdictions went in that direction, many did not in part because of the factual and logistical challenges of asking juries to decide such sentencing facts as a defendant’s role in the offense. This expanded role for the jury would have been particularly difficult for the federal sentencing system to adopt because it was both highly presumptive and required judges to make numerous factual determinations at sentencing, which now juries would have to make. The second option involved recasting the formerly presumptive guidelines as advisory. If the guidelines are just advisory, the judge has the ability to ignore the guidelines’ advice and the power to impose a sentence up to and including the traditional statutory maximum, which is

almost always higher than what the guidelines recommend. This is the path the federal system followed. However, neither the US Sentencing Commission nor the Congress made that change. Rather, it was the Supreme Court itself that recast the federal guidelines in this way. The Court construed the federal guidelines as “effectively advisory” in order to avoid the need to strike down the system as unconstitutional (*United States v. Booker*, 543 U.S. 220 (2005)). In doing so, it also broadened many people’s conception of what it means for a system to be “advisory,” as federal judges must still consider the guidelines and the sentences they impose are still subject to appellate review, albeit under a less stringent standard of review.

Broader Questions and Emerging Practices

Intermediate Punishments and Misdemeanors

Intermediate punishments have been the subject of discussion – and practice – for many years. Norval Morris and Michael Tonry have argued:

We are both too lenient and too severe; too lenient with many on probation who should be subject to tighter controls in the community, and too severe with many in prison and jail who would present no serious threat to community safety if they were under control in the community. (Morris and Tonry 1990, p. 3)

It is thus unfortunate and somewhat surprising given the size and scope of community corrections programs that guidelines relating to intermediate punishment are not common. Some jurisdictions – like Pennsylvania and North Carolina – have addressed this issue to various degrees, but it remains an underdeveloped area (Frase 2000, pp. 439–442; Frase 1999, pp. 77–78).

It is also interesting to note that relatively few guideline systems apply to both misdemeanors and felonies (Frase 2000, p. 429).

Guideline “Effectiveness”

The “effectiveness” of sentencing guidelines is often a topic of intense discussion and

disagreement. Part of the disagreement flows from differing definitions of success and varying methodological views on examining the relevant issues. Overall, many believe that sentencing guidelines can be and have been effective along particular metrics, including disparity reduction, in certain circumstances (U.S. Sentencing Commission 2004, xiv–xvi, 135, 141; Frase 2000, p. 443). For example, Professor Michael Tonry has written that “[g]uidelines promulgated by commissions have altered sentencing patterns and practices, reduced sentencing disparities and gender and race effects, and shown that sentencing policies can be linked to correctional and other resources, thereby enhancing governmental accountability and protecting the public purse” (Tonry 1993, p. 713). To be sure, not every jurisdiction succeeds on every metric all the time, but if one accepts that an increase in transparency and a decrease in judicial disparity were key goals of sentencing guidelines, there are reasons to be encouraged.

Sentencing guidelines have also had some success in the realm of controlling costs. At a minimum, many sentencing commissions have become trusted sources of information for legislatures as they debate broad penal strategies. Some commissions, particularly Minnesota and North Carolina, have tried to respect financial restraints as they promulgate guidelines. More recently, Missouri started providing information about the financial consequences of particular sentencing decisions to judges in individual cases.

However, if one is focused more on controlling prosecutorial discretion (and the disparity that may flow from that discretion) sentencing guidelines have not been effective. As some predicted early on, there is a widely held view that prosecutorial power has increased after the introduction of sentencing guidelines. “Since guidelines limit the range of sentences available for a given offense, the power to drop or not drop charges is the power to select the sentence range available to the court (that is, what ‘box’ on the grid the case ends up in). Thus, any disparity in charging translates into disparity in sentencing” (Frase 1999, p. 77). It must be recognized that most sentencing guidelines do not speak to the prosecutorial role directly

and were never intended to do so, despite the wishes of some.

Similarly, if one is focused on controlling prison populations, sentencing guidelines have not been particularly effective. The incarceration rate in the United States has risen precipitously in the last 30 years. Some jurisdictions have succeeded in curbing the growth of their prison population, but it is more than debatable whether the existence or nonexistence of guidelines is the dispositive feature. Sentencing guidelines are tools to implement broad policy in a manner that promotes systemic rationality and individual fairness. To that extent, there is reason to believe they can be – and often are – “effective.” Sentencing guidelines are not inherently severe or lenient any more than fire is inherently a force for good (warmth and cooking) or evil (arson and destruction). “The experience in Minnesota and Washington, progressive states generally regarded as having the most successful guidelines systems, teach that while those systems were effective at restraining the growth of prison populations when that was the policy of those states, they were equally effective at implementing policy judgments that more punitive sentences were appropriate” (Boerner 1993, p. 176).

Evidence-Based Practices and Risk

Selected jurisdictions have started to consider how, if at all, they should incorporate evidence-based practices, including risk assessment, in their sentencing systems. While there are some echoes from the pre-guidelines era’s interest in rehabilitation, the emphasis today is quite different. The focus now is on how reliable social science evidence can inform sentencing determinations – at the level of the guidelines and/or the individual sentencing judge – in a way that will reduce recidivism and thus improve public safety. Former Missouri Chief Justice Mike Wolff, who facilitated the inclusion of risk assessment information in Missouri’s Sentencing Assessment Reports (its version of presentence investigation reports), has argued that, “We must acknowledge that the reason for sentencing is to punish, but if we choose the wrong punishments, we make the crime problem worse,

punishing ourselves as well as those who offend” (Wolff 2008, p. 1395).

Virginia is the undisputed pioneer in integrating risk assessment into sentencing guidelines. Pursuant to a 1994 legislative directive, the Virginia Criminal Sentencing Commission explored whether an empirically based risk assessment tool could help judges divert “25 % of the lowest risk, incarceration-bound, drug and property offenders for placement in alternative (non-prison) sanctions” (Kern and Farrar-Owens 2004, p. 165). At the legislature’s request, the Virginia commission later broadened the eligibility to additional low-risk offenders. If the risk instrument indicates that the qualifying offender is of sufficiently low risk, the Virginia guidelines, which are advisory and without appellate review, recommend the offender for an alternative sanction. The judge remains free to accept or reject that recommendation. A few years later, the Virginia legislature asked the commission to develop a risk assessment tool for sex offenders with the goal of identifying those with the highest risk of reoffending. In 2001, this instrument was integrated into the Virginia guidelines by increasing the upper end of the recommended range by varying amounts for higher risk offenders (Kern and Farrar-Owens 2004, p. 167). Again, the judge retains discretion and may, but need not, impose a sentence that takes advantage of the expanded sentencing range while remaining in compliance with the guidelines.

Pennsylvania is also weaving risk assessment into its sentencing guidelines. Under Pennsylvania’s guidelines, the duration and intensity the recommended non-incarcerative sentence for certain drug-involved offenders has long reflected the treating of professional’s judgment concerning treatment needs. Pursuant to a 2010 legislative directive, the Pennsylvania commission is now working to integrate an actuarial risk assessment instrument into the guidelines. While Pennsylvania may not follow the precise path blazed by Virginia, it is also trying to use risk assessment “to assist in the transparent – and thus accountable – decision-making by both the Commission at the policy level and judges... at the individual level” (Hyatt et al. 2011, p. 748).

There are many reasons to be cautious about risk assessment, and many policy makers and judges are quite wary, if not suspicious, of this approach. By definition, risk assessment deals with predictions and not certainties. No actuarial instrument will ever be 100 % accurate. Yet, policy makers and judges have long engaged in predictions of risk based on their clinical judgment – or gut instinct – which itself is not only imperfect, but research indicates is often less accurate than actuarial instruments (Gottfredson and Moriarty 2006). Sentencing remains deeply normative and predictions about recidivism will never – and should never – be the only consideration. Actuarial risk assessment is simply a tool that guideline systems can use to help inform judicial discretion.

Not only does risk assessment raise crucial questions concerning how to distribute punishment, but it also can prompt reflection on the severity of punishments. If a less severe sentence does not diminish – and may even improve – public safety in certain situations, was society’s initial punitive judgment sound? There are no immutable answers, but the questions themselves can be important.

Concluding Observations

Sentencing guidelines have helped to refashion the landscape of American criminal justice, and they are continuing to do so. Even jurisdictions that do not yet have guidelines may be attracted by their promise of a more rational, transparent, and just approach to punishment. Indeed, the American Law Institute seems likely to recommend that all American jurisdictions adopt guidelines. Sentencing guidelines, however, remain very much a work in progress. All guidelines systems have flaws and many of them are glaring, but those same flawed systems often provide a mechanism for experimentation and, one hopes, positive refinement (Frase 2000, p. 445). Professor Michael Tonry summed it up this way:

Like all calls for just the right amount of anything, not too much and not too little, a proposal for sentencing standards that are constraining enough

to assure that like cases are treated alike and flexible enough to assure that different cases are treated differently is a counsel of unattainable perfection. Nonetheless, that is probably what most people would want to see in a just system of sentencing... (Tonry 1996, pp. 185–186).

Sentencing guidelines offer a real opportunity to strive for that “unattainable perfection.”

Related Entries

- ▶ [Penal Philosophy and Sentencing Theory](#)
- ▶ [Sentencing Commissions](#)
- ▶ [Sentencing Research](#)

Recommended Reading and References

- Allen F (1981) *The decline of the rehabilitative ideal: penal policy and social purpose*. Yale University Press, New Haven
- Bergstrom M, Mistick J (2003) The Pennsylvania experience: the public release of judge-specific sentencing data. *Fed Sentencing Rep* 16:57–61
- Berman D (2000) Balanced and purposeful departures: fixing a jurisprudence that undermines the federal sentencing guidelines. *Notre Dame Law Rev* 76:21
- Boerner D (1993) Bringing law to sentencing. *Fed Sentencing Rep* 6:174–177
- Breyer S (1988) The federal sentencing guidelines and the key compromises upon which they rest. *Hofstra Law Rev* 17:1–50
- Breyer S (1999) Federal sentencing guidelines revisited. *Fed Sentencing Rep* 11:180–186
- Chanenson S (2005) The next era of sentencing reform. *Emory Law J* 54:377–460
- Frankel M (1973) *Criminal sentences: law without order*. Hill and Wang, New York
- Frase R (1999) Sentencing guidelines in Minnesota, other states, and the Federal Courts: a twenty-year retrospective. *Fed Sentencing Rep* 12:69–82
- Frase R (2000) Is guided discretion sufficient? Overview of state sentencing guidelines. *St Louis Univ Law J* 44:425–446
- Frase R (2005) Punishment purposes. *Stanf Law Rev* 58:67–83
- Freed D (1992) Federal sentencing in the wake of the guidelines: unacceptable limits on the discretion of sentencers. *Yale Law J* 101:1681–1754
- Gottfredson S, Moriarty L (2006) Statistical risk assessment: old problems and new applications. *Crime Delinq* 52:178–200
- Hofer P, Allenbaugh M (2003) The reason behind the rules: finding and using the philosophy of the federal sentencing guidelines. *Am Crim Law Rev* 40:19–85
- Hyatt J, Chanenson S, Bergstrom M (2011) Reform in motion: the promise and perils of incorporating risk assessments and cost-benefit analysis into Pennsylvania sentencing. *Duq Law Rev* 49:707–749
- Kern R, Farrar-Owens M (2004) Sentencing guidelines with integrated offender risk assessment. *Fed Sentencing Rep* 16:165–169
- Kramer J, Ulmer J (2009) *Sentencing guidelines: lessons from Pennsylvania*. Lynne Rienner, Boulder
- Miller M (2004) Sentencing reform “reform” through sentencing information systems. In: Tonry M (ed) *The future of imprisonment*. Oxford University Press, New York
- Morris N (1974) *The future of imprisonment*. University of Chicago Press, Chicago
- Morris N, Tonry M (1990) *Between prison and probation: intermediate punishments in a rational sentencing system*. Oxford University Press, New York
- Reitz K (2005) The enforceability of sentencing guidelines. *Stanf Law Rev* 58:155–173
- Singer R (1978) In favor of “presumptive sentences” set by a sentencing commission. *Crime Delinq* 24:401–427
- Stith K, Cabranes J (1998) *Fear of judging: sentencing guidelines in the federal courts*. University of Chicago Press, Chicago
- Stith K, Koh H (1993) The politics of sentencing reform: the legislative history of the federal sentencing guidelines. *Wake Forest Law Rev* 28:223–290
- Tonry M (1993) The success of judge Frankel’s sentencing commission. *Univ Colo Law Rev* 64:713–722
- Tonry M (1996) *Sentencing matters*. Oxford University Press, New York
- United States Sentencing Commission (1987) *U.S. sentencing guidelines manual*. http://www.ussc.gov/guidelines/1987_guidelines/Manual_PDF/1987_Guidelines_Manual_Full.pdf
- United States Sentencing Commission (2004) *Fifteen years of guidelines sentencing: an assessment of how well the federal criminal justice system is achieving the goals of sentencing reform*. http://www.ussc.gov/Research/Research_Projects/Miscellaneous/15_Year_Study/15_year_study_full.pdf
- United States Sentencing Commission. An overview of the federal sentencing guidelines. http://www.ussc.gov/About_the_Commission/Overview_of_the_USSC/Overview_Federal_Sentencing_Guidelines.pdf
- Wolff M (2008) Evidence-based judicial discretion: promoting public safety through state sentencing reform. *N Y Univ Law Rev* 83:1389–1419

Sentencing Research

Jeffery T. Ulmer
The Pennsylvania State University, University
Park, PA, USA

Synonyms

[Penalties](#); [Punishment](#); [Sanctioning](#)

Overview

Research on sentencing has been an active field of inquiry for several decades. The 1980s, 1990s, and 2000s were particularly noteworthy. These decades saw advances in the quality of data, the sophistication of research methods, and improvements in theory. This entry provides a conceptual survey of research on noncapital sentencing outcomes since 2000. The entry first looks backward at the research agenda posed by reviews in the early 1980s and in 2000. Theoretical developments in the study of sentencing in the 1990s and 2000s are then discussed. An overview is next provided of recent sentencing research focused on the following: (1) court organizational and social contexts, (2) individual courtroom workgroup members, (3) disparity conditional on intersecting defendant characteristics, (4) victim characteristics, and (5) earlier case processing events and decisions.

The central focus of this entry is *noncapital sentencing decisions* as focal dependent variables. However, it is recognized that sentencing decisions are conditioned by and interdependent with a chain of earlier and later criminal justice processing decisions that are embedded in organizational and broader contexts. This entry also primarily focuses on research since 2000, discussed in the context of earlier reviews of sentencing research and theory in the 1980s and 1990s.

Sentencing Research in the 1990s: Responding to Issues and Opportunities

By far, the biggest development in the 1980s and 1990s was the availability of data collected by sentencing commissions in sentencing guideline states and the federal system. For the first time, researchers had at their disposal high-quality post-conviction, jurisdiction-wide data for very large numbers of cases. These data were disseminated by sentencing commissions with legal mandates to collect and make such data available. These data enabled progress on the following issues, raised by the important 1983 National Academy of Science assessment of sentencing research by Blumstein et al. (1983) and by Hagan and Bumiller in that report:

1. *The need for better measurement of legally relevant variables:* Sentencing commission data enabled substantially improved measures of offense severity and criminal history, which had often been measured crudely in earlier research.
2. *The need to study different sentencing outcome variables:* The Blumstein et al. (1983, p. 273) report called for "...more adequate treatments of the complexity of the dependent variable, sentence outcome."
3. *The need for larger samples from more varied jurisdictions* was particularly identified by Hagan and Bumiller (1983). Guideline data from states and the federal system provided very large samples drawn from large numbers of contextually diverse counties (for states) or district courts (with federal data).
4. *The need to investigate the effects of social contexts on sentencing and disparity* was noted by Hagan and Bumiller (1983) among others. As noted above, sentencing data from a relatively wide range of local jurisdictions enabled research on local variation in sentencing severity and disparity.

Several chapters of the Blumstein et al. (1983) report and others called for consideration of sentencing in the contexts of earlier criminal justice decisions, particularly pretrial detention, charging, and conviction. It has long been recognized that sentencing is conditioned by selection

processes in earlier justice system decisions. For all their benefits noted above, guideline sentencing data are limited to convicted offenders and do not include preconviction information such as the type and number of original charges (and rarely include pretrial detention). Datasets that include preconviction data for large numbers of cases and multiple local jurisdictions have traditionally been difficult to obtain. The years since the Blumstein et al. (1983) report have seen comparatively few studies that examine sentencing in the context of earlier decisions.

Relatedly, the Blumstein et al. (1983) report called for sentencing research to examine the effects of defendant socioeconomic status variables on sentencing outcomes and to disentangle the potential influence of race and ethnicity versus socioeconomic factors in sentencing. In addition, the Blumstein et al. (1983) report and others called for studies that examine sentencing variation between criminal justice actors, especially judges and prosecutors, and the effects of decision maker characteristics in explaining variation in sentencing. But, only a handful of studies since then have examined sentencing variation between judges, and the effect of judge characteristics, and even fewer studies have examined between-prosecutor variation. Also, defendant socioeconomic data and/or data on individual court actors has generally been hard to come by in the most widely used guideline datasets. Thus, studies that disentangle socioeconomic factors from other defendant characteristics (especially race or ethnicity) have not been as common.

Reviews of the 1990s

Two comprehensive reviews (Spohn 2000; Zatz 2000) summed up research on sentencing up to 2000, primarily focusing on studies of disparity based on defendant social characteristics. Both reviews noted that research had improved in quality since the 1983 report. Both reports also highlighted research that went beyond simply assessing whether “race/ethnicity mattered” or “gender mattered” to investigating *when* and

how such social statuses matter in sentencing – that is, investigating how the influence of race, ethnicity, and gender mutually conditioned one another and were also conditioned by other factors. Spohn’s (2000) review in particular provided a quantitative summary of racial/ethnic disparity studies with an eye to conditional influences.

Regarding the incarceration decision, Spohn (2000) found that 55 % of 45 effects in 31 state court studies she reviewed, and three of seven effects in eight federal court studies, showed significant black disadvantages. More importantly, Spohn (2000, see p. 462) found that eleven studies had looked at whether race and/or ethnic effects were conditioned by other social statuses or characteristics, especially gender and/or age. All eleven found evidence of such effects. In particular, four studies found sentencing disadvantages for young black and/or Hispanic *males*, four found sentencing disadvantages for unemployed young black or Hispanic males, one found disadvantages for poor minorities, and one found disadvantages for low-education minorities.

In addition, several studies in the 1990s showed that race or ethnicity effects on sentencing were conditional on criminal history, offense type or severity, case processing factors, or victim characteristics. Eight 1990s studies found that blacks or Hispanics with more serious criminal histories were sentenced more severely. Ten studies found that race/ethnic differences were conditioned by type or severity of crimes. Three found that racial disadvantages were greatest in less serious crimes, and six found that racial disadvantages were pronounced in drug offenses. In addition, six studies found that race/ethnicity effects were conditioned by case processing factors. Two studies found that blacks were sentenced more severely if they were detained prior to trial, one found that blacks were sentenced more severely if represented by a public defender rather than a private attorney, and three found that blacks were sentenced more severely if they were convicted by plea rather than trial. Finally, one study found that blacks who victimized whites received more severe sentences than other victim-offender combinations.

These reviews also spelled out directions for further research that they hoped would be undertaken in the ensuing decade. Several of these directions called for further progress on earlier recommendations from the Blumstein et al. (1983) report. Among other directions, these reviews called for (1) investigations of how the effects of race and ethnicity on sentencing vary by social context and over time and how the effects of race and ethnicity might be conditioned by other factors, (2) more investigations of the sentencing of Latinos, (3) research that utilized better measures of defendant socioeconomic status factors and disentangling the effects of socioeconomic status from race and ethnicity, (4) research on how court social and organizational contexts shaped sentencing severity and disparity, (5) investigations of sentencing and disparity in the context of earlier criminal justice (particularly prosecutorial) decisions, and (6) qualitative research that illuminated sentencing processes.

Theoretical Perspectives on Sentencing

Hagan and Bumiller (1983) as well as other contributors to the Blumstein et al. (1983) report noted that at that time there was little theoretical development in the sentencing literature. Spohn (2000, p. 458) also noted that most pre-1990s sentencing studies were based on “an overly simplistic version of conflict theory.” However, the 1990s and 2000s saw the emergence of new theoretical frameworks and the refinement of older theoretical propositions from conflict and labeling theories. These theories seem to be more complementary than competing, and some of them have mutually influenced each other. Below are brief sketches of these theoretical frameworks (readers seeking full treatments should consult the original works cited below).

Individual Cases and Actors

Celesta Albonetti articulated an *uncertainty avoidance* and *causal attribution* (1991) perspective on sentencing and court decision making. Albonetti applied insights from organizational theory to argue that sentencing suffers from operating in a context of bounded rationality in that

court actors make highly consequential decisions with insufficient information, which produces uncertainty. Albonetti (1991) particularly stressed uncertainty and insufficient information regarding the recidivism risk and rehabilitative potential of offenders. Albonetti drew from attribution theory in social psychology to argue that, as a means of reducing uncertainty, decision makers fall back on attributions about reoffending risk and/or rehabilitation potential that can be linked to race and gender and other social status stereotypes. This can then result in extralegal sentencing disparity connected to these statuses.

Similarly, Farrell and Holmes (1991) presented an *interpretive theory of legal decision making*, in which they emphasized the situational role of stereotypes linked to defendant social statuses in case processing. Farrell and Holmes (1991) generated ten propositions about the conditional, situation-specific role of status-linked stereotypes for routine and nonroutine cases and defendants. A related idea is the *liberation hypothesis* which has sometimes been applied to sentencing discretion (Spohn and Cederblom 1991). The liberation hypothesis implies that as the seriousness and/or visibility of the offense or case increases, sanctioning discretion is tightened and legally relevant variables are decisive, leaving little room for extralegal influences. By contrast, in less serious/visible cases, opportunities for discretion are greater and extralegal variables can influence outcomes more than in serious cases.

Because the sentencing guidelines movement of the 1980s transformed the sentencing landscape, the 1990s saw important treatments of sentencing discretion and guidelines, which involves the management of several dilemmas: between flexible discretion and rule-bound control, between uniformity and individualization, and between centralization and decentralized localism. Joachim Savelsberg (1992) helpfully conceptualized sentencing and sentencing guidelines as an attempt to impose a regime of what founding early sociologist Max Weber called formal rationality onto a traditionally “substantively rational” process. *Substantive rationality* refers

to an individualized decision process governed by criteria that are in service of ideological goals external to the law. The flexibility inherent in substantive rationality, however, permits the possibility of bias, discrimination, and unwarranted disparity.

The *focal concerns perspective* emphasizes particular kinds of substantively rational criteria at work in sentencing decisions. A recent review and statement of the focal concerns perspective in proposition form can be found in Kramer and Ulmer (2009). The focal concerns perspective argues that court actors' subjective definitions of offenders and offenses in relation to three focal concerns of punishment – blameworthiness, protection of the community, and practical constraints – determine punishment decisions. The focal concerns perspective argues that both legal and extralegal considerations affect the interpretation and prioritization of focal concerns through local substantive rationality (Kramer and Ulmer 2009). The influence of race, for example, may be conditional on defendant's gender, age, social class, legally relevant factors, and especially local contexts (see below).

Rational choice approaches have also been applied to sentencing in the times since the 1983 report (Piehl and Bushway 2007). Proponents of a rational choice approach to sentencing research would call for researchers to identify the formal and informal incentive structures of different court actors involved, make theoretical predictions about what sentencing outcomes would look like if actors acted rationally in pursuit of those incentives, and then compare these predictions to real-world data. Actors' incentives, of course, are strongly influenced by the organizational environments of prosecutors' offices, bench, and defense bar.

Court Organization and Multilayered Environments

It has long been recognized that sentencing practices varied between jurisdictions, even within states or the federal court system. Myers and Talarico (1987), Eisenstein et al. (1988), Ulmer (1997), and others drew attention to the

localization of sentencing. The *court community perspective* views courts as communities based on participants' shared workplace; interdependent working relations between key sponsoring agencies such as the prosecutor's office, judges' bench, and defense bar; and the court's relation to its larger sociopolitical environment (Eisenstein et al. 1988; Ulmer 1997). These court communities are said to foster their own locally varying sentencing norms which influence sentencing as least as much as formal policies and legal structures (Ulmer and Kramer 1996).

Interest in social contexts that foster negative racial/ethnic-based stereotypes and perceptions of threat has been fostered by *racial threat theory*. This theory argues that as minority racial groups grow in size relative to whites, they are likely to develop greater power, economic resources, and political influence in the community and are better able to compete with whites for power. In the context of sentencing, racial group threat theory implies that when perceptions of minority group threat are more pronounced and when courtroom actors perceive particular racial/ethnic groups as more dangerous or morally disrespectful, such minorities may receive harsher sentences.

Finally, *organizational efficiency/maintenance models* of sentencing have long been recognized (Dixon 1995; Engen and Steen 2000). In fact, the focal concerns model would also recognize organizational efficiency as a potentially important practical constraint faced by court actors (Kramer and Ulmer 2009). These organizational efficiency model views emphasize leniency for those who plead guilty and avoid time- and resource-intensive trials as an effort by courts to keep cases moving smoothly. Relevant to rational choice theory, an organizational efficiency proposition would suggest that rewarding guilty pleas and punishing trials is an organizationally rational response to the need to move cases efficiently. Furthermore, this differential punishment would be conditioned by caseload pressure – the greater the caseload pressure, the more courts would rely on such costs and incentives.

Research Since 2000

These theoretical developments have been instrumental in five broad and interrelated areas of empirical inquiry discussed below.

Social Contexts and Sentencing

The 2000s saw a burst of studies on contextual effects and variation and how the influence of individual-level factors was conditioned by court or social context factors. Overall, the recent literature on contextual variation in sentencing shows that local variation permeates many aspects of sentencing, both under sentencing guideline jurisdictions and non-guideline jurisdictions. Studies typically find that most sentencing outcome variation exists at the individual level and is most strongly predicted by individual-level factors. However, not only does sentencing severity (and related outcomes such as guideline departures) vary between local courts and their contexts but so too sometimes do the effects of other important legally relevant and extralegal sentencing predictors.

Racial/Ethnic Population Composition

Some of this research has investigated the role of local racial/ethnic composition in conditioning racial and ethnic sentencing disparity, drawing from both racial group threat theory and the focal concerns perspective. Many multilevel sentencing studies find that the effects of race and ethnicity in sentencing decisions do indeed vary significantly across courts. However, results have been decidedly mixed regarding racial threat theory's ability to explain this variation. That is, race/ethnic effects on sentencing tend to vary across contexts, but not always in ways predicted by racial threat theory. Some studies have found that the percentage of blacks in local populations has been found to increase racial/ethnic disparities in sentencing. Other studies reveal either no support for racial threat or evidence contrary to racial threat hypotheses (see Ulmer 2012).

Local Organizational Constraints

Other research has examined how local practical constraints such as caseloads or local criminal

justice resources (i.e., jail space) affect sentencing, research relevant for the organizational efficiency hypothesis, as well as the focal concerns and court communities' perspectives. Pennsylvania court caseloads were negatively related to sentencing severity (see in Kramer and Ulmer 2009). Local jail space in Pennsylvania counties affected the probability of incarceration and found that local jail space predicted the choice of county jail versus state prison (Kramer and Ulmer 2009).

Sociopolitical Influences on Sentencing Decisions

Though less common, studies have examined other sociopolitical influences on sentencing, including crime rates, political climate, neighborhood disadvantage, and local religion (see review by Ulmer 2012). Crime rates and broad political climate measures (such as percent Republican voters) have generally not been found to be strong predictors of sentencing patterns or to strongly and consistently condition individual-level predictors (an exception is Johnson 2006). Fearn (2005) and Ulmer et al. (2008) found evidence that local religious contexts may affect sentencing patterns. Fearn (2005) found that prison sentences were more likely in jurisdictions with greater proportions of evangelical Christians. Ulmer et al. (2008) found that local religious homogeneity fostered greater use of incarceration in Pennsylvania. Unlike Fearn (2005), Ulmer et al. (2008) did not find that evangelical Christian prevalence influenced sentencing but rather that local religious homogeneity interacted with percent Republican voters, in that counties that were religiously homogeneous *and* strongly Republican were most likely to incarcerate offenders.

Court Community Racial/Ethnic Composition

Farrell et al. (2009) studied the effects of federal district court community racial composition on variation in the effects of race on sentencing. They found that district US Attorney's Office's black representation was associated with significantly smaller racial disparities in incarceration, and interestingly, greater black Federal Probation Office representation was associated with greater

black/white sentencing outcome disparity. In addition, studies have found that greater black representation among county attorneys attenuated local black/white sentencing disparity (see in Ulmer 2012).

Court Personnel

Earlier research had investigated the race and gender of judges on sentencing, and this was reviewed in Spohn (2000). Johnson (2006) examined court contextual and inter-judge variation in sentences in Pennsylvania. He found that black and Hispanic judges sentenced all offenders, and particularly minority offenders, more leniently than white judges. Furthermore, male judges sentenced female offenders more leniently. Studies have also examined inter-judge variation in sentencing and also found wide variation between judges in the effects of race, gender, or financial means on their sentences. It has also been found that judges considered legal and extralegal factors quite differently in their sentencing decisions (see in Ulmer 2012).

Conditional Disparity

In addition, a key theme of Spohn's (2000) and Zatz's (2000) summaries of the 1990s literature was that the influence of social status factors such as race, ethnicity, and gender was conditional and mutually contingent. Dozens of studies in the 2000s have continued to confirm this insight. Many of these studies are noted above and below in the discussions of research on court contexts, courtroom actors, victim characteristics, and earlier case processing decisions. In sum, there appears to be substantial evidence that the effects of such extralegal social statuses are conditioned by court contextual factors and provocative evidence that disparity varies across individual court community actors. Furthermore, most studies that examine the issue find young black, and to a lesser extent Hispanic, male defendants to be sentenced more severely (see reviews by Mitchell 2005; Ulmer 2012). Also, the effects of defendant social statuses may interact with case processing, offense characteristics, and criminal history.

Victim Characteristics

It should be noted that existing studies that utilize data from sentencing guideline jurisdictions actually implicitly consider victim harm, financial loss, and often victim age and vulnerability. These factors are commonly included in creating guideline offense severity rankings. Thus, the strong effects of offense severity typically found in such studies in part incorporate victim impact and vulnerability. Such research, however, does not tell us about the effects of other specific victim attributes. Curry (2010) found that Texas offenders who victimized females received longer sentences, and females who victimized males received sentences about 10 months shorter than males who victimized females. Curry's research also found that violent offenders who victimized white and Hispanic females received longer sentences, and Hispanic and black homicide offenders who victimized whites got longer sentences than other combinations. Other studies have uncovered similar findings (see in Ulmer 2012). They found no race of victim/offender combination effects, but found that males who killed females received the longest prison sentences and that females who killed males received shorter sentences than males who killed females. Overall, there is mixed evidence that victim characteristics may matter in noncapital violent crime sentencing.

Sentencing in the Context of Earlier Case Processing

A comparatively smaller set of studies has examined sentencing as related to earlier case processing events, such as guilty pleas versus trial convictions, as well as charging decisions and pretrial release. This research provides some, albeit limited, insight into prosecutorial discretion as well as other joint courtroom workgroup outcomes, such as differentially punishing those convicted by trial relative to those who plead guilty.

Charging Decisions

Other researchers have investigated prosecutors' charging decisions and their impact on eventual sentencing. Studies have found that the number

of indictment charges filed by federal prosecutors significantly increased federal sentence lengths, and decreased the likelihood of substantial assistance or other downward departures. Shermer and Johnson (2010) examined the likelihood of federal prosecutors reducing charges for defendants. They found that while about 12 % of federal cases in their sample involved charge reductions, race/ethnicity and gender (along with offense severity and criminal history) influenced the likelihood of those charge reductions in drug and violent offenses. These charge reductions, in turn, resulted in lower sentences for those who received them and charge reductions-mediated race/ethnic effects on sentences.

Pretrial Detention/Release

Though not as plentiful as studies of sentencing outcomes, there is a literature focusing on the consequences of these decisions for defendants' sentencing. Studies have found that defendants subject to pretrial detention were more likely to receive more severe sentences than defendants who had been released before case disposition. Other studies have Spohn examined federal offenders in three district courts and found that both legal and extralegal factors predicted pretrial status and that pretrial status significantly impacted sentencing. These found that race and gender had indirect effects on sentence severity through their effects on pretrial status.

Mandatory Minimums

Sentencing involving mandatory minimums has proved useful for understanding how earlier decisions shape and constrain sentencing choices. Bjerk (2005) found that prosecutors used their charge reduction discretion to circumvent three strikes mandatories for some defendants. He found that such circumvention of three strikes mandatory minimums was moderately less likely to occur for men, Hispanics, and to a lesser extent, blacks. Multilevel analyses of prosecutorial discretion in applying mandatory minimums among mandatory-eligible offenders sentenced for drug crimes or as "three strikes" offenders (see in Kramer and Ulmer 2009). They found that prosecutors' decisions to apply mandatory

minimums were significantly affected by mode of conviction (negotiated guilty pleas greatly assisted defendants in avoiding mandatories), the type and characteristics of offenses and guideline sentence recommendations (the greater the difference between the mandatory and the applicable guideline sentence, the less likely the mandatory was applied), prior record, and gender. In addition, Hispanic males were substantially more likely to receive mandatory minimums.

Federal Substantial Assistance Departure Motions

Some studies have examined prosecutorial discretion in decisions to file motions for substantial assistance guideline departures in federal court crack and powder cocaine sentences in 2000. These found that the likelihood of substantial assistance motions increased with offense severity and criminal history and that having multiple concurrent charges reduced the likelihood of such motions for crack cocaine cases. In powder cocaine cases, black and Hispanic males had the lowest odds of receiving substantial assistance departures, while white females and males had the greatest likelihoods. The results were similar for crack cocaine cases, except that Hispanic females had the greatest odds of substantial assistance departures.

In addition, Johnson et al. (2008) examined interdistrict variations in the application of both substantial assistance and other downward departures among a variety of offense types in federal sentences from 1997 to 2000. Findings indicated considerable between-district variation in the probability of these prosecutor-initiated substantial assistance departures. This variation was explained, in part, by organizational court contexts such as caseload pressures and by environmental considerations such as the racial composition of the district.

Spohn and Fornango (2009) examined variation between federal prosecutors in three districts in their likelihood of moving for substantial assistance departures. They found substantial interprosecutor variation in the likelihood of substantial assistance departures, net of the influence of individual case and defendant characteristics. Specifically, about 24 % of the variation between

prosecutors in the likelihood of substantial assistance departures was unaccounted for by case or defendant factors, and even more between-prosecutor variation was unexplained in nondrug cases.

Mode of Conviction

Finally, several studies have examined how mode of conviction affects sentencing; that is, sentencing differences between types of guilty pleas and types trials. This research exhibits some limitations in that studies typically use data on convicted offenders only, which presents issues of selection and potential missing-variable bias absent data on the likelihood of acquittal (see Bushway and Piehl 2007). Still, one overarching lesson from the recent studies below is that sentencing guidelines seem to provide a discretionary framework within which to differentially reward guilty pleas and punish trials.

Some studies have found that mode of conviction moderated the role of race/ethnicity in predicting guideline departures, with blacks and Hispanics experiencing different odds of receiving downward or upward departures, depending on their modes of conviction. Studies have looked at trial penalties for serious violent offenders and less serious offenders using hierarchical models with cases sentenced under Pennsylvania's guidelines (Kramer and Ulmer 2009). They found that defendants were substantially penalized if they were convicted by trial relative to those with negotiated or open guilty pleas. Furthermore, this jury trial penalty varied depending on the seriousness and type of offense (more severe offenses had lesser trial penalties), defendant criminal history (offenders with more substantial criminal histories actually experienced less of a trial penalty), race (blacks experienced greater trial penalties) and court contextual characteristics such as court community size (larger trial penalties in larger courts), local violent crime rates (the higher the crime rate, the greater the trial penalties), and the size of local black populations (greater trial penalties in counties with larger black populations). Studies have also found that federal trial penalties could not be fully explained by US Sentencing

Guideline factors that were relevant to mode of conviction. They also found that higher district court caseload pressure was associated with greater trial penalties, while higher district trial rates were associated with lesser trial penalties (see in Ulmer 2012).

New Methodological Extensions

The past decade has seen the application of new statistical methods to sentencing research questions, and several useful quantitative modeling alternatives now confront sentencing researchers. Bushway et al. (2007) present a very useful discussion of the merits and demerits of two-part models (modeling incarceration and sentence length separately), tobit, and Heckman two-step corrections with ordinary least squares regression in addressing problems of censoring and/or selection surrounding the incarceration and sentence length decisions. Several researchers have also demonstrated the usefulness of using multinomial logistic regression to predict different types of incarceration, such as county jail versus state prison, either in an individual level (see review in Ulmer 2012). In addition, Britt (2009) proposes *quantile regression* as an interesting alternative for assessing variation in the effects of predictors of interest (legally relevant, extralegal, case processing, etc.) across the distribution of sentence length/severity. That is, the quantile regression approach allows the researcher to separate the sentence length distribution into quantiles and examine how the strength of predictors' effects varies across those quantiles. Finally, propensity score methods appear to be a promising alternative for examining sentencing disparity in a way that attempts to create balanced, comparable samples of offenders that differ only on a "treatment" or characteristic of interest (e.g., race or mode of conviction). This therefore allows the construction of useful counterfactuals, for example, "what would the sentence be if this case involved a white rather than a black defendant."

The Need for More International Research

Most sentencing research is limited to the contemporary North American – particularly the US – context. However, particularly useful overviews

of the comparatively small amount of international literature on sentencing disparity that does exist can be found in Tonry and Frase (2001). Also, Johnson et al. (2010) studied 1,613 Dutch homicide offenses from 1993 to 2004 and found that homicide offenders that victimized youth under 18, elderly people, women, and Dutch (vs. foreign) victims were sentenced more severely than those victimizing other types of victims. Some other suggestive information on ethnic disparity comes from Canada and Australia (see reviews in Tonry and Frase 2001; Dawson 2006).

Comparatively little research exists on sentencing in non-Western contexts, particularly, in Asian countries, which are growing in global prominence. Lu and Kelly (2008) provide a useful summary of research on courts and sentencing in China. Two studies focus on sentencing in South Korea, a country which very recently adopted sentencing guidelines. These found that female drug offenders were sentenced more leniently than their male counterparts, though this gender difference disappeared among those with prior criminal records. Also, research has found that legal factors connected to the offense and past criminal behavior primarily determined the length of sentences for Korean marijuana and methamphetamine offenders but that males and older offenders received longer sentences. They also found that admitting guilt in court resulted in shorter sentences (see review in Ulmer 2012).

Finally, a great deal of research has been done on sentencing under sentencing guidelines, particularly, the US Sentencing Guidelines, and state guideline systems in Minnesota, Pennsylvania, Washington, and to a lesser extent Florida (see reviews by Spohn 2000; Zatz 2000; Kramer and Ulmer 2009; Ulmer 2012). More recently, research has examined the extent to which US federal sentencing and disparity changed in the wake of important Supreme Court decisions that rendered the federal guidelines advisory (see Ulmer et al. 2011). This research is mixed regarding whether racial disparity has increased under the advisory guidelines – the US Sentencing Commission’s analyses argue that racial disparity increased, but independent researchers have disputed their findings, arguing that overall, the data

do not support the notion that advisory guidelines have aggravated racial disparity (see Ulmer et al. 2011). Research has examined sentencing disparity, patterns, and trends within guideline systems, but relatively little research has compared differences in sentencing disparity between guideline systems (Engen 2009). Very little research has examined differences in sentencing between guideline and non-guideline jurisdictions, likely because of the lesser availability of comparable sentencing data from non-guideline systems and because of the methodological difficulties of comparing measures of key variables such as offense severity and criminal history across guideline and non-guideline contexts.

In sum, since 2000, research has made some progress on certain issues identified by the Blumstein et al. (1983) report and the Spohn (2000) and Zatz (2000) reviews. For example, in many sentencing studies published since 2000, researchers have moved beyond examining the traditional incarceration and length to look at more refined outcomes, such as distinctions between types of incarceration (e.g., probation, county jail, state prison). Also, studies have examined the imposition of mandatory minimums, which are tightly linked to prosecutorial discretion, adjudication waivers, first-time offender waivers, and special sex offender sanctions, and different types of sentencing guideline departures (see Ulmer 2012). Further, research in the 2000s generally heeded the call to include Hispanic defendants, and many, if not most, of the recent studies on US racial sentencing disparity include black, white, and Hispanic comparisons. Other directions, such as examining sentencing in the context of preconviction decisions and prosecutorial discretion, as well as research on sentencing in non-US and non-Western contexts, have seen comparatively less progress.

Related Entries

- ▶ [Comparative Courts and Sentencing](#)
- ▶ [Drug Courts](#)
- ▶ [International Sentencing](#)
- ▶ [Mitigation and Aggravation at Sentencing](#)

- ▶ [New Penology](#)
- ▶ [Plea Bargaining](#)
- ▶ [Problem-Solving Courts](#)
- ▶ [Prosecutorial Discretion](#)
- ▶ [Sentencing as a Cultural Practice](#)
- ▶ [Sentencing Commissions](#)
- ▶ [Sentencing Guidelines in the United States](#)
- ▶ [Sex Offenders and Criminal Policy](#)
- ▶ [Structured Sentencing Outside the United States](#)
- ▶ [Theories of Punishment](#)

Recommended Reading and References

- Albonetti C (1991) An integration of theories to explain judicial discretion. *Soc Probl* 38:247–266
- Bjerk D (2005) Making the crime fit the penalty: the role of prosecutorial discretion under mandatory minimum sentencing. *J Law Econ* 48:591–625
- Blumstein A, Jacqueline C, Susan M, Michael T (eds) (1983) *Research on sentencing: the search for reform*. National Academy Press, Washington, DC
- Britt C (2009) Modeling the distribution of sentence length decisions under a guideline system: an application of quantile regression models. *J Quant Criminol* 25(4):341–370
- Bushway S, Brian J, Slocum LA (2007) Is the magic still there? The relevance of the Heckman two-step correction for selection bias in criminology. *J Quant Criminol* 23:151–178
- Bushway SD, Piehl AM (2007) Social science research and the legal threat to presumptive sentencing guidelines. *Criminology & Public Policy* 6:461–482
- Curry T (2010) The conditional effects of victim and offender ethnicity and victim gender on sentences for non-capital cases. *Punishm Soc* 12(4):438–462
- Dawson M (2006) Intimacy and violence: exploring the role of victim-defendant relationship in criminal law. *J Crim Law Criminol* 96(4):1417–1449
- Dixon J (1995) The organizational context of criminal sentencing. *Am J Sociol* 100:1157–1198
- Eisenstein J, Flemming R, Nardulli P (1988) *The contours of justice: communities and their courts*. Little, Brown, Boston
- Engen R (2009) Assessing determinate and presumptive sentencing: making research relevant. *Criminol Public Policy* 8:323–335
- Engen R, Steen S (2000) The power to punish: discretion and sentencing reform in the war on drugs. *Am J Sociol* 105:1357–1395
- Farrell R, Holmes M (1991) The social and cognitive structure of legal decision making. *Sociol Q* 32(4):529–542
- Farrell A, Ward G, Rousseau D (2009) Race effects of representation among federal court workers: does black workforce representation reduce sentencing disparities? *Ann Am Acad Pol Soc Sci* 623:121–133
- Fearn NE (2005) A multilevel analysis of community effects on criminal sentencing. *Justice Quart* 22:452–487
- Hagan J, Bumiller K (1983) Making sense of sentencing: race and sentencing outcomes. In: Blumstein A, Cohen J, Martin S, Tonry M (eds) *Research on sentencing: the search for reform*. National Academy Press, Washington, DC
- Johnson BD (2006) The multilevel context of criminal sentencing: integrating judge and county level influences in the study of courtroom decision making. *Criminology* 44:259–298
- Johnson B, Ulmer J, Kramer J (2008) The social context of Guideline circumvention: the case of federal district courts. *Criminology* 46:711–783
- Johnson BD, van Wingerden S, Nieuwebeerta P (2010) Sentencing homicide offenders in the Netherlands: offender, victim, and situational influences in criminal punishment. *Criminology* 48(4):981–1018
- Kramer J, Ulmer J (2009) *Sentencing guidelines: lessons from Pennsylvania*. Lynne Rienner Publishers, Boulder
- Lee M, Ulmer JT, Park MR (2011) Drug sentencing in South Korea: the influence of case processing and social status factors in an ethnically homogeneous context. *J Contemp Crim Justice* 27(3):378–397
- Lu H, Kelly B (2008) Courts and sentencing research on contemporary China. *Crime Law Soc Change* 50:229–243
- Mitchell O (2005) A meta-analysis of race and sentencing research: explaining the inconsistencies. *J Quant Criminol* 21(4):439–466
- Myers M, Talarico S (1987) *The social contexts of criminal sentencing*. Springer, New York
- Piehl AM, Bushway SD (2007) Measuring and explaining charge bargaining. *J Quant Criminol* 23(2):105–125
- Savelsberg JJ (1992) Law that does not fit society: sentencing guidelines as a neoclassical reaction to the dilemmas of substantivized Law. *Am J Sociol* 97:1346–1381
- Shermer LON, Johnson B (2010) Criminal prosecutions: examining prosecutorial discretion and charge reductions in U.S. Federal District Courts. *Justice Quarterly* 27(3):394–430
- Spohn C (2000) Thirty years of sentencing reform: the quest for a racially neutral sentencing process. In: *Policies, processes and decisions of the criminal justice system*. Vol. 3, criminal justice 2000. U.S. Department of Justice, Washington, DC
- Spohn C, Cederblom J (1991) Race and disparities in sentencing: a test of the liberation hypothesis. *Justice Q* 8:305–327
- Spohn C, Fomango R (2009) US Attorneys and substantial assistance departures: testing for interprosecutor disparity. *Criminology* 47(3):813–847

- Tonry M, Frase R (eds) (2001) *Sentencing and sanctions in Western countries*. Oxford University Press, New York
- Ulmer JT (1997) *Social worlds of sentencing: Court communities under sentencing guidelines*. State University of New York Press, Albany
- Ulmer JT (2012) Recent developments and new directions in sentencing research. *Justice Q* 29(1):1–40
- Ulmer JT, Kramer JH (1996) Court communities under sentencing guidelines: Dilemmas of formal rationality and sentencing disparity. *Criminology* 34:383–408
- Ulmer JT, Bader C, Gault M (2008) Do moral communities play a role in criminal sentencing? Evidence from Pennsylvania. *Sociol Q* 49(4):737–768
- Ulmer JT, Light M, Kramer J (2011) Racial disparity in the wake of the *Booker/Fanfan* decision: an alternative analysis to the USSC's 2010 Report. *Criminol Public Policy* 10(4):1077–1118
- Zatz M (2000) The convergence of race, ethnicity, gender, and class on court decision making: looking toward the 21st Century. In: *Policies, processes and decisions of the criminal justice system*. Vol. 3, criminal justice 2000. U.S. Department of Justice, Washington, DC

the Good Lives Model (GLM). It then provides a discussion of some of the relevant process issues, namely, whether treatment should be conducted in groups, and the problems of confrontation and coercion. The entry concludes with a description of the primary components of group treatment approaches and a discussion of treatment efficacy.

Treatment Models

In order to fully understand contemporary sex offender treatment, it is useful to first summarize the models which underlie treatment. Obviously, it is not possible to describe all of the relevant models herein; thus, two models which are considered to be especially influential will be discussed, namely, the Risk Needs Responsivity model (Andrews and Bonta 2003a) and the Good Lives Model (Ward and Stewart 2003).

Sequential Lineups

- ▶ [Eyewitness Research](#)

Serial Victims

- ▶ [Multiple Victims and Super Targets](#)

Sex Offender Treatment

Jo Thakker
School of Psychology, University of Waikato,
Hamilton, New Zealand

Overview

This entry provides a brief summary of three key aspects of the treatment of sexual offenders. It begins with a description and critique of the two of the most influential treatment models, namely, the Risk Needs Responsivity (RNR) model and

The Risk Needs Responsivity Model

Andrews and Bonta's (2003a) Risk Needs Responsivity (RNR) approach has been important in guiding treatment across a variety of offender groups, including sexual offenders. Since its development in the late 1990s, it has played a significant role in the way that treatment has been conceptualized and delivered. As indicated by the name of the model, there are three main components. The risk aspect of the model refers to the type and degree of risk that an offender poses and suggests that the risk level should determine the intensity of treatment. Thus, low-risk offenders should receive less intensive treatment, and high-risk offenders should receive higher intensity treatment. As explained by Andrews and Bonta, intensity essentially refers to the duration of treatment.

The need aspect of the RNR model refers to the offender's specific rehabilitative requirements, which are variously termed criminogenic needs or dynamic risk factors. Essentially, these are the psychological characteristics and/or behaviors of the individual which require treatment. According to the model, treatment should target the specific characteristics that contribute

to the individual's offending so that the likelihood of reoffending will be reduced. As outlined by Andrews and Bonta, some common examples of these sorts of characteristics include antisocial attitudes and beliefs, substance abuse, and relationship problems.

The third part of the RNR model is the responsivity principle which is concerned with the actual delivery of treatment. As stated by Andrews and Bonta, the responsivity principle acknowledges the significance of the therapeutic relationship and assumes that an offender's response to treatment will be dependent on this, as well as a range of other variables. Also, the model takes the view that in order for treatment to be effective, it needs to identify and manage barriers to treatment and build on the individual's unique strengths. Furthermore, the responsivity principle espouses that in order for treatment to be successful, it needs to be designed according to best practice guidelines; in other words, it is important that there is evidence that a particular therapeutic approach is effective in treating the offender's unique range of needs.

According to Ward et al. (2008, p. 180), the RNR model has "constituted a revolution in the way that criminal conduct is managed in Canada, Britain, Europe, Australia, and New Zealand." While the model is general, insofar as it has been applied to a range of offender types, it is nonetheless pertinent to this discussion as it has been widely used in the design of treatment of sexual offenders. One of the key strengths of the model is the ease with which it can be translated into treatment. It provides clear guidelines in terms of selecting appropriate treatment for offenders, and it assists in the identification of treatment targets. Also, it encourages the therapist to consider the offender's strengths and weaknesses and to design treatment accordingly. Furthermore, it emphasizes the link between theory, research, and practice and is thus consistent with the principles of the scientist-practitioner model.

However, some researchers have criticized certain aspects of the RNR model. For instance, it has been argued that the RNR model places too much emphasis on rehabilitative needs and not

enough emphasis on other aspects of treatment. For example, Wilson and Yates (2009) argue that the RNR model tends to overlook the importance of therapist qualities such as empathy and respect due to its focus on the needs of the offender. However, as outlined above, in a recent explication of the RNR model, Andrews and Bonta highlighted the significance of the therapeutic relationship in treatment success. Ward argues (e.g., Ward and Stewart 2003) that the RNR model is typically implemented in an inflexible manner which fails to adequately deal with individual needs. However, this may be an unfair criticism insofar as this may reflect an inadequate implementation of the model rather than a failing of the model per se. In their critique of the RNR model, Ward and Stewart (2003) also state that the model takes an overly negative approach to therapy by focusing on risk avoidance. They suggest that the model assists offenders in understanding what to avoid, but does not assist them in establishing new pro-social behaviors.

The RNR model continues to provide the framework upon which many contemporary treatment programs for offenders are designed and administered, and this is true also for sex offender programs. Furthermore, there is now mounting empirical support for the model. According to Andrews and Bonta (2010, p. 39), "...programs that adhere to the Risk Needs Responsivity (RNR) model have been shown to reduce offender recidivism by up to 35 %."

The Good Lives Model

The Good Lives Model (GLM) by Ward and colleagues (Ward and Stewart 2003) is another model that is influential in the area of offender rehabilitation. Furthermore, the GLM has been widely discussed in relation to sexual offenders. As illustrated above, Ward and various colleagues have criticized the RNR model, and they have developed the GLM, at least in part, to respond to some of the weaknesses of the RNR model. The GLM takes a very different approach to offender rehabilitation; rather than examining the risk level of an offender, it begins by asking what purpose the offending is serving. Specifically, the question is: What gap or need is the

offending filling? Ward and colleagues (e.g., Serran et al. 2007) suggest that in order to bring about long-term change in offenders' behavior, therapy should identify and respond to the underlying motivation for the behavior.

The GLM is founded on the proposition that all human activity revolves around the desire to fulfill basic needs or "goods." For instance, eating nutritious food and engaging in regular physical exercise might fulfill the basic good of maintaining physical well-being. Painting a picture or acting in a play may fulfill the basic good of creativity. When these sorts of "goods" are maintained pro-socially in an individual's life that it can be said that he or she has a "good" life. Ward also argues that the drive to attain such "goods" is inherent by virtue of one's humanity. For example, he states that "...both sexual and non-sexual offenders are naturally disposed to seek a range of primary human goods that if secured will result in greater self-fulfilment and sense of purpose" (Ward et al. 2009, p. 304).

In explicating the GLM, Ward draws a distinction between primary and secondary goods. While primary goods are essentially ends in themselves, secondary goods relate to the means of attaining primary goods. For example, as shown below, knowledge is viewed as a primary good; thus, secondary goods that might bring this about would include activities such as reading a book and engaging in educational activities. Drawing on theoretical work from a range of research areas, including anthropology, evolutionary psychology, and ethics, Ward and Stewart (2003) propose that there are at least ten primary human goods, namely, "Life (including healthy living and functioning), knowledge, excellence in play and work (including mastery experiences), excellence in agency (i.e., autonomy and self-directedness), inner peace (i.e., freedom from emotional turmoil and stress), friendship (including intimate, romantic and family relationships), community, spirituality (in the broad sense of finding meaning and purpose in life), happiness, and creativity" (p. 356).

Ward proposes that offenders, like all human beings, value these primary goods and seek to

manifest them in their daily lives. However, Ward stresses that for most offenders, and indeed most non-offenders, the drive to attain these goods is not necessarily part of one's conscious awareness. Thus, the drive to attain these goods is often natural and ingrained and not typically part of a carefully thought-out plan. Ward suggests that while offenders attempt to realize the same goods as those who do not commit crimes, they go about this in a problematic way and he describes four key types of problems. These are the use of *inappropriate means* (i.e., goods are sought in ways that are inappropriate and counterproductive), a *lack of scope* (i.e., only some goods are sought), *conflict or lack of coherence* (i.e., the ways some goods are sought directly reduces the chances of others being secured), and a *lack of capacity* (i.e., individuals lack the skills, opportunities, and resources to achieve a certain good in specific ways).

With regard to sexual offenders, the use of inappropriate means would be a common example. While a male sexual offender may seek a sexual partner as part of his intention to attain the "good" of "friendship," he may seek an inappropriate sexual partner. For instance, he may engage in sexual activity with a minor rather than an adult. Furthermore, his sexual offending may be related to a lack of capacity; he may lack the social skills required to establish and maintain an appropriate sexual relationship with an adult.

As outlined by Ward, a GLM approach to offender rehabilitation requires a comprehensive assessment of the relationship between the offending and goods that the offending is being used to attain. Then, a Good Lives Plan is developed which includes setting goals for the future which will ultimately bring about the manifestation of goods in a pro-social manner. Thus, treatment involves the acknowledgement that all human beings have basic needs and the encouragement of offenders to strive to meet these in their lives but to do so in the manner that is consistent with societal norms and laws. According to the GLM, when these basic needs are met, the risk for reoffending will be greatly reduced because the offender will no longer use offending as a means of achieving his goals.

The GLM approach to offender rehabilitation is consistent with the work of Maruna (e.g., 2001) which stresses the importance of taking a positive approach in therapy. The GLM is also compatible with the strengths-based approach which has become increasingly popular within the field of mental health treatment (Wong 2006). However, as yet, there is insufficient empirical evidence that applying the GLM to offender treatment is advantageous. Also, the suggestion put forward by proponents of the GLM that the more traditional RNR model has a negative focus and is not strengths-based has been challenged. For example, it has been noted that the RNR framework does have positive and forward-looking components (New Zealand Department of Corrections 2009). Furthermore, Andrews and Bonta (2003b) take issue with Ward's proposition that a focus on criminogenic needs and risk factors is incompatible with a focus on more positive factors.

At this time, in light of current research and theory, it would seem that a combination of the RNR model and a "good lives" approach is the best way forward in the treatment of sexual offenders. While the GLM still requires empirical support, given its alignment with positive and strengths-based approaches, which have already proven their worth, it is arguably a useful model which is likely to prove efficacious.

Process Issues

Over recent years researchers have predominantly focused on the content of sex offender treatment programs rather than the process. However, with regard to the process of treatment, there are a number of important issues which require attention.

Group Versus Individual Approaches

Treatment of sexual offenders is typically conducted in groups with between eight and ten participants with either one or two facilitators (Marshall 2001). As stated by Marshall, group treatment is more economical and has been found to be more effective. Marshall points out that in group settings, clinicians are less likely to

collude with offenders. Collusion is more likely with this offender group as due to society's strong aversion to sexual offending; it is especially difficult for these sorts of offenders to admit to their crimes. Sexual offenders are more likely to deny their offending or to minimize their degree of culpability, and this can pose a challenge in terms of establishing rapport in individual treatment. In contrast, in a group setting, participants are able to challenge each other which assists in maintaining the rapport between the facilitator and the offender.

As explained by Harkins and Beech (2008), effective group treatment is dependent on the presence of several qualities within the group. One that has been found to be particularly important is cohesiveness (Harkins and Beech 2008) which is essentially the positive feelings that group members have for each other and the extent to which they are able to work together. As stated by Harkins and Beech, the role of the therapist is crucial in establishing cohesiveness within the group. In particular, the therapist needs to earn the respect of the participants by displaying a range of positive characteristics, such as flexibility, warmth, and empathy. However, it is equally important that the therapist is firm and directive and is able to challenge the participant when necessary, although this should be done in a gentle and compassionate manner (Serran et al. 2003).

Many contemporary sex offender treatment programs utilize mixed groups, that is, they treat rapists and child sexual offenders together, and some researchers have questioned whether this is appropriate. For example, Harkins and Beech (2008) compared the "therapeutic climate" in mixed groups and in homogenous groups. They measured therapeutic climate with the Group Environment Scale (GES) which targets a range of phenomena, including expressiveness, cohesion, and leader support. Harkins and Beech found that there was no difference in the quality of the therapeutic climate between the group types. Furthermore, they found that both types of group appeared to have very positive therapeutic climates. Also, recidivism data examined in the study showed that the type of group that child

sexual offenders participated in had no impact on the likelihood of reoffending (however, similar data was not available for rapists).

Some researchers have argued that even when treatment is delivered within a group setting, it should still be tailored to the unique needs of the individual. For example, Marx et al. (1999) suggest that facilitators should adapt their approach to suit the various personality traits and learning styles of group members. Thus, it is proposed that even within a group setting, therapists can respond to individual needs. It is important to note though that such an approach would be dependent on a thorough and accurate assessment of the group participants which would allow for the identification of individual factors that would then facilitate appropriate tailoring of the therapist's interaction with group members.

Many group-based sex offender treatment programs provide individual therapy sessions as an adjunct to group treatment. For example, the Te Piriti Sex Offender Treatment Program in New Zealand includes individual therapy sessions alongside group-based treatment. These sessions allow the facilitator to provide additional support to the offender and to discuss any problematic issues that may arise during group sessions. They also provide an opportunity for the offender to raise issues that he may not feel comfortable raising within the group. For example, often offenders feel apprehensive about the idea of disclosing the details of their offending to their fellow group members, and thus individual sessions provide an opportunity to discuss such concerns prior to disclosure and to provide appropriate guidance and support.

Confrontation

There is an ongoing debate about the merit of using outright confrontation in the treatment of sexual offenders. This sort of approach is questioned because it is incompatible with the sorts of qualities which are considered to be fundamental to effective therapy. However, it is sometimes considered to be necessary because sexual offenders often deny or minimize their offending. Marques et al. (2005) state: "...too many sex offender treatment providers appear to

believe that it is necessary to be extremely confrontative when working with these clients" (p. 1098). However, as Marque and colleagues explain, such an approach is inconsistent with the principles of motivational interviewing (MI), principles which the research shows are very helpful in guiding offender treatment. Although one of the key tasks of MI is to "develop discrepancy," this should be done carefully and in conjunction with the other components, which include the expression of empathy and avoiding arguments.

Marshall and others (Marshall et al. 2003) examined the effect of therapist qualities on the efficacy of sex offender treatment and concluded that there are several therapist characteristics which are associated with treatment success. For instance, they found that emotional warmth, empathy, directiveness, and positive reinforcement were all associated with positive treatment outcomes. Arguably, these types of therapist qualities are not compatible with a confrontational approach; thus, Marshall's research raises further doubts about the usefulness of such an approach in sex offender treatment. Overall, research suggests that therapists working with sexual offenders should probably be supportively challenging. That is, they should challenge in a way that allows them to remain positive and encouraging. Supportive challenges differ from confrontational challenges in that they are typically expressed in a warm and understanding manner which displays a genuine interest in the individual and concern for his or her well-being.

Coercion

Another process issue, which is pertinent to the rehabilitation of sexual offenders, is the impact of coercion on treatment effectiveness. Although many sexual offenders are recommended by judges, parole boards, and psychologists to engage in offense-related treatment, many choose not to. A comprehensive study by Langevin (2006) found that approximately 50 % reported that they wished to undertake treatment and about 13 % successfully completed treatment. Thus, a high number of sexual offenders

decline treatment and only a small proportion complete treatment. Furthermore, with regard to those who do engage in treatment, there are a variety of reasons that may underlie their decision. For example, while some offenders may believe that they need treatment and will benefit from it, others simply undertake treatment in order to gain early release from prison. Thus, the provision of consent in a prison setting may not reflect a genuine motivation to engage in therapy but may simply reflect other motivations. Also, the offender may feel coerced by others; he may believe he has no choice but to abide by the wishes of the judicial system.

As explained by Burdon and Gallagher (2002), the coercion that forms the backdrop to offender treatment is part of broader societal efforts to coerce the offender, including imprisonment, registration, and community supervision. Thus, coercion reflects society's desire to control the offender in order to prevent reoffending. But of course, in terms of treatment some degree of motivation and cooperation is necessary in order for treatment to progress, and this is recognized by most therapists who are involved in selecting treatment participants. Also, often there are constraints on program delivery which limit the number that can be run at any one time, and this in turn limits the number of places that are available. Thus, it is often necessary to choose participants that are most likely to benefit from the program, and these are often those who express a desire to undertake treatment. However, it is interesting to note that involvement in sex offender treatment tends to increase an offender's openness to the idea of engaging in treatment and can lead to success in treatment, even in individuals who are initially unenthusiastic (Burdon and Gallagher 2002).

Treatment Components

There are many different components that may be used in the treatment of sexual offenders. The components that are covered herein are those that are commonly used in group treatment

approaches and which have been described in detail in the research literature.

Deviant Sexual Arousal

It is not surprising that deviant sexual arousal is often targeted in treatment as it has been found to be one of the most significant risk factors for sexual reoffending. However, it has been reported that deviant arousal is now not routinely targeted in treatment as it was previously, due to the acknowledgement that problematic sexual preferences do not always play a significant role in an individual's offending (e.g., Marshall 2006). Also, the factors that lead to sexual offending are many and varied and differ across offender and offense type. For example, in rape offenses, anger and aggression may play a more significant role than problematic sexual preferences. Further, research suggests that sexual arousal to stimuli associated with rape has been found in men who have not committed rape; thus, it is not necessarily associated with deviant behavior. Thus, it may be the case that offense-related sexual preference is more likely present in child sexual offenders. Nonetheless, many sexual offender treatment programs, for those with adult and juvenile victims, continue to include a component that focuses on deviant sexual arousal (Marshall 2006).

As explained by Marshall et al. (2009), there are two general types of behavioral approaches to treating deviant sexual preferences: aversion techniques (such as covert sensitization and satiation therapy) and techniques that utilize positive reinforcement (such as masturbatory reconditioning). Covert sensitization involves pairing the deviant sexual behavior with an aversive stimulus via imaginal exposure so that the offender eventually finds the sexual behavior to be aversive. In contrast, satiation therapy involves combining the deviant stimulus with prolonged masturbation so that over time the stimulus loses its appeal. Masturbatory reconditioning involves replacing the deviant sexual stimulus (e.g., child) with a pro-social stimulus (e.g., an adult) which is then reinforced through masturbation. While these sorts of techniques have been in use for many years, there is

little research available that has examined their efficacy. Also, much of the research that has been conducted has used single-case designs which of course do not allow comparison with other treatment methods.

Marshall et al. (2009) conclude that based on the available evidence, the best approach to treating deviant sexual arousal would be to use a combination of masturbatory reconditioning and satiation therapy. In this way the individual's arousal to the pro-social stimulus would be strengthened, while the arousal to the deviant stimulus would be weakened.

Emotional Regulation

Research indicates that those who commit sexual offenses tend to have difficulties regulating their emotions and may commit sexual crimes as a means of avoiding experiencing negative emotion (Mandeville-Nordon and Beech 2004). Evidence suggests that, in particular, sexual offenders often struggle to cope adaptively with stressful life events. It seems that when faced with stressful situations, sexual offenders often use maladaptive and ineffective coping strategies which increase the likelihood of reoffending (Cortoni and Marshall 2001). For example, it has been observed that rapists may experience high levels of anger and resentment in response to stress which may contribute to their offending.

Other evidence indicates that sexual offenders often use masturbation with either appropriate or inappropriate sexual fantasies as a means of escaping the experience of negative emotion (Cortoni and Marshall 2001), and such responses could increase the likelihood of sexual offending, especially, if it reinforces the idea of inappropriate sexual interaction. Thus, researchers have suggested that treatment should include a component that teaches offenders how to respond constructively to the negative emotions that may arise from stressful life events (Beech and Fisher 2002). However, to date, there is little research available that shows whether the inclusion of such components in treatment leads to a reduction in sexual reoffending. A study by Serran and colleagues (2007) found that sex offenders who had completed treatment which

included a coping skills component appeared to be better equipped (than the wait-listed controls) at dealing with high-risk situations. However, the study did not include a follow-up looking at recidivism; thus, it is unclear whether this would translate into a reduction in reoffending.

As mentioned above, in terms of rape, anger is a particularly prominent negative emotion. Howitt (2006) states: "Anger control problems have to be seen as an issue with rapists for whom issues of anger are common" (p. 369). However, some researchers have questioned the use of anger management components in treatment programs for rapists due to the fact that anger does not appear to be a problem for all rapists. For example, Loza and Loza-Fanou (1999) investigated anger levels across a range of offender groups and found no difference in levels between rapists and non-rapists (or between violent and nonviolent offenders). The authors then question the usefulness of anger management in the treatment of rapists. Arguably, this study would have been more worthwhile if it had also included a non-offender population as it may be the case that offender populations have high levels of anger problems. If so, then rapists' anger levels, though similar to other offenders, may still be problematic and potentially contribute to their risk of reoffending.

Cognition

It is widely acknowledged that sexual offenders may harbor a range of offense-related attitudes and beliefs. Cortoni (2009) states: "...an overall predisposition to be tolerant of sexual offending is related to sexual recidivism" (p. 44). Thus, the targeting of offense-related attitudes and beliefs is a mainstay of rehabilitation approaches. One common type of belief that is endorsed by sexual offenders is a belief of "entitlement" to sex. This can include specific beliefs such as "men need more sex than women do" and "everyone is entitled to sex" (Pemberton and Wakeling 2009). Also, child sex offenders may believe that children are sexual objects and that having sex with children is good because it educates them about sex. In contrast, research has shown that those who commit rape offenses tend to objectify and

sexualize women and that they hold quite specific beliefs about specific situations. For instance, rapists often endorse beliefs along the lines of “if a woman flirts, then she wants to have sex” or “if a woman accepts a free drink, then she wants to have sex.”

Within the sex offender field these problematic attitudes and beliefs are typically referred to a “cognitive distortions,” and they are frequently used by offenders to justify their offending. In this way, they may be used by the offender to feel comfortable with continued offending. For example, a child sex offender might justify his ongoing abuse of a child with the belief that the sexual interaction with the child is appropriate because the child is choosing to engage in the sexual activity and shows no sign of being unwilling. In association with his belief, he may also believe that the child is not being harmed in any way. In this way the abuser attempts to minimize his level of culpability, and this contributes to his decision to continue with the behavior. While sexual offenders will usually report that they had an awareness of wrongdoing, they will often use their cognitive distortions to distract themselves from this awareness.

As mentioned above, sexual offenders often find it difficult to admit to their offending due to the social ramifications. For example, they may believe that if they admit to what they have done, they will lose the support of family and friends. Also, in some cases, the sense of shame and embarrassment make it very difficult for them to talk to others about it, including those involved in correctional and rehabilitative services. It has been reported that most sexual offenders either deny that they committed the crime or at least attempt to minimize their level of responsibility (Marshall 2006). Thus, it might be assumed that denial and minimization would be important targets for treatment. In fact, many treatment programs do not allow deniers to participate because some components of treatment require the offender to admit to their offending. For example, often offenders are expected to describe their offense in detail in order to identify points at which they could have behaved differently and avoided offending.

There may be a range of other reasons that treatment programs may preclude those who deny their offending. For instance, deniers may have a negative impact on other individuals in the group; if they are not willing to be honest, then others may be less inclined to be upfront and open. Also, those who are unwilling to admit to and discuss their offending may be less likely to benefit from treatment; thus, they may take the place of someone who may have had more to gain from taking part in treatment. This issue arises because most sex offender treatment programs have limited places available. As they require significant resources to run, organizations (typically prisons) are only able to run a limited number. However, Marshall (2001, 2006) argues that excluding those who deny their sexual offending has the effect of putting the public at greater risk because those offenders are then released untreated.

When denial is addressed in treatment, there are two general approaches. One approach is to include a specific module that targets denial during the initial stages of the treatment program, while the other approach is to place deniers in a specially designed treatment program. When the former approach is taken, the issue of denial is usually discussed in an individual setting, prior to the beginning of treatment proper, so that the offender can consider and perhaps reconsider the extent to which they will be willing to discuss the offense, prior to having to communicate with fellow group members.

Intimacy and Social Skills

Evidence suggests that sexual offenders find it difficult to establish and maintain consensual intimate relationships and that subsequently they are vulnerable to experiencing social isolation and loneliness (Mandeville-Norden and Beech 2004). Mandeville-Norden and colleague also report that individuals who commit sexual crimes tend to lack confidence and assertiveness skills; thus, it may be hypothesized that the lack of confidence and problems being assertive contribute to the intimacy difficulties. Also, research has shown that sex offenders report significant feelings of loneliness even within the context of

intimate relationships. Therefore, social skills training is a mainstay of sex offender treatment programs. As explained by Marshall (2006), social skills training is usually a broad-based aspect of treatment which may include a range of components, including problem solving skills, assertiveness training, self-esteem building, and sex education.

The Extended Sex Offender Treatment Programme (ESOTP) in the United Kingdom is an example of a program that addresses social functioning. This program for high-risk offenders, which involves 140 h of treatment over 68 sessions, includes a component titled “intimacy skills.” This component includes skill development in the areas of jealousy management, conflict resolution, and giving and receiving social support. A recent survey of sex offender treatment programs in Canada reported that three quarters of their programs included components that targeted intimacy, relationships, and social skills. Overall, research shows that both intimate relationships and more general social functioning are addressed in most sex offender treatment programs.

Some research has examined the relationship between sexual offending and adult attachment styles. For instance, Lyn and Burton (2004) examined attachment styles in a “Midwestern United States” sample of sex offenders and found that the majority (85 %) were assessed as having insecure patterns of attachment. Furthermore, when Lyn and colleague compared sex offenders with non-sex offenders, they found that the former were significantly more likely to display signs of insecure attachment. In a further study Lyn and Burton (2005) explored the particular types of problematic attachment that sexual offenders tended to display, and they found that anxiety and avoidance were seen most frequently.

William Marshall is well known for his work on the role of attachment in sexual offending. He has long argued that attachment problems contribute to the development of sexual offending. Furthermore, he has examined some of the childhood difficulties which may be associated with attachment problems. For example, he has

reported that sexual offenders often have problematic childhoods, characterized by disruption, neglect, and abuse (both physical and sexual) (e.g., Marshall and Marshall 2000). Also, as explained by Marshall and Marshall, while such aversive early experiences may lead to attachment problems, they may also lead to a range of more general difficulties, such as low self-esteem and limited relationship skills. Thus, attachment theory offers a psychological mechanism that may explain the intimacy and relationship problems that are seen in sexual offenders.

Victim Empathy

Many, if not most, contemporary sex offender treatment programs include a component that focuses on the development of victim empathy. Research examining empathy in sexual offenders has found that while they often lack empathy for their victims, they do not typically have more generalized empathy deficits. Thus, treatment that aims to develop sexual offenders’ empathy is usually victim-focused. For instance, some treatment programs will require offenders to listen to a voice recording of a victim talking about the impact of his or her own sexual abuse experience. Alternatively, the offenders may be required to read stories of victims’ experiences or to read accounts written by their own victims.

Empathy is a complex concept and, historically, there have been a variety of approaches to defining it. While some theorists have conceptualized it as essentially involving cognitive processes, others have construed it as the ability to recognize emotion in others and to take on the perspective of another person. More recently, it seems that many theorists have viewed it as a multifaceted concept involving thoughts, emotions, and behaviors. With regard to sexual offenders, Marshall (2006) suggests that empathy deficits are frequently associated with the presence of victim-related cognitive distortions. For example, a child sex offender may believe that his sexual activity with his victim is educational and, therefore, not harmful. Or he may believe that if the child is quiet during the sexual interaction, then the child is enjoying it; thus, the offender may misinterpret the child’s emotional response.

Given the apparent connection between empathy deficits and cognitive distortions, Marshall proposes that it may be possible to address empathy deficits simply by challenging an offender's cognitive distortions, rather than directly targeting empathy problems.

Other researchers have suggested that empathy deficits in sexual offenders may be associated with difficulties with the identification and awareness of emotion (e.g., Gannon et al. 2008). Although, this seems to contrast with the suggestion that sexual offenders' empathy deficits are specific to their offending. Arguably, if they have difficulty with their own emotional awareness, then one might expect that they would have more general problems empathizing with others. One commonly used technique for the development of emotional awareness in group treatment is encouraging participants to identify and describe the emotion that they are experiencing at that moment in treatment. This is believed to facilitate the development of an ability to put a name to one's emotional experience and to describe that experience to others.

Some treatment programs utilize victim role plays as a way of enhancing offenders' empathy for their victims. A study by Webster et al. (2005) examined the effectiveness of offense reenactments and concluded that their inclusion in treatment appeared to enhance offenders' victim empathy. The approach used in the study, which was tailored to each participant, included the development of an offense map which was then used as a guide for walking the offender through the offense while he was role-playing the victim. Any sexual or violent acts were depicted symbolically, using light touch on a nonsexual area of the body. The study included a control group who carried out "empathy deficit role play scenarios" (p. 67) but did not engage in any victim role plays. While, as mentioned above, the researchers found this approach had clinical benefits, they also found that the method seemed to be more effective in treating rapists than child sex offenders and that the differences between the two approaches were small.

It is important to point out that although, as outlined above, victim empathy is a frequently

utilized component of sex offender treatment programs, research has not been able to show that it leads to a decrease in reoffending. However, demonstrating this sort of causal connection will always be a challenging task because treatment is typically multifaceted; therefore, it is difficult to know for sure which elements of treatment are bringing about change. Furthermore, given that sexual offenders present with victim-specific empathy deficits, it makes clinical sense to address this in treatment.

Substance Abuse

There is now a significant body of research that demonstrates a link between alcohol use, other drug use, and various types of criminal behavior. Further, as stated by Fridell and colleagues (2008, p. 800), "Alcohol abuse precedes or accompanies a large proportion of violent crime." Also, research has shown that problematic alcohol use is associated with a tendency to reoffend. With regard to sexual offenders, evidence suggests that a significant proportion have drug and or alcohol-related problems. For example, a study by Marshall (1996) found that 50 % of sexual offenders were intoxicated with alcohol at the time that they committed their most recent sexual offense. Also, as with general offending, drug and alcohol abuse has also been found to be associated with sexual reoffending. For instance, research suggests that alcohol abuse at least doubles the likelihood that a sexual offender will reoffend. Some researchers have observed that drug and alcohol problems are often associated with problematic lifestyle choices which lend themselves to offending behavior. Thus, drug and alcohol issues and offending may have a common origin.

While this may indeed be the case, it is important to acknowledge that intoxication with any substance has a range of consequences that can directly impact on offending. For example, alcohol can significantly diminish an individual's ability to understand and respond appropriately to a social situation. Specifically, it impairs one's judgment, thus making it more likely that an individual will respond in an antisocial manner. Furthermore, research has shown that alcohol can increase one's tendency for impulsivity, meaning

that an offender will be more likely to respond to a situation quickly and without a proper appraisal of the consequences of his behavior. Also, there are well-established links between excessive alcohol use and aggression in those who have already demonstrated aggressive tendencies.

A study by Abracen et al. (2006) compared sexual and violent (nonsexual) offenders in terms of the presence of alcohol abuse, using the Michigan Alcohol Screening Test (MAST). Results showed that sexual offenders had significantly higher scores thereby indicating that they were more likely to have problematic patterns of drinking. In their discussion of the results, they suggest that myopia theory may offer an explanation of why alcohol use may contribute to sexual offending. Myopia theory suggests that when an individual is considering engaging in a risky behavior, alcohol may have the effect of decreasing their inhibition via a reduction in their information processing capacity. This is similar to effect of alcohol use on judgment which is suggested above.

Evidence suggests that there are a variety of approaches to addressing drug and alcohol problems. While some programs include a substance abuse component, others do not address it directly but direct individuals to appropriate drug and alcohol programs. Often drug and alcohol problems are construed as responsivity barriers and are thus seen as being best addressed prior to offense-specific treatment. For example, even in many prison environments, drugs are available and, therefore, it is considered that in order for an offender to make the most of offense-specific treatment, he should first address his drug use. Obviously, in community settings drug and alcohol use is more prevalent; thus, the same approach is often taken.

Treatment Efficacy

There is now a growing body of research that has examined the efficacy of sex offender treatment programs. A study a decade ago (Hanson et al. 2002) examined the outcome evaluations of 43 sex offender treatment programs with a mean

follow-up period of 76 months. Hanson et al. reported a sexual recidivism rate of 12.3 % for sex offenders who had completed treatment and a 16.8 % recidivism rate for untreated sex offenders. They also found that the treated sex offenders committed significantly fewer general offenses than their untreated counterparts (27.9 % and 39.2 %, respectively). Note that the majority of treatment programs that were included in this study had a cognitive-behavioral orientation; thus, this lends support to the efficacy of cognitive-behavioral programs.

Similarly, Lösel and Schmucker (2005) carried out a meta-analysis of 69 studies of sex offender treatment efficacy that contained 80 separate comparisons of treated and untreated offenders. The authors concluded that treatment led to a mean reduction in sexual recidivism of almost 37 % (when low base rates are taken into account). The actual difference in recidivism rates between treatment groups and control groups (who were untreated or had completed another type of treatment) was 11.1 % and 17.5 %, respectively. They also found that sex offender treatment led to a significant reduction in general reoffending. In summarizing their findings, Lösel and colleague state: "The most important message is an overall positive and significant effect of sex offender treatment" (p. 135).

Another study (Seager et al. 2004) looked at recidivism rates in 109 sex offenders who had completed offense-specific cognitive-behavioral treatment and 37 who had not completed any treatment for their sexual offending. Note that of the 109, only 81 were assessed as having successfully completed the program. The recidivism rates that were examined included sexual and violent offenses. Results showed that regardless of whether individuals were considered to have been successful in completing the program, those who completed were found to have a lower recidivism rate. Specifically, 4 % of successful completers and 7 % of unsuccessful completers were reconvicted within the 2-year follow-up period. This contrasted with 18 %, 42 %, and 100 % (respectively) of those who withdrew from treatment, declined it, or were terminated from the program. However, the authors

concluded that there was no evidence of a connection between offenders' change in clinical phenomena (such as empathy) and their reoffending. For example, they found that the participants risk scores on the Static 99 (an actuarial risk measure) predicted their likelihood of reoffending regardless of whether they completed. Seager and colleagues concluded that "participation in the sex offender program did not reduce recidivism rates for those who complied with treatment but merely enabled motivated offenders to concretely demonstrate their commitment to not reoffend" (p. 609).

It is unclear whether the two aforementioned meta-analyses took into consideration the various ways in which treatment completers and those who did not complete treatment might have differed. The study by Seager et al. suggests that apparent treatment effects may simply be a result of preexisting differences between those who undertake such programs and those who do not. However, it is important to note that Seager's study was comparatively small and only included a small number of offenders. Nonetheless, it does raise questions about the way in which studies of treatment efficacy are conducted and how results are interpreted.

It seems reasonable to conclude that sex offender treatment is often effective and that at the very least it does no harm. Furthermore, there is growing evidence that cognitive-behavioral approaches may be especially useful. However, further studies are needed to determine if the effects that are being found are indeed due to treatment or to the fact that particular individuals tend to undertake and complete treatment.

Summary

This entry has discussed two of the key models that underlie sex offender treatment. It has also examined some of the relevant process issues, such as whether treatment should be delivered in group settings. Further, it looked at the components that are typically found in sex offender treatment programs. The entry concluded with a brief discussion of treatment efficacy. It is

hoped that this overview will provide the reader with some understanding of how the treatment of sexual offenders is ordinarily delivered and of some of the issues surrounding such treatment. Obviously, there is a need for ongoing research in many areas as many findings remain preliminary.

Related Entries

- ▶ [Actualizing Risk-Need-Responsivity](#)
- ▶ [Good Lives Model](#)
- ▶ [Sexual Recidivism](#)
- ▶ [Specialization and Sexual Offending](#)
- ▶ [Treatment of Sex Offenders](#)

Recommended Reading and References

- Abracen J, Looman J, Di Fazio R, Kelly T, Stirpe T (2006) Patterns of attachment and alcohol abuse in sexual and violent non-sexual offenders. *J Sex Aggress* 12:19–30. doi:10.1080/13552600600722963
- Andrews DA, Bonta J (2003a) *The psychology of criminal conduct*, 3rd edn. Anderson, Cincinnati
- Andrews DA, Bonta J (2003b) A commentary on Ward and Stewart's model of human needs. *Psychol Crime Law* 9:215–218. doi:10.1080/10683/16031000112115
- Beech AR, Fisher DD (2002) The rehabilitation of sex offenders. *Aust Psychol* 37:206–214. doi:10.1080/00050060210001706886
- Burdon WM, Gallagher CA (2002) Coercion and sex offenders: controlling sex-offending behavior through incapacitation and treatment. *Crim Justice Behav* 29:87–109. doi:10.1177/0093854802029001006
- Cortini F (2009) Factors associated with sexual recidivism. In: Beech AR, Craig LA, Browne KD (eds) *Assessment and treatment of sex offenders: a handbook*. Wiley-Blackwell, Chichester, pp 39–52
- Cortoni F, Marshall WL (2001) Sex as a coping strategy and its relationship to juvenile sexual history and intimacy in sexual offenders. *Sex Abuse J Res Treat* 13:27–43. doi:10.1023/A:1009562312658
- Fridell M, Hesse M, Jaeger MM, Kühllhorn E (2008) Anti-social personality disorder as a predictor of criminal behaviour in a longitudinal study of a cohort of abusers of several classes of drugs: relation to type of substance and type of crime. *Addict Behav* 33:799–811. <http://org/10.1016/j.addbeh.2008.01.001>
- Gannon TA, Collie RM, Ward T, Thakker J (2008) Rape: psychopathology, theory, and treatment. *Clin Psychol Rev* 28:982–1008. doi:10.1016/j.cpr.2008.02.005
- Hanson RK, Gordon A, Harris AJR, Marques JK, Murphy W, Quinsey V, Seto M (2002) First report of the collaborative outcome data project on the

- effectiveness of psychological treatment for sexual offenders. *Sex Abus J Res Treat* 14:169–194. doi:10.1023/A:1014624315814
- Harkins L, Beech AR (2008) Examining the impact of mixing child molesters and rapists in group-based treatment for sexual offenders. *Int J Offender Ther Comp Criminol* 52:31–45. doi:10.1177/0306624X07300267
- Howitt D (2006) Introduction to forensic and criminal psychology. Pearson Education, Harlow
- Langevin R (2006) Acceptance and completion of treatment among sex offenders. *Int J Offender Ther Comp Criminol* 50:402–417. doi:10.1177/0306624X06286870
- Lösel F, Schmucker M (2005) The effectiveness of treatment for sexual offenders: a comprehensive meta-analysis. *J Exp Criminol* 1:117–146. doi:10.1007/s11292-004-6466-7
- Loza W, Loza-Fanous A (1999) The fallacy of reducing rape and violent recidivism by treating anger. *Int J Offender Ther Comp Criminol* 43:492–502. doi:10.1177/0306624X99434007
- Lyn TS, Burton DL (2004) Adult attachment and sexual offender status. *Am J Orthopsychiatry* 74:150–159. doi:10.1037/0002-9432.74.2.150
- Lyn TS, Burton DL (2005) Attachment, anger, and anxiety of male sexual offenders. *J Sex Aggress* 11:127–137. doi:10.1080/13552600500063682
- Mandeville-Nordon R, Beech A (2004) Community-based treatment of sex offenders. *J Sex Aggress* 10:193–214. doi:10.1080/1355260042000261760
- Marques JK, Wiederanders M, Day DM, Nelson C, van Ommeren A (2005) Effects of a relapse prevention program on sexual recidivism+. Final results from California's sex offender treatment and evaluation project (SOTEP). *Sex Abus J Res Treat* 17:79–107. doi:10.1007/s11194-005-1212-x
- Marshall WL (1996) Assessment, treatment, and theorizing about sex offenders: developments during the last 20 years and future directions. *Crim Justice Behav* 23:162–199. doi:10.1177/0093854896023001011
- Marshall WL (2001) Adult sexual offenders against women. In: Holin CR (ed) *The handbook of offender assessment and treatment*. Wiley, Chichester, pp 333–348
- Marshall WL (2006) Diagnosis and treatment of sexual offenders. In: Weiner IB, Hess AK (eds) *The handbook of forensic psychology*, 3rd edn. Wiley, Hoboken, pp 790–818
- Marshall WL, Marshall LE (2000) The origins of sexual offending. *Trauma Violence Abuse* 1:250–263. doi:10.1177/152483800001003003
- Marshall WL, Serran GA, Fernandez YM, Mulloy R, Mann RE, Thornton D (2003) Therapist characteristics in the treatment of sexual offenders: tentative data on their relationship with indices of behaviour change. *J Sex Aggress* 9:25–30. doi:10.1080/1355260031000137940
- Marshall WL, O'Brien MD, Marshall LE (2009) Modifying sexual preferences. In: Beech AR, Craig LA, Browne KD (eds) *Assessment and treatment of sex offenders: a handbook*. Wiley, Chichester, pp 311–327
- Maruna S (2001) *Making good: how ex-convicts reform and rebuild their lives*. American Psychological Association, Washington, DC
- Marx BP, Miranda R, Meyerson LA (1999) Cognitive-behavioral treatment for rapists: can we do better? *Clin Psychol Rev* 19:875–894. doi:10.1016/S0272-7358
- Milner RJ, Wakeling HC, Mann RE, Webster SD (in press) Clinical impact of a socio-affective functioning programme for high risk sexual offenders
- New Zealand Department of Corrections (2009) What works now? A review and update of research evidence relevant to offender rehabilitation practices within the Department of Corrections. <http://www.corrections.govt.nz/research>
- Pemberton AE, Wakeling HC (2009) Entitled to sex: attitudes of sexual offenders. *J Sex Aggress* 15:289–303. doi:10.1080/13552600903097212
- Seager JA, Jellicoe D, Dhaliwal GK (2004) Refusers, dropouts, and completers: measuring sex offender treatment efficacy. *Int J Comp Criminol* 48:600–612. doi:10.1177/0306624X04263885
- Serran GA, Fernandez Y, Marshall WL, Mann R (2003) Process issues in treatment: application to sexual offender programs. *Prof Psychol Res Pract* 34:368–374. doi:10.1037/0735-7028.34.4.368
- Serran GA, Moulden H, Firestone P, Marshall WL (2007) Changes in coping following treatment for child molesters. *J Interpers Violence* 22:1199–1210. doi:10.1177/0886260507303733
- Ward T, Stewart CA (2003) The treatment of sexual offenders: risk management and good lives. *Prof Psychol Res Pract* 34:353–360. doi:10.1037/0735-7028.34.4.353
- Ward T, Collie RM, Bourke P (2009) Models of offender rehabilitation: the good lives model and the risk-need-responsivity model. In: Beech AR, Craig LA, Browne KD (eds) *Assessment and treatment of sex offenders: a handbook*. Wiley, Chichester, pp 293–310
- Ward T, Gannon TA, Yates PM (2008) The treatment of offenders: current practice and new developments with an emphasis on sex offenders. *Int Rev Victimol* 15:179–204. doi:10.1177/026975800801500207
- Webster SD, Bowers LE, Mann RE, Marshall WL (2005) Developing empathy in sexual offenders: the value of offence re-enactments. *Sex Abus J Res Treat* 17:63–77. doi:10.1007/s11194-005-1211-y
- Wilson R, Yates P (2009) Effective interventions and the good lives model: maximising treatment gains for sexual offenders. *Aggress Violent Behav* 14:157–161. doi:10.1016/j.avb.2009.01.007
- Wong YJ (2006) Strength-centred therapy: a social constructionist, virtues-based psychotherapy. *Psychother Theory Res Pract Train* 2:133–146. doi:10.1037/0033-3204.43.2.133
- Yates PM, Goguen BC (2000) National sex offender treatment programme description. Correctional Service of Canada, Ottawa

Sex Offenders and Criminal Policy

Richard Tewksbury and David P. Connor
Justice Administration, University of Louisville,
Louisville, KY, USA

Overview

Over the past several years, individuals convicted of criminal acts have been subjected to increasingly severe sanctions. In particular, efforts have been made to increase the supervision of criminals and decrease offender opportunities to further perpetrate crimes. Without a doubt, many emerging laws and criminal justice policies have targeted sex offenders, especially those convicted of victimizing children. In addition to progressively long incarcerations, sex offenders potentially face civil commitment, registration and community notification, chemical castration, polygraph testing, and residency restrictions. The United States has held a leading position in passing these laws and enacting policies aimed at preventing future sex crimes.

Explanations for the renewed interest in punitive sex offender legislation may be found in the high-profile sexual attacks on children in the 1990s. Child sexual assault victims whose names saturated media headlines and generated castigating policies toward sex offenders included Jacob Wetterling, Megan Kanka, and Adam Walsh. Although not new concepts, the revival of criminal registration and civil commitment for sex offenders reflect growing public and political concerns. Technological innovations may also be responsible for intensified sex offender legislation. Chemical castration laws and polygraph testing, for example, are likely the result of expanding scientific know-how.

Background Description

Civil Commitment Statutes

In response to high-profile sex crimes, the first American sexual psychopath laws were passed in

the 1930s. Community protection, as well as treatment and incapacitation of sex offenders, were the explicit goals of these laws. Often calling for the civil commitment of so-called mentally disordered sex offenders to public mental hospitals, these statutes rested on the assumption that mental health professionals were capable of identifying, confining, and treating sexual psychopaths.

Criticism surrounding sexual psychopath laws has always existed. Those offenders determined to be sexual psychopaths were often committed to mental institutions indefinitely with few procedural safeguards. These laws may have also been aimed at sex offenders other than violent recidivists and included exhibitionists, voyeurs, and homosexuals. Moreover, public disenchantment with rehabilitation generally, and lack of confidence in effective treatment methods for sexual psychopaths specifically, eventually dissuaded efforts to utilize these laws. By the late 1960s, many states started to repeal, intentionally disregard, and seldom employ these sexual psychopath laws.

Registration and Community Notification

Registration of individuals charged or convicted of various crimes has been used for decades. In 1994, the Jacob Wetterling Act put into practice the registration of sex offenders in statewide databases. Culpable for transforming sex offender registries into publicly available online domains, Megan's Law, passed in 1996, requires state law enforcement agencies to make public information about sex offenders. Affording a sense of safety and control to the public, such laws have subsequently also been passed unanimously by many legislatures. Today, all states, as well as the federal government, have enacted and maintain publicly available, Internet-based sex offender registries.

Castration Laws

For many centuries, sex offenders were punished through physical castration. In the United States, castration was first popularly used for slaves suspected of having sex with white women. The eugenics movement (1905–1935) supported

castration and sterilization of criminals and the mentally ill. Currently, voluntary chemical castration (combined with therapy) has largely replaced physical castration, as medical doctors believe that similarly effective results can be reached through treatment with medication.

Concerns over body mutilation, intrusiveness, and the lasting results of surgery have rendered physical castration of sex offenders in the United States as largely unacceptable. The widespread availability of medications as an alternative to physical castration may also explain this change in practice. Although some laws concerning sex offenders allow both physical and chemical castration as a substitute or adjunct to punishment, chemical castration seems to be the more socially acceptable solution.

Polygraph Testing

Over the last two decades, polygraph testing has been presented as yet another tool to manage, supervise, and treat sex offenders under community supervision. Numerous states require individuals convicted of two sex offenses to undergo periodic mandatory polygraph supervision. Until recently, *Frye v. United States* (1923) governed the admissibility of polygraph results in American courts. In this case, the Court of Appeals of District of Columbia ruled that when questions of fact require special knowledge, the opinions of witnesses skilled in the subject to which the questions relate are admissible in evidence.

In 1993, however, the US Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* ruled that the Federal Rules of Evidence should provide the standard for admitting expert scientific testimony. To determine the admissibility of polygraph evidence, trial courts must now consider factors such as whether lie detector testing has been scientifically tested, largely accepted within the scientific community, and exposed to review and publication. As a result, most states currently do not consider polygraph evidence as admissible; however, if both parties stipulate, a few courts may admit polygraph results.

Appellate courts in *Himes v. Thompson* (2000) and *Kansas v. Lumley* (1999) have ruled that polygraph evidence is admissible, as well as

sufficiently reliable as evidence, at probation and parole revocation hearings. Courts in *State v. Flores-Moreno* (1994) and *State v. Riles* (1997) have also held that polygraph testing limited to topics related to crimes perpetrated by sex offenders is permissible as a condition of community supervision. Moreover, the Washington Court of Appeals in *State v. Eaton* (1996) expressed that polygraph testing as a requirement for probationers and parolees was a necessary and effective way in which to monitor compliance with terms of supervision.

Residency Restrictions

Residency restriction laws have been created as a result of increasing concern about registered sex offenders living in communities. After the murder of Jessica Lunsford in Florida by a previously convicted sex offender, housing restrictions for sex offenders in the United States became more widespread. Despite punishment, treatment, and supervision, many believe that sex offenders continue to pose a serious danger to society. These housing restrictions prevent sex offenders from living near schools, parks, daycare centers, and other designated “child congregation” locations. In this way, these laws seek to limit contact between registered sex offenders and children, and subsequently reduce sex offender recidivism.

State of the Art

Civil Commitment Statutes

The premise of civil commitment statutes is that many sex offenders cannot be rehabilitated; consequently, such criminals must be incapacitated to the greatest extent possible. Under civil law, sex offenders may be committed to institutions. Modern civil commitment statutes, commonly known as sexually violent predator laws, allow governments to confine particular sex offenders to secure mental health facilities upon their release from prison or a judicial finding of incompetency to stand trial. Upon the decision of a court, sex offenders considered to have mental abnormalities or other psychological disorders that may prompt harmful sexual conduct in the

future may be subjected to civil commitment. As of 2009, 20 states and the federal government have civil commitment laws aimed at sex offenders.

Following the placement of committed sex offenders to secure institutions, mental health clinicians and other professionals are assigned the responsibility of evaluating offenders at regular intervals. During periods of civil commitment, medical and psychological assessments of sex offenders are performed to assess improvements in mental status. Examinations and subsequent reports prepared for the court concerning the mental conditions of committed sex offenders are typically completed on an annual basis. After these examinations by medical and mental health officials, sex offenders civilly committed to institutions may petition for release, appear before a judge, and ask the court to determine whether their commitments continue to be necessary to protect the public. In order to be released, psychologists and the court must agree that mental abnormalities or personality disorders once exhibited by sex offenders no longer pose a threat to society.

Whereas earlier sexual psychopath laws may have centered on rehabilitation, current civil commitment laws focus on the social control and incapacitation of sex offenders. Depicting the social control function of civil commitment laws, the US Supreme Court in *Kansas v. Hendricks* (1997) upheld a sexually violent predator statute, which sent sex offenders determined to be violent predators likely to recidivate to state mental hospitals. Although Hendricks challenged the civil commitment law as a violation of due process and prohibitions against double jeopardy and (*ex post facto*) laws, the Court found such to be constitutional as the civil commitment statute was not deemed to be punishment. In 2002, the Supreme Court ruled in *Kansas v. Crane* that mental abnormalities displayed by sex offenders must differentiate committed individuals from ordinary recidivists. Further, besides showing the likelihood that offenders will commit sex crimes upon release, it must be shown that offenders have a serious inability to control their behaviors.

Registration and Community Notification

Sex offenders have consistently faced stringent sentencing laws; moreover, society has consistently looked upon them with disdain. Spotlighting society's harsh treatment of sex offenders, one of the most recent developments has been the creation of sex offender registries. Sex offender registries are utilized in every jurisdiction in the United States, and these repositories of information provide online access to a wide array of facts about convicted sex offenders and their offenses. Individuals convicted of sex crimes are typically required to provide local law enforcement and corrections authorities with name, photograph, address, birth date, Social Security number, fingerprints, offense history, date of convictions, and other information. Sex offenders must verify the accuracy of this information on a routine basis for the duration of their registration, which may range from 10 years to life.

Most arguments supporting sex offender registries emphasize public safety, particularly the protection of children. Proponents also contend that registration will permit law enforcement officials to quickly and easily ascertain the locations of sex offenders in their communities, facilitating sex crime investigations. Because Internet databases reveal the identification of sex offenders to the community, it is further maintained that sex offender registries reduce opportunities for recidivism. Despite evidence suggesting little or no effect of registration and community notification on recidivism rates, there remains general public approval concerning sex offender registries and a belief that most sex offenders will reoffend.

Community notification typically occurs in neighborhood meetings, door-to-door visits by the police, newspaper advertisements, online notices, and flyers circulated throughout a jurisdiction. Many states that use community notification have a three-tiered system based on the purported dangerousness of sex offenders that determines the degree of notification that will take place. When sex offenders are categorized as the lowest risk to public safety, notification is typically reserved for law enforcement officials only. Schools, daycares, and other neighborhood organizations are notified of the presence of sex

offenders posing a medium risk to public safety. Those sex offenders considered the most dangerous, designated at high risk, will generate the most widespread notification, as the general public is notified. Some states use formal assessment tools to classify sex offender risk levels. Other jurisdictions utilize committees of clinicians and other professionals to evaluate offenders and decide upon risk level.

Four major laws are associated with contemporary sex offender registration and notification policies. The Jacob Wetterling Crimes Against Children and Sexually Violent Offender Registration Act (1994) is the first federal law that mandated statewide registration of sex offenders. This statute requires all states to establish procedures for sex offender registration. If particular jurisdictions fail to comply with the Wetterling Act, they risk the loss of crime prevention funding from the federal government. As a result of the Wetterling Act, each state has a mandatory registration law that obligates sex offenders to provide their information to law enforcement officials.

Despite the fact that Washington enacted the first community notification law in 1990, it was Megan's Law (New Jersey Stat. Ann. §§ 2C: 7-1 to 2C: 7-11, 1996) in New Jersey that created registration and community notification legislation that was ultimately replicated nationwide. Megan Kanka was sexually assaulted and murdered by a twice-convicted sex offender on parole. Amending the Jacob Wetterling Act, Megan's Law calls for the registration of sex offenders with law enforcement at various time intervals. The frequency of registration is contingent upon assumed risk levels of sex offenders. In many states, Megan's Law also requires the local police to inform communities of the presence of sex offenders moving into neighborhoods. The jurisdiction in which sex offenders reside, as well as their risk status, will determine the degree of information that is made available to the public.

In 2000, the Campus Sex Crimes Prevention Act further amended the Wetterling Act, obligating registered sex offenders studying and working at colleges and universities to provide notice

of their status as sex offenders to these institutions of higher learning. College and university officials are required to inform the campus community where information regarding registered sex offenders may be obtained. In fact, many colleges and universities maintain their own distinctive online sex offender registry. At the very least, these registries provide the full names of students and university employees convicted of sex offenses. More often, campus sex offender registries include demographic information, including physical descriptions and dates of birth. Some university registries also contain the offender's home address, conviction information, description of victims, and special conditions imposed at sentencing. Further, in addition to the academic institutions with campus-based registries, many institutions of higher learning provide online links to the state sex offender registry from the webpages of campus police or the public safety department. This information is also provided to law enforcement agencies in the jurisdiction of the college or university, and it is entered into state registry records.

The abduction of Polly Klaas generated the development of three-strikes laws. Polly's murder by a repeat sex offender who was not required to register infuriated the public and generated intense media coverage. This resulted in the overwhelming public support for Proposition 184 in California. Passed in 1994, this legislation allowed a third felony conviction to result in a sentence of 25 years to life. Although those championing three-strikes laws contend that such legislation serves as a general deterrent to potential three-time offenders, critics note that there is little evidence that shows reduction in serious crime has resulted from implementation of three-strikes laws.

Accessible to the public through the Internet, a nationwide databank of registered sex offenders was created in 2005. The Dru Sjodin National Sex Offender Public Registry, as it was named in 2006, was designed by the Department of Justice to provide more efficient access to individual state sex offender registries.

The Adam Walsh Child Protection and Safety Act was signed into law in 2006. This federal law

includes the Sex Offender Registration and Notification Act (SORNA) and seeks to make the archiving and monitoring of sex offenders more efficient by creating a comprehensive and national system for sex offender registration. The legislation categorizes sex offenders into three tiers of risk. Tier three sex offenders are the most serious and must verify their location with law enforcement every 3 months for the rest of their lives. Individuals defined as tier two offenders must update their whereabouts every 6 months for 25 years, while sex offenders described as tier one lawbreakers must verify their location annually for 15 years.

The Adam Walsh Act includes additional provisions for those convicted of sex crimes. Failing to register and update information with authorities are increased to felony offenses. Failure to adhere to registration guidelines under SORNA is punishable by 10 years in prison and a \$250,000 fine. The federal law also eliminates the statute of limitations concerning the prosecution of child abduction and felony child sexual offense cases. It allows victims of child abuse to utilize civil proceedings to seek monetary damages from perpetrators. Further, the Adam Walsh Act sets forth mandatory minimum sentencing for particular sex crimes. For instance, the law dictates a minimum sentence of 30 years for the rape of a child and calls for increasing the minimum prison terms for offenders traveling between states with minors.

All jurisdictions must comply with the Adam Walsh Act and the provisions of SORNA or face reduced federal grant funding. In September 2009, Ohio was the first state to comply with federal sex offender provisions under SORNA. However, as of 2011, 43 states have yet to achieve compliance with the mandates of the Adam Walsh Act.

The constitutionality of sex offender registration and community notification has been challenged in two US Supreme Court cases. *Smith v. Doe* (2003) considered the Fifth Amendment double-jeopardy clause and whether registration and notification laws amount to further punishment of sex offenders. The Supreme Court found that sex offender registration and notification is

regulatory and consequently is not additional punishment. Moreover, the Court ruled that individuals convicted of sex crimes prior to the establishment of registration and notification for sex offenders can still be obligated to comply with these laws. The second Supreme Court case regarding sex offender registration and community notification examined the issue of cruel and unusual punishment. The Court in *Connecticut Department of Public Safety v. John Doe* (2003) ruled that the posting of sex offender photographs online is constitutional. Registration and notification laws, therefore, are not considered to be obstacles to personal freedom.

Castration Laws

In the last two decades, medication as a means of controlling the behavior of sex offenders has gained renewed interest. Chemical castration consists of injecting Depo-Provera, the synthetic hormone medroxyprogesterone acetate, to reduce the blood serum testosterone levels in males. Taken on a regular basis, Depo-Provera is supposed to reduce sexual impulses, erections, and ejaculations. The potential side effects of Depo-Provera include migraines, nausea, weight gain, insomnia, fatigue, and loss of body hair, although most of these are thought to be rare and reversible.

Presently, many chemical castration laws call for the forced dispensing of medication to control the behavior of recidivist sex offenders. In 1997, California was the first jurisdiction to require chemical castration for repeat sex offenders (with victims under the age of 13). Under California law, repeat sex offenders are obligated to receive chemical injections prior to their release on parole. These injections persist until offenders complete their criminal sentences. Although sex offenders may refuse chemical castration, parole will be immediately denied to those deciding not to participate. Following California's example, other states including Colorado, Florida, Georgia, Louisiana, Montana, Texas, and Wisconsin have similar sex offender castration laws. However, unlike California, many state statutes allow judicial discretion in regards to implementation of castration.

Chemical castration laws regarding sex offenders generally do not require complete medical or psychiatric assessments prior to the injections of medication. Further, there is no distinction between types of sex offenders most appropriate for chemical castration in the statutes.

Polygraph Testing

Polygraph (or “lie detector”) testing with sex offenders focus on disclosure polygraphs performed after sentencing, denial and specific issues examinations, and maintenance polygraphs administered during sex offender treatment for purposes of reducing offender denial and gauging compliance with conditions of supervision. Following sentencing, the disclosure polygraph involves questions asked of sex offenders about their history of sexual deviance. Polygraph examiners identify offender deception through responses of intensified physiological arousal to relevant questions. If there are discrepancies among crime descriptions, the denial polygraph examination will likely be administered. Also, when sex offenders refuse to accept responsibility for their crimes, probation and parole officials may perform the denial examination. Specific issues polygraph examinations usually center on a specific accusation or suspicion. The maintenance polygraph detects how sex offenders manage inappropriate thoughts and measures compliance with conditions of probation or parole. In total, these applications of polygraph technology serve to manage, supervise, and treat sex offenders living in the community.

Residency Restrictions

Well over one-half of all states and numerous municipalities have sex offender residency restriction laws. Residency restriction laws often feature nebulous language to restrict registered sex offenders from living near locations described as “child congregation” areas. Such places are typically defined to include schools, parks, school playgrounds, daycare centers, bus stops, and recreational facilities. Fluctuating between 500 ft and 2,500 ft, residency restriction laws assert that specific distances must be

preserved between a sex offender’s residence and various landmarks in the community.

Restrictive zoning laws prohibiting sex offenders from residing near places frequented by children are often associated with registration statutes and typically involve all registered sex offenders without regard to victims’ ages. However, other residency restriction laws concern only sex offenders under community supervision or those with child victims. For instance, the first state law addressing sex offender residency restrictions, which was passed in 1995, applied only to Florida sex offenders on probation with child victims.

In 2005, after a registered sex offender confessed to sexually assaulting and killing 9-year-old Jessica Lunsford in Florida, sex offender housing restrictions became more intensified nationwide. For example, California voters overwhelmingly passed Proposition 83 in 2006. Among other initiatives, Proposition 83 prevents sex offenders from living 2,000 ft from child congregation locations. Further, municipal sex offender housing ordinances became more prevalent. Enacted in December 2005, a Dyersville, Iowa, ordinance is among the most extreme of residency restrictions for sex offenders. The law essentially prohibits any sex offender from living anywhere in the city.

Possible Controversies in the Literature

Civil Commitment Statutes

Research highlights many disputes surrounding the civil commitment of sex offenders. Critics assert that in reality sex offenders committed to mental facilities are almost never released. Washington and Minnesota, the two states with the longest contemporary commitment programs, never discharged committed sex offenders. Others oppose civil commitment for sex offenders because such laws apparently punish individuals who have already paid their debt to society. Despite the Supreme Court decision in *Kansas v. Hendricks* (1997), Friedland (1999) contends that the civil confinement seems to

hinge on the desire to continue punishment and incapacitation.

Besides apprehensions relating to sex offenders themselves, civil commitment statutes have also raised issues concerning mental health professionals. Wettstein (1992) argues that civil commitment laws regarding sex offenders inappropriately utilize experts in the mental health field. Although many sex offenders are committed due to their mental deficiencies or personality disorders, he notes that this does not necessarily mean committed sex offenders are genuinely mentally ill.

In addition, the increasing number of sex offenders and the need to commit them may exhaust government money allotted for the entire mental health field (Friedland 1999; Wettstein 1992) and may not be the most effective use of mental health resources already in short supply. If mental health professionals are trusted to confine sex offenders, it is argued that funding for treatment and other mental health services for committed sex offenders and other patients will become even scarcer. When appropriate mental health treatment is absent, some scholars note that it is possible that many mental health patients likely to be responsive to treatment may gravitate toward criminal activities.

Registration and Community Notification

Controversies over sex offender registration and community notification also exist. Those supporting sex offender registration and community notification often contend that these laws are not punishment; instead, any punitive actions resulting from registration and notification are only related to the public protection function (Brooks 1995). Further, as a result of registration and community notification, there is evidence that probation and parole officials and the community are working together to actively monitor sex offenders (Zevitz and Farkas 2000b).

And yet, it has been found that notification laws associated with sex offender registration negatively impacted probation and parole officers (Zevitz and Farkas 2000a). In particular, officers assigned to sex offender caseloads described a loss of staff, time, and financial resources as

a result of a new community notification program. In addition, research on sex offender registration has shown unintended outcomes for the general community. Scholars have noted that sex offender registration and notification “can have the effect of leaving neighborhood residents frightened but feeling powerless to do anything about it” (Zevitz and Farkas 2000b, p. 405). Although registration laws were meant to assist the community by providing information about sex criminals, these requirements have also increased a sense of alarm among residents attending community notification meetings.

Registration and community notification may prevent successful reintegration into the community. Mental health professionals have expressed concern that reactions from the community resulting from registration and notification may intensify the anxiety of sex offenders, ultimately generating recidivism through poor choices (Billings and Bulges 2000).

Available literature concerning sex offender registration and community notification shows that registered sex offenders often experience numerous deleterious consequences. Specific offender aftermath associated with registration and community notification includes loss of family contact, destruction of friendships, employment difficulties, hostile confrontations, and threats of violence and personal harm. Another consequence of sex offender registration and community notification is the stigma that is associated with labeling as a registered sex offender (Tewksbury and Lees 2006, 2007). Making reintegration more challenging for sex offenders, powerful and enduring stigmas that result from public disclosure are strong obstacles to employment, education, and community activity (Uggen et al. 2004). Registered sex offenders in numerous studies have reported these marks of disgrace as common experiences (Tewksbury 2004, 2005; Tewksbury and Lees 2006). The most commonly reported collateral consequences for registered sex offenders are feelings of vulnerability, stigmatization, and housing difficulties (Tewksbury 2004, 2005; Tewksbury and Lees 2006). Tewksbury and Mustaine (2009) explain that these experiences are mutually

influential, as feelings of vulnerability may intensify as registered sex offenders find themselves unable to locate housing.

Castration Laws

Debates continue over the effectiveness of sex offender castration laws. Research has suggested that sex offenders treated with both medication and counseling may have improved control of their sexual behaviors (Melella et al. 1989), and sex offenders under chemical castration and participating in therapy are less likely to recidivate than those individuals discontinuing medication (Meyer et al. 1992). Supporters of castration laws regarding also contend that chemical castration does not violate the fundamental right of individuals to procreate. The effects of Depo-Provera are temporary and completely revocable upon termination of the injections (Melella et al. 1989).

However, despite Depo-Provera's only temporary obstruction to sexual activity, Spalding (1998) suggests that the ability to procreate is still infringed upon during the period of treatment, which may last for years. Further, some chemical castration laws, such as Florida's statute, permit the courts to order injections of Depo-Provera for the lifetime of the offender. In this way, sex offenders sentenced to lifetime chemical castration may have their right to procreate removed permanently. According to Meyer and Cole (1997), the long-term effects of Depo-Provera on sexual functioning are still very much in question.

Some opponents argue that judges may not be the most appropriate individuals to impose sentences of chemical castration and doubt the capabilities of medical doctors working in corrections to make impartial decisions concerning the well-being of offenders. Interestingly, Fitzgerald (1990) suggests that injections of Depo-Provera will not have any meaningful influence on many sex offenders. These sex offenders include those that deny committing the offense or refuse to acknowledge the criminal nature of their actions, those blaming their actions on environmental factors, and violent individuals motivated by nonsexual elements. Moreover, the use of medications for castration

purposes is widely considered effective only when employed in combination with psychotherapy specifically designed for sex offenders (Meyer and Cole 1997). Other research warns that the administrative procedures for involuntary medication outlined in *Washington v. Harper* (1990) may not protect inmates from uninformed and arbitrary decisions by prison psychiatrists. Ryan argues that the objectivity of a committee consisting of correctional employees and the cooperative nature of the hearing afford little protection from ill-advised psychiatric diagnosis.

Polygraph Testing

There are disputes in the literature surrounding sex offender polygraph examinations. Polygraph testing has been reported to improve the management of sex offenders in the community, as it purportedly reduces denial, extracts confessions of sexual offenses, and improves treatment results (Wilcox 2000). On the other hand, Iacono and Lykken (1997) raised questions concerning the possibility of false positives and negatives, in spite of the use of control questions present in polygraph testing. They also reported no evidence for the validity of polygraph results in their review of field studies. Further, the fear of being incorrectly accused causing innocent subjects to have physiological responses that indicate guilt has been considered as a flaw of the polygraph. Proponents of polygraph testing have described accuracy rates beyond 90 %; however, critics (Iacono and Lykken 1997) contend that these studies contain fatal methodological flaws.

Residency Restrictions

Research dedicated to sex offender residency restrictions has only emerged in the past decade. The available research points to controversies in the literature. Finding that nearly half of sex offenders with child victims lived in close proximity to child congregation zones, Walker et al. (2001) studied one Arkansas county and suggested that child sex offenders may be more likely to intentionally reside near schools, parks, and daycares. And yet, when compared with nonrecidivists, research has reported that recidivating child sex offenders under

community supervision did not live closer to these congregation zones (Colorado Department of Public Safety 2004). Tewksbury and Mustaine (2006) also showed that only approximately one in five registered sex offenders live in close proximity to such locations. Further, a Minnesota study found that sex offenders were more likely to seek victims in neighborhoods other than their own to avoid detection (Minnesota Department of Corrections 2003).

Research has recognized the problematic nature of residency restrictions aimed at sex offenders. Lawmakers believe that residency in close proximity to potential (child) victims influences recidivism (Sample and Kadleck 2008). However, research generally shows such restrictions have little or no effect on sex offender recidivism. And yet, laws limiting housing opportunities for sex offenders potentially exacerbate the limited choices for home placement already facing many ex-convicts. Studies concerning the limitations of where registered sex offenders may live have been widely shown to impose negative consequences on both offenders (Tewksbury 2007; Tewksbury and Lees 2006; Zandbergen and Hart 2006, 2009) and their families (Farkas and Miller 2007; Tewksbury and Levenson 2009). For instance, research confirms that residency restrictions placed on registered sex offenders significantly decreases housing opportunities perhaps to as little as 2 % of all housing stock (Zandbergen and Hart 2006, 2009). As a result of residency restrictions, registered sex offenders are also likely to be concentrated in very dense, socially disorganized communities or in rural locations with limited employment, treatment, and transportation opportunities (Minnesota Department of Corrections 2003; Tewksbury and Mustaine 2006, 2008; Zandbergen and Hart 2006, 2009).

Conclusion

Clearly there are a number of criminal justice policies that are specific to sex offenders and which establish procedures and practices that

are unique for offenders convicted of sexual offenses. What may be most interesting about the ways that policies have developed and been rapidly implemented across the country are both that they tend to be in response to particular, individual cases that have generated significant public (e.g., media) attention and that these policies and their outlined practices are largely untested.

As research has started to address sex offender criminal justice policies what is quickly becoming clear is that these policies are not being shown to be very effective, they are expensive and inefficient and in many cases seem to be overreaching and including many offenders who may not need to be responded to in such ways. The future holds many questions for how and why the types of policies discussed above can and should be continued to not. However, at the present time there is little to suggest that such policies will be scaled back. Rather, it seems that criminal justice policies regarding sex offenders are only continuing to be expanded and made more harsh, despite emerging evidence suggesting that they are not necessary effective and are very inefficient.

Related Entries

- ▶ [Sex Offender Treatment](#)
- ▶ [Sexual Recidivism](#)

Recommended Reading and References

- Billings J, Bulges C (2000) Comment: maine's sex offender registration and notification act. Wise or wicked? *Maine Law Rev* 52:175–259
- Brooks AD (1995) Megan's law: the legal issues. *Crim Justice Ethics* 14(2):12–16
- Colorado Department of Public Safety (2004) Report on safety issues raised by living arrangements for and location of sex offenders in the community. Sex Offender Management Board, Denver
- Farkas MA, Miller G (2007) Reentry and reintegration: challenges faced by the families of convicted sex offenders. *Fed Sentencing Rep* 20:88–92
- Fitzgerald EA (1990) Chemical castration: MPA treatment of the sexual offender. *Am J Crim Law* 18(1):1–60

- Friedland S (1999) On treatment, punishment, and the civil commitment of sex offenders. *Uni Colorado Law Rev* 70(1):73–154
- Iacono WG, Lykken DT (1997) The validity of the lie-detector: two surveys of scientific opinion. *J Appl Psychol* 82:426–433
- Melella JT, Travin S, Cullen K (1989) Legal and ethical issues in the use of antiandrogens in treating sex offenders. *Bull Am Acad Psychiatr Law* 17(3):223–232
- Meyer WJ, Cole C (1997) Physical and chemical castration of sex offenders: a review. *J Offender Rehabil* 25:1–16
- Meyer WJ, Cole C, Emory E (1992) Depo-Provera treatment for sex offending behavior: an evaluation of outcome. *Bull Am Acad Psychiatr Law* 20(3):249–259
- Minnesota Department of Corrections (2003) Level three sex offenders' residential placement issues. Author, St. Paul
- Sample LL, Kadleck C (2008) Sex offender laws: legislators' accounts of the need for policy. *Crim Justice Policy Rev* 19(1):40–62
- Spalding LH (1998) Article: Florida's 1997 chemical castration law. A return to the dark ages. *Fla State Uni Law Rev* 25:117–139
- Tewksbury R (2004) Experiences and attitudes of registered female sex offenders. *Fed Probat* 68(3):30–33
- Tewksbury R (2005) Collateral consequences of sex offender registration. *J Contemp Crim Justice* 21(1):67–81
- Tewksbury R (2007) Exile at home: the unintended collateral consequences of sex offender residency restrictions. *Harv Civ Rights – Civ Lib Law Rev* 42:531–540
- Tewksbury R, Lees MB (2006) Perceptions of sex offender registration: collateral consequences and community experiences. *Sociol Spectr* 26:309–334
- Tewksbury R, Lees MB (2007) Perceptions of punishment: how registered sex offenders view registries. *Crime Delin* 53(3):380–407
- Tewksbury R, Levenson JS (2009) Stress experiences of family members of registered sex offenders. *Behav Sci Law* 27:611–626
- Tewksbury R, Mustaine EE (2006) Where to find sex offenders: an examination of residential locations and neighborhood conditions. *Crim Justice Stud* 19:61–75
- Tewksbury R, Mustaine EE (2008) Where registered sex offenders live: community characteristics and proximity to possible victims. *Vict Offender* 3:86–98
- Tewksbury R, Mustaine EE (2009) Stress and collateral consequences for registered sex offenders. *J Public Manag Soc Policy* 15:215–239
- Uggen C, Manza J, Behrens A (2004) Less than the average citizen: stigma, role transition and civic reintegration of convicted felons. In: Maruna S, Immerigeon R (eds) *After crime and punishment: pathways to offender reintegration*. Willan Publishing, Devon
- Walker JT, Golden JW, VanHouten AC (2001) The geographic link between sex offenders and potential victims: a routine activities approach. *Justice Res Policy* 3(2):15–33
- Wettstein WM (1992) A psychiatric perspective of Washington's sexually violent predator statute. *Univ Puget Sound Law Rev* 15:597–633
- Wilcox DT (2000) Application of the clinical polygraph examination to the assessment, treatment, and monitoring of sex offenders. *J Sex Aggress* 5(2):134–152
- Zandbergen PA, Hart TC (2006) Reducing housing options for convicted sex offenders: investigating the impact of residency restrictions laws using GIS. *Justice Res Policy* 8(2):1–24
- Zandbergen PA, Hart TC (2009) Availability and spatial distribution of affordable housing in Miami – Dade County and implications of residency restriction zones for registered sex offenders. American Civil Liberties Union of Florida, Miami
- Zevitz R, Farkas MA (2000a) The impact of sex-offender community notification on probation/parole in Wisconsin. *Int J Offender Ther Comp Criminol* 44(1):8–21
- Zevitz R, Farkas MA (2000b) Sex offender community notification: managing high-risk criminals or exacting further vengeance? *Behav Sci Law* 18:375–391

Sex Offending and Criminal Mobility

Eric Beauregard

School of Criminology, Simon Fraser University, Burnaby, BC, Canada

Overview

For more than 60 years now, sex offenders have been the object of scrutinized attention by policy makers and the police (Barnes et al. 2009). However, it is only since 1994 that specific legislations regarding sex offenders have been passed throughout the United States to counter sex offender recidivism and alleviate community safety concerns. The Jacob Wetterling Act requires all states to track sex offenders through the use of registration systems. Such policy forces sex offenders to report to the authorities their home address. This was soon followed by Megan's Law, which requires law enforcement to notify communities when a sex offender is returned to the area (Zevitz 2006). Through the use of different means (e.g., flyers, posters, public meetings, automated phone messages), the public is informed of the identity and the residential

address of the sex offender integrating into their community, allowing for an informal network of surveillance. Residence restriction laws constitute the most recent type of policy specifically targeting sex offenders. Such policies restrict sex offenders from living within close proximity (e.g., between 1,000 and 2,500 ft) of places where children are typically present (e.g., schools, parks, playgrounds, bus stops), in order to prevent them from getting access and making direct contact with potential victims (Duwe et al. 2008). What is particular to these policies is the fact that they all share a common focus – that is, *where* the offender lives. Thus, the rationale for such policies is based on the same underlying assumption – that sex offenders are geographically stable, meaning that they do not travel when committing their crimes. But is this really the case? The current paper aims to address three related questions: (1) What do we know about sex offenders' criminal mobility? (2) Is journey to crime an adequate measure of criminal mobility? (3) Are there any other ways to look at criminal mobility of sex offenders?

Criminal Mobility: The Case of the Sex Offender

Most studies interested in the criminal mobility of offenders use the journey to crime as their main measure. These studies generally show that offenders travel a longer distance to commit property offense compared to violent crimes (see for instance Brantingham and Brantingham 1981). When looking at violent crimes in particular, a total of 21 studies were identified that examined crime trips of sex offenders in North American and European cities (Beauregard et al. 2005). For all of these studies, it is interesting to note that the crime trip distance traveled by different types of sex offenders varies between no distance traveled (the offender committing the crime at home) and 40 km. On average, the distance traveled by different types of sex offenders between their home base and the crime location was a little more than 2 miles.

Interestingly, some authors have attempted to further our understanding of the criminal mobility of sex offenders by examining two types of factors influencing their journey to crime: the offender and the offense characteristics. The underlying principle of the relationship between offender characteristics and criminal mobility is that these characteristics influence the cognitive map of individuals. According to Brantingham and Brantingham (1993), cognitive maps are a representation of the awareness of space, which consists of subjective images of an individual's environment that are fundamental in determining the areas where the criminal's offense will be carried out. Furthermore, cognitive maps vary with the characteristics of individuals. For instance, age is one of the characteristics associated with criminal mobility. Most studies of sex offenders that investigate the relationship between age and distance come to the same conclusion as for other types of criminals: Younger men tend to offend nearer to home (Warren et al. 1998). According to these studies, this difference could be attributed to a greater impulsivity in the offense behavior of younger offenders, a greater access to vehicles by older offenders, or simply because of the age-related development of the cognitive map. However, recent studies have not confirmed this age-distance relationship (see Rossmo et al. 2004).

There appears to be a clear relationship between race and the criminal mobility of sex offenders. Findings show that white rapists traveled farther than nonwhite offenders (Canter and Gregory 1994; Warren et al. 1998). Even if this relationship was once again not confirmed by Rossmo et al. (2004), it is hypothesized that this finding may reflect class distinctions or cultural differences in the cognitive mapping of space. Moreover, Gabor and Gottheil (1984) found that those with a criminal record were substantially more likely to be transient than those without one, suggesting a positive relationship between criminal career and mobility. Others have suggested that sexual fantasy is another factor related to longer sexual crime travel distance. These offenders spend long periods of time

prowling for victims, and sometimes record a diary of their movements. In addition, they are willing to travel long distances to commit a crime that will reflect their fantasies (Dietz et al. 1990). Finally, a study showed that psychopaths displayed greater geographic mobility than did nonpsychopaths (Hunter 2004). The psychopath's impulsivity, short-term relationships, unstable employment, and need for stimulation/proneness to boredom may predispose them to geographic mobility (Cooke 1998). Another explanation suggests that psychopaths frequently move locations, as their tendency to con and exploit others eventually becomes known and they are no longer able to take advantage of people in their surroundings.

As to the offense characteristics, Lebeau (1987a) focused on how the journey to rape varies as a function of the offender's approach method. Results revealed that offenders traveled the shortest mean distance to assault their victims when they illegally entered the victim's residence, suggesting that offenders travel shorter distances when using a method linked to crimes against property. In a different study, Canter and Gregory (1994) found that rapists who offend during the weekend travel farther than those who commit rape during a weekday and that rapists who attack outdoors traveled approximately 2.7 times farther to offend as those who raped indoor (e.g., in a house). Davies and Dale (1995) suggested also that rapists who target victims from a particular area (e.g., prostitutes from a red-light district), who commit sophisticated property offenses during a sexual assault, who spend large amounts of time roaming and using public transportation, and who are familiar with numerous neighborhoods (previous habitation, locations of significant people, current or past workplace locations), travel longer distances to commit their crimes. This was echoed by Warren et al. (1998) who found similar results in relation to the sophistication of the crime. Their results showed that rapists who had more extensive criminal antecedents, who used forced entry, and who burglarized the victim during the assault tended to travel farther. According to the authors, this could reflect a more generalized criminal

motivation and a more experienced offender in terms of nonsexual crimes.

The research on the journey to crime of sex offenders has not only permitted to confirm that, as with other violent offenders, sex offenders do not travel long distances to commit their crime, but it has also allowed to examine the factors that could influence their criminal mobility. However, it appears that all the studies have taken for granted that journey to crime was the only and probably the most adequate measure of criminal mobility. Whether or not this is true, none of these studies have questioned the use of the journey to crime to measure their criminal mobility.

Journey to Crime: A Valid Measure of Criminal Mobility?

Previous studies on the criminal mobility of sex offenders show that their journey to crime is mostly constant, being not very far from their home base. This information is not only useful for environmental criminology but also for criminal investigations, more specifically geographic profiling techniques (Rossmo 2000). However, relying solely on the measure of journey to crime to characterize their criminal mobility can be problematic for three main reasons.

First, the studies examining the journey to crime have all used the home location of the offender as the starting point of this crime trip. Although this is congruent with theoretical models such as the routine activity theory (Cohen and Felson 1979), such a measure does not take into account a crime trip that might have originated from the offender's workplace for instance. A rapist could identify and target a victim while coming back from a friend's house. Moreover, as Bernasco (2010) found, offenders are more likely to target former residential areas to commit their crimes if they lived there for a long time. Moreover, as mentioned by Michaud and Morselli (2011), homeless offenders and many street criminals may not have a fixed home address, therefore choosing to base their criminal activities from other social activity locations such as bars or pool halls

(Rengert 1996). Again, as stated by Michaud and Morselli (2011), using the offender's home as the exclusive platform from which crimes are committed is even more surprising given the fact that most people spent around half of their days outside home (see for instance Wikström et al. 2010). Moreover, Pettitway (1995) is the only one to our knowledge who showed through interviews with crack users the true origin of the crack trip. Interestingly, he found that only 26 % of crack-purchasing trips originated from the user's place of residence. The lack of studies taking into consideration other points of origin in the journey to crime might be partly explained by the fact that most of these studies rely exclusively on police data and other anchor points of the offender might not be known.

Second, journey to crime research suggests that the whole criminal event takes place all at the same location. Although this may be the case for property crimes such as burglary, the reality is often different for crimes where the victim is mobile (Beauregard et al. 2010). In crimes such as sexual assaults, the offender may encounter a victim at a certain location, decide to attack her at another location, take her to a different location to commit the assault, and take her to another location to release her, the so-called EAMD classification used in geographic profiling (Rossmo 2000; see also Lebeau 1987b for rape). Michaud and Morselli (2011) explain that the non-consideration of multiple crime sites in journey to crime research is problematic, given that it may distort the estimation of the criminal mobility involved in the process. To illustrate their point, they provide the example of an offender who encounters a woman in a bar, takes her to a hotel, rapes and kills the woman there, to finally go back to his place with the dead body in order to dispose of her body parts in garbage bags. If the journey to crime was calculated with this case, police data would record the address where the body parts were found (offender's residence; distance-to-crime = 0 mi), thus greatly underestimating the criminal mobility that was involved in the process. Such example is not unusual in sex-related crimes. For instance, it has been shown that sexual murderers dispose

of the victim's body on average 17.2 miles from the murder scene (Häkkinen et al. 2007).

Third, research on criminal mobility seems to suggest that the concept of mobility can and should only be measured in terms of distance traveled. However, the measure of distance traveled represents only one dimension of the mobility concept – that is, the movements or the mobility performances (Canzler et al. 2008). However, offenders also present other “movements” that may characterize their criminal mobility. Using the same example presented above, an offender may travel from one location to another during the same event. For instance, some sexual murderers decide to move the victim's body to a different location after the murder (Beauregard and Field 2008). Although such finding does not provide a measure of *distance* per se, it nonetheless provides some information on the mobility of the offender during the criminal event. This is not to suggest that the measure of journey to crime is not appropriate to examine criminal behavior and that this research is not important for our understanding of criminal mobility. However, this suggests that there is a need for criminology to look at complementary measures as well.

Alternative Measures of Criminal Mobility in Sex Offending

In addition to the journey to crime measure, a few alternative methods to conceptualize criminal mobility have been used in the research on sex offenders. One such approach has been to classify sex offenders according to their criminal mobility patterns. Although using different labels (see, for instance, Canter and Larkin 1993; Rossmo 1997), these typologies can be grouped under two main categories: the *geographically stable* and the *geographically mobile* offenders (Beauregard et al. 2005). This typological approach is interesting as it presents other characteristics (e.g., offender characteristics, modus operandi) associated with each mobility pattern. However, such approach presents also certain limitations. The majority of these typologies have been identified intuitively,

without being tested empirically, which carries problems of validity and reliability. Moreover, the types are mainly descriptive and offer little information to explain the criminal mobility patterns of offenders (Michaud and Beauregard 2010). Also, using typologies to classify criminal mobility patterns presents the same problem as with any other typology: They assume that crime-commission processes are stable instead of being dynamic (Beauregard et al. 2007).

One way to overcome these specific limitations is to conceptualize geographic mobility as the use of multiple locations for the purpose of repetitive sexual contact with the same victim. In their study, Leclerc, Wortley, and Smallbone (2010) set out to examine whether offending differences existed between perpetrators who used multiple locations for sexual contact and those who used a single location for the entire crime-commission process. Overall, the results demonstrate that mobile offenders are more likely to isolate their victims, use violence, involve the victim in several sexual episodes, abuse the victim for over a 1-year period, and make the victim participate and perform sexual behaviors on them during sexual episodes. The authors concluded that by examining mobility of pedophiles from a location angle rather than measures of distance and direction provides a different perspective on the crime-commission process of these offenders (Leclerc et al. 2010).

Following Leclerc et al. (2010) study, Beauregard and Busina (2013) used a similar approach with serial sex offenders. They proposed that criminal mobility can be defined as the number of changes of location during the criminal event. As discussed previously, rape events present different stages – that is, encounter, attack, crime, and victim release – that may be associated with different locations. Although some sex offenders decide to commit all their action at the same location (i.e., stable offender with zero change of location), other mobile offenders may change location up to three times during the same event. As criminal mobility can be interpreted as a purposive action necessary to successfully commit a crime, the aim of their study was to predict the criminal mobility

patterns exhibited in serial rape events from situational and modus operandi characteristics. The situational characteristics of the rape events and the modus operandi used in serial sex crimes might explain why some offenders need to be mobile and change location during the criminal event while others do not. Using negative binomial regression, the authors found that events which involve child or adolescent victims, committed during daytime, when the offender did not use pornography prior to crime, and where victim resistance is observed, should display more criminal mobility. Moreover, when the victim is selected, the victim is alone when approached by the offender, and the crime is characterized by sexual penetration and a lack of premeditation are exhibiting more criminal mobility. These results point toward the fact that criminal mobility is a goal-oriented action taken by serial sex offenders in order to complete successfully their crime and to avoid detection and apprehension (Beauregard and Busina 2013).

Rossmo, Lu, and Fang (2011) examined the spatial-temporal patterns of a group of reoffending parolees (many of which were sex offenders) on the Florida Department of Corrections electronic monitoring and global positioning system program. Their travel over a period of at least 8 days, including the offending day, was mapped and analyzed. This allowed analyzing the spatial activity patterns of criminals prior to, during, and after offending. At the aggregate level, results showed that the mean distance traveled 37.8 mi (min = 3.5 mi, max = 79.5 mi), covering an area of 27.2 mi². The average time spent traveling was 10.6 h (min = 1.0 h, max = 15.8 h), compared to an average of 12.4 h spent at home (min = 4.5 h, max = 20.9 h). On average, these offenders traveled 4.1 sites during their daily travel (min = 2.0, max = 6.5). When disaggregating these results, the findings are also very interesting, especially to further our knowledge on criminal mobility. Rossmo et al. (2011) described the spatial movements of a sex offender. This offender visited the offense location twice in the week before the crime. His movements started late in the day and went through to late in the night or early in

the morning of the next day. The analysis revealed also that his routine began with a trip from his home to an activity site northwest of his residence and that he passed by the crime location for the first time during the study period, but did not stop. Moreover, their analysis showed that the actions of the offender were different when he next visited the site. On that day, the offender passed the crime location again on his way to his routine nighttime activity sites northwest of his home but he made extra turns and stops. The same travel pattern was repeated the day of the offense. As mentioned by Rossmo et al. (2011), “a better understanding on this level can assist practitioners and academics in several ways, such as providing early warning cues for offending, better knowledge of how criminals hunt for their targets, and an enhanced understanding of offender spatial behavior” (p. 39).

Finally, another alternative to the traditional journey to crime measure of criminal mobility has been suggested by Michaud and Beauregard (2010). They conceptualized the criminal mobility of offenders as having two distinct dimensions: the *criminal migration* – which is the effective distance traveled by a mobile offender between the activity spaces in which he commits crimes – and the *criminal nomadism* – which corresponds to the offender propensity to change the activity space in which he commits crimes, either by exploring new activity spaces or by traveling back to former ones. Instead of looking only at one criminal event, Michaud and Beauregard (2010) consider the entire criminal career of their sex offenders. As such, they have analyzed 3,003 criminal events (sex and non-sex crimes) committed by 461 sex offenders. Their findings showed that in terms of criminal migration, only a third of the sex offenders traveled a short distance (<150 km) during their criminal career and approximately one quarter traveled more than 1,000 km. As to the criminal nomadism, one third of the sample changed activity space in at least 70 % of their criminal career. In order to understand what could explain this criminal mobility, they examined different offender characteristics. Their findings showed that sex offenders who are psychopaths,

Caucasian, educated, and specialized in sex offending present more nomadism than the sex offenders who are not.

Conclusion

Recent legislations specific to sex offenders – whether it is registration, community notification, or residence restriction – have all focused on one aspect of the crime-commission process: location. The rationale for such a focus is that by knowing where sex offenders live and by forbidding them to reside near certain location, we can prevent sexual recidivism. Although in theory, this all makes sense, all these approaches make the same mistake in assuming that criminals – more specifically sex offenders – are geographically stable and are not likely to travel when contemplating a criminal opportunity. Research shows the opposite. Sex offenders are rational individuals who may decide to travel to a different location either to avoid detection and apprehension, to complete successfully the sexual assault, or both. This suggests that even if a sex offender cannot live near a park because of a residence restriction law, nothing prevents him from leaving his residence, going to the park to encounter potential victims, and take them back to his place to sexually assault them. Studies have already shown that these legislations are ineffective to prevent sexual recidivism (e.g., Duwe et al. 2008). What has been presented here contribute further to this conclusion by showing how criminal mobility is an adaptive response to the criminal situation aiming to decrease the risks of detection and apprehension, while maximizing the chances of successfully completing the crime (e.g., obtaining sexual gratification, being able to perform sexual penetration actions).

Moreover, although journey to crime research has been helpful to the understanding of the geography of crime and to the development of investigative tools such as geographic profiling, this unique measure of criminal mobility is no longer adequate to capture the real spatial movement of offenders when committing crimes. New ways to

investigate and conceptualize criminal mobility in sex offenders have been suggested which, it is hoped, will spark some interest in academics involved in the crime and place research for other types of crimes as well. The research on criminal mobility needs to be extended in order to investigate additional aspects that journey to crime has overlooked. In addition to what has been presented, two interesting concepts could further our understanding of criminal mobility. The first concept comes from Frank, Andresen, and Felson's (2012) study who found strong evidence for geodiversity (i.e., variations in the amount of area covered by various crimes depending on the variations of criminal opportunity) in cases of co-offending or co-victimization. The second concept of interest, which appears even more important to consider is directionality (Bernasco and Block 2009). Although it has been under-researched by environmental criminologists, some research has shown that directional knowledge is important for spatial decision making (Frank et al. 2012).

Related Entries

- ▶ [Crime Location Choice](#)
- ▶ [Criminal Profiling](#)
- ▶ [History of Geographic Criminology Part I: Nineteenth Century](#)
- ▶ [Offender Decision Making and Behavioral Economics](#)
- ▶ [Situational Crime Prevention](#)
- ▶ [Theories for Situational and Environmental Crime Prevention](#)

Recommended Reading and References

- Barnes JC, Dukes T, Tewksbury R, De Troye TM (2009) Analyzing the impact of a statewide residence restriction law on South Carolina sex offenders. *Crim Justice Pol Rev* 20:21–43
- Beauregard E, Field J (2008) Body disposal patterns of sexual murderers: implications for offender profiling. *J Polic Crim Psychol* 23:81–89
- Beauregard E, Proulx J, Rossmo K (2005) Spatial patterns of sex offenders: theoretical, empirical, and practical issues. *Aggress Violent Behav* 10:579–603

- Beauregard E, Proulx J, Rossmo K, Leclerc B, Allaire J-F (2007) Script analysis of hunting process in serial sex offenders. *Crim Justice Behav* 34:1069–1084
- Beauregard E, Rebocho MF, Rossmo K (2010) Target selection patterns in rape. *J Investig Psychol Offender Profil* 7:137–152
- Beauregard E, Busina I (2013) Journey “during” crime: Predicting criminal mobility patterns in rape. *J Interpers Violence*. OnlineFirst
- Bernasco W, Block R (2009) Where offenders choose to attack: A discrete choice model of robberies in Chicago. *Criminology*, 47:93–130
- Bernasco W (2010) A sentimental journey to crime: effects of residential history on crime location choice. *Criminology* 48:389–416
- Brantingham PJ, Brantingham PL (1981) *Environmental criminology*. Waveland Press, Prospect Heights
- Brantingham PL, Brantingham PJ (1993) Environment, routine and situation: toward a pattern theory of crime. In: Clarke RV, Felson M (eds) *Routine activity and rational choice*. Transaction, New Brunswick, pp 259–294
- Canter D, Gregory A (1994) Identifying the residential location of rapists. *J Forensic Sci Soc* 34:169–175
- Canter D, Larkin P (1993) The environmental range of serial rapists. *J Environ Psychol* 13:63–71
- Canzler W, Kaufmann V, Kesselring S (2008) Tracing mobilities – an introduction. In: Canzler W, Kaufmann V, Kesselring S (eds) *Tracing mobilities: towards a cosmopolitan perspective*. Ashgate, Aldershot, pp 1–10
- Cohen L, Felson M (1979) Social change and crime rate trends: a routine activity approach. *Am Sociol Rev* 44:588–608
- Cooke DJ (1998) Psychopathy across cultures. In: Cooke DJ, Forth AE, Hare RD (eds) *Psychopathy: theory, research, and implications for society*. Kluwer, Dordrecht, pp 13–45
- Davies A, Dale A (1995) Locating the stranger rapist (Special Interest Series: Paper 3). Police Research Group, Home Office Police Department, London
- Dietz PE, Hazelwood RR, Warren J (1990) The sexually sadistic criminal and his offences. *Bulletin of the American Academy of Psychiatry and Law*, 18:163–178
- Duwe G, Donnay W, Tewksbury R (2008) Does residential proximity matter? A geographic analysis of sex offense recidivism. *Crim Justice Behav* 35:484–504
- Frank R, Andrrsen MA, Brantingham P (2012) Criminal directionality and the structure of urban form. *J Environ Psychol* 32:37–42
- Frank R, Andresen MA, Felson M (2012) The geodiversity of crime: evidence from British Columbia. *Appl Geogr* 34:180–188
- Gabor T, Gottheil E (1984) Offender characteristics and spatial mobility: an empirical study and some policy implications. *Can J Criminol* 26:267–281
- Häkkinen H, Hurme K, Liukkonen M (2007) Distance patterns and disposal sites in rural area homicides committed in Finland. *J Investig Psychol Offender Profil* 4:181–197

- Hunter SM (2004) Geographical mobility among male psychopaths. Unpublished Honour's thesis, Simon Fraser University, Burnaby, Canada
- Lebeau JL (1987a) The journey to rape: geographic distance and the rapist's method of approaching the victim. *J Polic Sci Admin* 15:129–136
- Lebeau JL (1987b) Patterns of stranger and serial rape offending: factors distinguishing apprehended and at large offenders. *J Crim Law Criminol* 78:309–326
- Leclerc B, Wortley R, Smallbone S (2010) Investigating mobility patterns for repetitive sexual contact in adult child sex offending. *J Crim Justice* 38:648–656
- Michaud P, Beauregard E (2010) Criminal nomadism: a neglected dimension of sex offender criminal mobility. Paper presented at the annual meeting of the American Society of Criminology, San Francisco
- Michaud P, Morselli C (2011) Should we stay or should we go? A critical review of journey-to-crime research. Submitted for publication
- Pettitway LE (1995) Copping crack: the travel behavior of crack users. *Justice Quart* 12:499–524
- Rengert GF (1996) *The geography of illegal drugs*. Westview Press, Boulder
- Rossmo DK (1997) Geographic profiling. In: Jackson JL, Bekerian DA (eds) *Offender profiling: theory, research and practice*. Wiley, Chichester, pp 159–175
- Rossmo DK (2000) *Geographic profiling*. CRC Press, Boca Raton
- Rossmo DK, Davies A, Patrick M (2004) Exploring the geo-demographic relationship between stranger rapists and their offences (Special Interest Series). Policing and Reducing Crime Unit, Home Office, London
- Rossmo DK, Lu Y, Fang TB (2011) Spatial-temporal crime paths. In: Andresen MA, Kinney JB (eds) *Patterns, prevention, and geometry of crime*. Routledge, New York, pp 16–42
- Safarik ME, Jarvis JP, Nussbaum KE (2002) Sexual homicide of elderly females: linking offender characteristics to victim and crime scene attributes. *J Interpersonal Violence* 17:500–525
- Warren J, Reboussin R, Hazelwood RR, Cummings A, Gibbs N, Trumbetta S (1998) Crime scene and distance correlates of serial rape. *J Quant Criminol* 14:35–59
- Wikström P-OH, Ceccato V, Hardie B, Treiber K (2010) Activity fields and the dynamics of crime. Advancing knowledge about the role of the environment in crime causation. *J Quant Criminol* 26:55–87
- Zevitz RG (2006) Sex offender community notification: its role in recidivism and offender reintegration. *Crim Justice Stud* 19:193–208

Sexual Assault and Violence Against Females

- [Feminist Theory in the Context of Sexual Violence](#)

Sexual Offender Treatment (SOT)

- [Treatment of Sex Offenders](#)

Sexual Recidivism

Brian Francis and Keith Soothill
Department of Maths and Statistics, Lancaster University, Lancaster, UK

Synonyms

[Sexual reconviction](#); [Sexual reoffending](#)

Overview

Sexual recidivism is concerned with the reoffending of sexual offenders who have already had contact with the criminal justice system. This entry emphasizes the need to carefully define what is meant by recidivism and discusses some of the pitfalls in this area. It then focuses on recidivism rates, following on with a consideration of risk factors for recidivism, both static and dynamic, before commonly used risk scores are introduced. A discussion on specialization and glances towards possible developments in this area concludes the section.

Fundamentals

This entry is concerned with sexual recidivism, that is, the “persistence” of sexual offending once an offender has been arrested or convicted and addresses the future risks posed to the public by sex offenders after conviction.

Defining Recidivism

To begin, definitional aspects of recidivism and its measurement are considered, including the outcome measure, the type of recidivism, the

time horizon, and the definition of what constitutes a sexual offense.

Recidivism: Recidivism is defined as the reoffending of a known offender; however, Falshaw et al. (2003) have noted that “there is little consistency in the way this term is used and the specific behaviour it refers to.” They remind the reader that Maltz, in his classic book on recidivism, pointed to the use of nine different indicators of recidivism across a total of 90 studies in the United States, i.e., absconding, arrest, incarceration, parole violation, parole suspension, parole revocation, reoffense, reconviction, and probation violation.

Focusing explicitly on sexual recidivism, Falshaw et al. (2003) distinguished between sexual reconviction, sexual reoffending, and sexual recidivism. Their view was that recidivism is primarily concerned with lapsing into previous patterns of behavior and thus encompassed other forms of potential sexual-offense-related behavior such as loitering outside a school. They defined sexual reconviction as a subsequent conviction for a sexual offense, sexual reoffending as the perpetration of another illegal sexual act (whether caught or not), and sexual recidivism as the commission of a behavior related to a sexual offense, legal or illegal, with a clear sexual motivation. A more common use of the term sexual recidivism in the literature and one which is used in this entry, however, is some kind of official contact with the criminal justice system for a sexual offense, whether it be arrest, charge, or reconviction.

There is an implicit hierarchy within these distinctions. A “sexual reconviction” is the strictest measure as, to count, one needs the endorsement of a court of law. “Sexual recidivism” is a broader measure and may include arrests or charges as well as convictions for a sexual offense. “Sexual offending” has difficult counting and legal implications as it includes illegal sexual acts, whether caught or not. It is crucial to understand the importance of such analytical distinctions as the use of different outcome measures can produce very different results.

Type of Recidivism: Many studies on the recidivism of sexual offenders use as their

outcome measure any form of recidivism – a general measure of subsequent contact with the criminal justice system for *any* offense. Others are interested in whether *sexual* recidivism has taken place, whereas yet others are concerned with recidivism for *dangerous* offenses which is usually taken to be sexual or violent recidivism.

Time Horizon for Follow-Up: Some studies have a fixed time horizon or a fixed series of time gates whereby recidivism rates, say, 5 or 10 years after release can be established. Other studies use an average follow-up time – these are less valuable as their results cannot be generalized in systematic reviews.

Nature of the Sample Taken: Some recidivism studies relate to those incarcerated in prison or in a treatment facility; others involve those with community sentences as well. Other aspects are also important – are females as well as males included, and is the sample restricted to a particular age range? Specifically, are juvenile offenders included?

Definition of a Sexual Offense: Normally, legal statutes which categorize sexual offenses (e.g., rape, indecent assault, exhibitionism) are used, but these can differ from country to country and from US state to US state. Most of these differences currently relate to laws against consensual homosexual activity (e.g., Malaysia, Zimbabwe). In addition, the age of sexual consent varies from country to country, varying from 12 (e.g., parts of Mexico) to 20 (Tunisia), and in some countries sexual behavior outside marriage is illegal (e.g., Saudi Arabia, UAE). The age of consent also varies from state to state in the USA and in Australia. There is a further issue relating to long-term or historical recidivism studies in that the definition of what is a sexual offense will change over time. A specific example relates to the laws of sodomy, or “unnatural” sex, which was in effect a legal prohibition in 14 US states until 2003. This makes comparison of recidivism rates across time, across countries, and across states problematic.

There are two further issues relating to whether an offense can be classified as sexual. Some offenses may be classified as a sexual

offense but may indeed have no sexual motivation. For example, some might regard “bigamy,” which until recently was classified by the Home Office as a sexual offense in England and Wales, as more of a deception than a sexual offense. In contrast, other offenses captured under “theft” (e.g., stealing underwear) could be regarded as more indicative of a sexual than acquisitive motivation but are unlikely to be included. Less well recognized is the fact that most reconviction studies are based on the principal offense committed. Hence, someone convicted of both murder and rape, for example, is recorded for murder as the principal offense, while the offense of rape is masked.

Risk and Prediction: A *risk* is a chance or probability that some event (usually undesirable) will happen in the future. Thus, one can talk about the risk of a sexual offender being reconvicted for a sexual offense in the next 5 years on release from an English prison. Good risk statements should include some element of location (risks for offenders in the Netherlands may well be lower than in the UK) and some indication of a future time horizon (the next year, the next 10 years, etc.) – however, these are sometimes left implicit, leaving the reader to judge what the statement means. A *prediction* usually relates to a particular person and uses the risk measure to make a judgment about an individual. This prediction may, in turn, be acted upon and some judgment made about the individual

based on the prediction. Mostly, the assessment of risk in criminology is about potential dangerousness.

The Risk of Sexual and General Recidivism

Both large numbers and a long-term follow-up are thought to be crucial in the quest for a definitive study about recidivism rates. Although systematic reviews suggest that the observed sexual recidivism rates are only 10–15 % after 5 years (Hanson and Bussière 1998), the rates continue to increase gradually with extended follow-up periods. The study by Prentky et al. (1997), using a follow-up to 25 years from a treatment center in Massachusetts, demonstrated that, if they had restricted the follow-up to less than 24 months, they would have missed up to 45 % of new charges. While even with a 5-year follow-up period, they would still miss 30 % of the charges they finally identified. In fact, they examined the cumulative failure rates for new charges for sexual offenses only as well as for new charges of any offense. They used nine time gates, broken down by charge, conviction, and imprisonment. Table 1 shows the recidivism rate for new *sexual* charges.

As Table 1 shows, the conviction rates after 2, 5, 10, and 25 years were 6 %, 11 %, 16 %, and 24 % for rapists, while the comparable figures for

Sexual Recidivism, Table 1 Cumulative sexual recidivism rates over nine time gates

Rate at given time gate									
Type of recidivism	1 year	2 years	3 years	4 years	5 years	10 years	15 years	20 years	25 years
<i>Rapists (N = 136)</i>									
Sexual reoffense charge	.09	.12	.15	.17	.19	.26	.31	.36	.39
Sexual offense conviction	.04	.06	.08	.09	.11	.16	.20	.23	.24
Sexual offense conviction with prison disposal	.04	.06	.07	.08	.10	.14	.17	.19	.19
<i>Child molesters (N = 115)</i>									
Sexual reoffense charge	.06	.10	.14	.17	.19	.30	.39	.46	.52
Sexual offense conviction	.04	.07	.10	.12	.14	.23	.31	.37	.41
Sexual offense conviction with prison disposal	.04	.07	.09	.11	.13	.21	.28	.33	.37

Source: Prentky et al. (1997)

child molesters were 7 %, 14 %, 23 %, and 41 %. The work of Prentky and his colleagues provides a very clear case in demonstrating the importance of long-term follow-ups.

More recently, the results of Prentky et al. (1997) have been validated by Cann et al. (2004), who followed up all 419 male sexual offenders discharged from prison in 1979 in England and Wales until 2000, giving a follow-up period of 21 years. The sexual reconviction rates after 2, 5, 10, and 20 years were 10 %, 16 %, 20 %, and 25 %; the equivalent figures for general reconviction were 34 %, 49 %, 56 %, and 62 %. The sexual conviction rates are higher at the 2-year point than either of Prentky et al.'s subsamples, but after that point, the reconviction rates are very similar.

Risk Factors and Sexual Recidivism

As already stated, the risk of recidivism is the probability or chance that the offender, who has already offended once, will take place again in the future within some time horizon. The risk will depend on characteristics of the offender, and it is therefore important to identify a set of factors or variables which are known to be associated with the risk of recidivism. Five such factors are highlighted below.

Sexual Recidivism and Gender

The extent of female sex recidivism is still under-researched; however, differences between male and female sexual offenders in terms of recidivism have been observed. Freeman and Sandler (2008) took a matched series of 780 female and male sex offenders in New York State and demonstrated that male sex offenders were significantly more likely than female sex offenders to be rearrested for both sexual and nonsexual offenses. A later study by the same authors (Sandler and Freeman 2009) extended this work and took 1,466 female offenders who were then followed up for 5 years. The 5-year rearrest rate for these offenders was a very low 1.8 % for a sexual offense and 26.6 % for any offense. Cortoni et al. (2010) concurred with the general

conclusion of low rates for female sexual recidivism and, in a meta-analysis using 10 studies, found that the rates were typically under 3 % with an average follow-up of six and a half years.

Sexual Recidivism and Age

The age variable is also theoretically very relevant, particularly in relation to developmental psychology. For instance, Hanson (2002) identifies three broad factors relevant to sexual offending – deviant sexual interests (motivation), opportunity, and low self-control – and uses these to help explain the variation in the recidivism rates of various offenses. So, he maintains that, for rapists, all three factors should decline with age – “self-control should increase in young adulthood, deviant sexual drives should decrease in late adulthood, and opportunities should gradually decline throughout the life span” (Hanson 2002). Consequently, Hanson is not surprised that most rapists are young and that their recidivism rate steadily decreases with age. In contrast, in explaining that child molesters are generally older than rapists, he points to competing factors influencing recidivism risk during early to middle adulthood with, for instance, self-control probably improving but the opportunities for child molesting may well be increasing.

Thornton (2006) has noted that “sexual offenders released at a younger age tended to be more general criminals while those released at an older age tended to be sexual specialists.” But Barbaree et al. (2003), focusing more specifically on sexual recidivism, produce an extra twist. They argue that “if libido [seen as one of the important determinants of sexual aggression] decreases with aging, then it follows that sexual aggression should show similar aging effects.” Certainly their results suggested that “offenders released at an older age were less likely to recommit sexual offences and that sexual recidivism decreased as a linear function of age-at-release” (Barbaree et al. 2003).

Following reviews and meta-analyses, the consensus is that there is an inverse relationship between sexual offenders' age at the time of their release from incarceration and their sexual recidivism risk (Thornton 2006; Hanson 2002; Hanson

and Bussière 1998). However, Doren (2006) notes some recent challenges to this “iron law” – Doren found a series of “study-specific conclusions . . . that were often mutually exclusive.” Reanalysis of existing data showed numerous potential interacting variables, such as participation in treatment, type of risk measure used, type of sexual offender, jurisdiction, and even a different measure of offender age. In other words, to explain the wide disparity of findings among the studies he reviewed, there were perhaps confounding variables in all of the reviewed research. While the age variable is definitely on the agenda, Doren concludes that “we have a lot of work to do before we can say we understand how to consider offender age in sexual recidivism assessments” (p. 156).

Sexual Recidivism of Juveniles

In law and policy, the age divide between juveniles and adults is particularly crucial. However, the activity of juvenile sex offending has only comparatively recently attracted attention and research interest. In fact, over the past decade or so, sexually abusive juveniles/adolescents have increasingly been distinguished from adult offending. Worling and Långström (2003) provide a useful and detailed review of studies indicating criminal recidivism risk with adolescents who have offended sexually. They identify, for example, that their own study on 117 sexual adolescents reported a 30 % sexual reconviction rate after a mean follow-up of 9.5 years. Such rates are higher than the Prentky et al.'s 10-year reconviction rates reported earlier in this entry.

An important counterweight to the concern that juvenile sex offenders are likely to be particularly dangerous is provided by Caldwell (2007). Caldwell compared 249 juvenile sex offenders and 1,780 nonsexual offending delinquents who were released from secured custody and contrasted the sexual charge recidivism figures for juvenile sexual offenders and nonsexual offenders and found nonsignificant differences – 6.8 % compared to 5.7 % after a 5-year follow-up. This result has important implications for policy

concerns. Quite simply, most juvenile sex offenders may not be the future threat that many might have expected.

The low rates of sexual recidivism among juveniles and adolescents are certainly noteworthy. Vandiver (2006) confirms how nonsexual offenses predominate in recidivism among juvenile sex offenders. Vandiver focuses on 300 registered male sex offenders who were juveniles at the time of their initial arrest for a sex offense. The series is followed through for 3–6 years after they reached adulthood, and while more than half of the series is arrested at least once for a nonsexual offense during this adult period, only 13 (or 4 %) were rearrested for a sex offense. Similar results are portrayed in Nisbet et al. (2004) showing relatively low rates of detected adult sexual recidivism but high rates of detected nonsexual recidivism, among young men who committed sexual offenses as adolescents. Hence, one is identifying relatively stable patterns of general antisocial conduct in adolescent sex offenders, but continuing aberrant *sexual* behavior is not much in evidence. Curiously, with this study, while age at assessment was found to predict sexual recidivism, rather counterintuitively, it was an older age at assessment that predicted adult sexual offense charges. In fact, “the likelihood of being charged with sexual offences as an adult increased by 60 % with each year increase in age at assessment” (p. 230). This is an important finding as it confirms that very early sexual crime is generally unlikely to be a precursor to persistent sexual offending in adult life.

Sexual Recidivism and Type of Sexual Offending

In general, large numbers of sex offenders are needed when testing the notion that different types of sex offenders have different kinds of outcome. Earlier work such as Prentky et al. (1997) had already hinted that some types of sex offenders were more likely to commit new *sexual* crimes than others, suggesting that child molesters have a substantially higher sexual recidivism rate compared to rapists. Other authors have disagreed with this

conclusion. For example, Serin et al. (2001) have indicated that recidivism rates are higher for rapists than child molesters. Hanson (2002) notes that “rapists were younger than child molesters, and the recidivism risk of rapists steadily decreased with age.” Also differences in recidivism within the same kind of sex offense have been noted – for example, pedophilic versus non-pedophilic child molesters.

Sexual Recidivism and Prior Criminal History

The final risk factor highlighted is that of the offender’s prior criminal history. There are many ways of summarizing such information, but interest has been focused on the number of prior convictions, the number of prior sexual convictions, the age of criminal career onset, and the relationship between the victim and the offender of prior sexual offending. In general, such measures have been incorporated into risk scores, and criminal history is discussed more fully below.

Risk Scores and Sexual Recidivism

A sexual recidivism risk score, put simply, identifies the risk of future recidivism of a sex offender based on a set of risk factors. For example, the risk score might classify a typical offender with a set of characteristics into one of five categories – highly unlikely to reoffend, unlikely to reoffend, equally likely to reoffend or not, likely to reoffend, and highly likely to reoffend. A prediction of recidivism, in contrast, is using that risk probability or score (and possibly other information) to make a judgment as to whether one particular person will reoffend. Less commonly, a prediction might estimate how many of a group of offenders will reoffend. Thus, a prediction might be made that a male sex offender in California released from prison who has three prior sexual convictions has a 40 % chance of reoffending with another sexual offense in the next 5 years. Based on this prediction, a decision might be made on the form of post-release supervision needed by the offender.

Actuarial and Clinical Assessment Tools

There are two methods of assessing the risk of an individual for subsequent sexual recidivism. Actuarial measures use summary measures of criminal career data on a large set of offenders together with other information about the age, gender, and circumstance of the offender to estimate a risk score. The offenders are followed up for the required period of time, and each offender is then identified as a recidivist or not. This risk score estimate is made through the application of statistical modeling methods such as logistic regression or other score-building models. The performance of the measure can be assessed by dividing the set of offenders into two – using one part of the data to build the risk score and the second part to assess how well the measure performs.

Clinical risk assessment, on the other hand, takes the judgment of professionals such as psychiatrists, probation officers, or parole board members to make a relatively informal judgment on the likelihood of an offender to reoffend. Tools such as the Violence Risk Score for Sexual Offenders can be used to guide the judgment. Although these professionals will have access to the same information on past criminal career history, they will also take into account a whole set of personal factors such as degree of remorse, demeanor, family support, stable residential status, etc., to determine risk (Milner and Campbell 1995).

Both measures would tend to be used at the start of some process – for example, in presentence reports presented to the court or in considering release from an indeterminate prison sentence. Which of these two approaches appears to give better predictions? An important study by Grove and Meehl took 136 separate studies – a mix of clinical and actuarial studies. In general, the clinical studies had a great deal of extra information available compared to the actuarial studies (Grove and Meehl 1996). Their study came to two important conclusions. Firstly, in studies which compared practitioners, there was little agreement between them. Secondly, despite using less information, the actuarial studies were either equal or superior to clinical risk assessment.

However, there have been criticisms of the actuarial approach (Quinsey et al. 1995). Firstly, it can categorize whole groups of individuals as high risk – even though it is recognized that personal circumstances will mean that some individuals are at low risk within a high-risk group. Furthermore, actuarial measures fail to take account of factors such as the intention of the offender to desist from crime. There are also other criminological arguments which need to be considered – that actuarial risk measures focus on the individual to the exclusion of other causes of crime, notably economic and social deprivation. Nevertheless, actuarial measures are perceived by many as providing a reliable estimate of risk and are increasingly used in daily criminal justice practice. For dangerous offenders, current practice relies increasingly on some combination of a structured clinical judgment and actuarial measures.

“Static” Versus “Dynamic” Factors

The main focus in predicting recidivism risk from its outset has been on historical, or “static,” factors. These are factors relating to the prior criminal history of the offender such as age of first conviction; the number of previous sexual and nonsexual convictions; demographic factors, such as age and gender; historical factors in the early life of the offender such as whether both parents were present when the offender was a child; as well as victim characteristics (stranger, familial) and type of offending (child molestation, rape, etc.). Age is regarded as a “static” factor as it is taken at some fixed point in time, such as conviction or date of release from prison or treatment center. As such factors are “static,” it is not possible to intervene and to change these to improve outcome.

Dynamic factors, in contrast, are subject to intervention. They can be divided into stable dynamic factors, which change slowly over time (such as job responsibility and attitudes to the opposite sex), and acute dynamic factors, which can vary day by day or hour by hour (such as day to day drinking behavior).

Building Risk Scores

Many risk scores in common use will apply statistical techniques to build a recidivism risk score. Data on a cohort of sex offenders is collected and each offender in the sample is followed up for a fixed period of time – typically 2 or 5 years – although longer follow-up periods can also be used. Typically, official data is used to assess whether the offender has recidivated by looking at arrest or court conviction records. Information on offender characteristics is also collected. These can be obtained from the criminal history of case notes of the offender. For scores involving dynamic factors, in addition, psychometric tests (e.g., for psychopathy) may need to be administered as they often contribute to the recidivism test score. Typically, logistic regression is then used to build a score and to determine the important risk factors. If follow-up times vary, then Cox regression may be used.

Once the score is built, its performance on new samples of offenders is assessed. Performance is usually measured by estimating the receiver operating characteristic curve of the score and calculating the AUC (area under the curve) which is expressed as a proportion or percentage. The AUC can be interpreted as follows: if two offenders are taken at random – one reconvicted and one not reconvicted – then the AUC gives the probability that the reconvicted offender will have a higher risk score than the unconvicted offender. Most risk scores used in assessing sexual recidivism have AUCs of around 0.70 or 70 %.

Risk Scores for Sexual Recidivism

Table 2 presents the sexual recidivism risk scores in common use, together with typical items which make up the score. As can be seen, four of the seven scores presented consist solely of static items, and only three scores attempt in addition to include dynamic factors.

Three measures are highlighted: the Risk Matrix 2000/Sexual (RM2000/S), the Rapid Risk Assessment of Sexual Offense Recidivism (RRASOR), and the Violence Risk Score for Sexual Offenders (VRS-SO).

The RM2000/S score consists of seven static items which relate to previous criminal history

Sexual Recidivism, Table 2 Some commonly used risk scores for assessing sexual recidivism for sex offenders

Measure	Full name	Number of items	Examples of items	Developers	Year
Static-99	Static-99	10 static	Conviction and sexual conviction history, relationship to victim, victim gender, lack of long-term intimate relationship, aged under 25	Hanson and Thornton	2000
RM2000/S	Risk Matrix 2000/Sexual	7 static	Conviction and sexual conviction history, age at release, relationship to victim, victim gender, marital status	Thornton, Mann, Webster, Blud, Travers, Friendship, and Erickson	2003
SORAG	Sex Offender Risk Appraisal Guide	12 static 2 dynamic	Age at offense, criminal history, failure on prior release, marital status, history of alcohol abuse, psychopathy score, phallometric test	Quinsey, Harris, Rice, and Cornier	1998, 2006
RRASOR	Rapid Risk Assessment of Sexual Offense Recidivism	4 static	Prior sexual offenses, age at release, relationship to victim, victim gender	Hanson	1997
MnSOST-R	Minnesota Sex Offender Screening Tool – Revised	17 static	Sexual and nonsexual criminal history, substance abuse, discipline history and sex offender treatment history, age at release	Epperson, Kau, and Hesselton	1998
VRS-SO	Violence Risk Score for Sex Offenders	7 static 17 dynamic	Conviction and sexual conviction history; age at release; age of first conviction; sexual deviance, criminality, treatment responsivity	Wong, Olver, Nicholaichuk, Gordon	2003
J-SOAP	Juvenile Sex Offender Assessment Protocol	23 dynamic	Sexual drive; impulsive, antisocial behavior; clinical; community adjustment	Prentky, Harris, Frizzel, and Righthand	2000

(number of previous court appearances, number of previous court appearances for sexual offenses, any conviction for a sexual offense against a male and against a stranger, any conviction for a noncontact sexual offense) together with measures of age at assessment and marital status. The test yields four risk categories – very high, high, medium, and low. The score is easy to administer and is used extensively by the UK Prison Service. AUC values of 0.75–0.77 have been reported.

RRASOR is an even briefer score and consists solely of four items and is, as its name suggests, easy to administer. The items are the number of prior sexual offenses, age of offender, gender of victim, and relationship to victim. AUC values of between 0.65 and 0.79 have been reported by various authors.

More recently, risk scores involving both static and dynamic factors have started to be

introduced. The VRS-SO is one example of this development. The score includes both static items (conviction and sexual conviction history, age at release, age of first conviction) and a range of 17 dynamic items (sexual deviant lifestyle, sexual compulsivity, interpersonal aggression, and cognitive distortions are four of the items). The 17 dynamic items represent three underlying components of sexual deviancy, criminality, and treatment responsivity, together with additional items for intimacy deficits and emotional control. Theoretically, it uses “stages of change” to assess sexual-offending-related attitudes and behaviors and is administered twice, for example, pretreatment and posttreatment. The VRS-SO developers highlight its sensitivity to treatment-related change as well as to other forms of intervention. In terms of predictive validity at the time of development, AUC values ranged from 0.66 to 0.74. Beggs and Grace (2010) recently assessed

the score and reported an AUC of 0.80 and also stated that the dynamic items provided additional predictive power after controlling for the static items. However, it needs to be stressed that researchers need training in the use of the instrument and the administration of the instrument is lengthy.

Researchers are divided about the utility of including dynamic measures in a risk score. The developers of VRS-SO argue that dynamic variables do not necessarily have to add to the predictive efficacy of static variables to be useful, stressing that their utility in treatment and in assessing risk change is also important.

Another area of controversy is whether age needs to be included in a risk score which includes dynamic items. The argument is that as most databases used to construct tests are cross-sectional, then age differences in the sample may represent birth cohort effects as well as age, and the distinct contribution of age cannot be determined (Harris and Rice 2007). An additional concern is that the dynamic items include measures of self-control, antisocial traits, and sexual potency, which will decline with age; hence, the inclusion of age in addition to these trait measures will overpredict sexual recidivism. However, Barbaree et al. (2009) have determined that age provides additional explanatory power over antisocial and sexual deviance measures in assessing recidivism.

In general, the choice of score depends on whether assessors have the skill, training, and time to assess offenders on the dynamic items. Scores involving dynamic items are best suited to those offenders who are incarcerated where there may be an interest in possible change, while court-based assessment is probably carried out more efficiently by using a static measure.

Recidivism and Specialization

A separate but related debate to sexual recidivism is that of specialization, which is defined as the tendency to commit the same type of offense. Specialist sexual offenders therefore have a

tendency to recidivate with another sexual offense to a greater extent than the average offender.

Lussier (2005) explains that two major hypotheses have been put forward to describe the criminal activity of sexual offenders in adulthood. The first of these states that sexual offenders are specialists who tend to repeat sexual crimes. The second describes sexual offenders as generalists who are versatile in their offending. He goes on to state that the current state of knowledge provides empirical support for both the specialization and the generality hypothesis. A recent study has examined the specialization of sexual offenders both pre- and post-commitment (Harris et al. 2011). They found strong evidence of versatility but also found that those offenders who specialized prerelease were more likely than versatile offenders to specialize in sexual offending on release. They also found that child molesters were more likely to specialize than rapists or incest offenders.

The results suggest that sexual offenders are both versatile and specialist – a hypothesis originally suggested by Soothill et al. (2000). This study compared the criminal records of the 6,097 males convicted in 1973 in England and Wales of one of four offense categories – indecent assault against a female, indecent assault against a male, indecency between males, and unlawful sexual intercourse (USI) with a girl under 13 – over a 32-year observation period looking backward 10 years and forward 22 years, and very different patterns emerged for the four groups (Soothill et al. 2000). It showed that each offense group had very different criminality patterns in terms of their likelihood of being convicted of any (general) offense on another occasion – ranging from 37 % of those convicted of indecency between males to 76 % for those convicted of USI under 16. These sex offenders were shown to differ greatly in terms of general offending behavior – a *higher* proportion of those committing *heterosexual* offenses (i.e., indecent assault on female or USI with a girl under 16) tend to be convicted for violence against the person, property offenses, and criminal damage compared with the other two groups; however, a *lower*

proportion of these offenders commit *sexual* offenses on other occasions. In contrast, those committing *homosexual* offenses (i.e., indecent assault on a male or indecency between males) are the mirror image. They are much less likely, compared with other sex offenders, to be convicted of violence or property offenses, while those committing indecent assault on a male are much more likely to be convicted of sexual offenses on other occasions.

At one level, therefore, sexual offenders may or may not specialize within their general criminal career, while at a more specific level offenders may or may not specialize in specific kinds of sex offending within their sexual criminal career. These two levels may act quite independently insofar that an offender could be a specialist at one level and a generalist at another. Miethe et al. (2006) have confirmed this picture by identifying, within low levels of general specialization, that 37 % of serial child molesters and 27 % of serial rapists were identified as specialists, offending with their specific types of sexual crime in the early middle and late third of their criminal careers.

Future Directions in Risk and Sexual Recidivism

This entry concludes with an indication of possible future directions in the study of sexual recidivism, focusing on new data sources, improvements in the methodology of risk score building, and better understanding of the criminal careers of sex offenders.

New Data Sources: One major difficulty with the construction of risk scores is that, in general, they are both constructed and validated on data sets containing a relatively small number of cases. Risk scores are, moreover, often built on specific samples such as sexual offenders in prison. However, the potential is there to construct longitudinal data sets through record linkage which would take complete birth cohorts of sex offenders and follow them up in terms of their conviction or arrest records for long periods of time. For example, one potential source for such

data would be the population registers of the Scandinavian countries, which, when linked together, would provide information on police contacts, convictions, employment, income, and demographic variables.

New Methodologies for Assessing Recidivism: Improved data sets will allow more sophistication in the type of models which can be built to explain recidivism. Firstly, new forms of dynamic factors can be introduced. Marital status, for example, is currently included as a static variable, but having dynamic information on changes of marital status, as well as changes in job status and responsibility for children, will provide additional explanatory power which may prove useful. Modeling techniques such as the Cox discrete-time model can help to build prediction scores with time-varying covariates. Secondly, group-based trajectory modeling, which has proved its worth in identifying latent trajectories of offending frequency over the criminal history, and group membership of such classes is starting to be used as a predictor for future recidivism (Lussier et al. 2010). Thirdly, new methods of prediction from the machine-learning and data-mining disciplines (neural networks, support vector machines, random forests) will offer additional sophistication which can deal with the inherent nonlinearity of covariates on recidivism outcome.

Specialization and Recidivism: Most work on specialization and sexual offending has dealt with the whole of the criminal career. However, alternative methods are now available which propose that offenders may favor certain offense types during the short term, largely because of opportunity structures, but that because of changing situations and contexts over the life course, their offending profiles aggregate to versatility over their criminal career as a whole. This notion has not been considered specifically for sex offending. However, it seems possible that offenders may well have particular sex crime preferences in relatively narrow time periods and then transition to other kinds of behavior over time.

Reintegration into the Community: Recently, work has been carried out on how long after

sentence or release general offenders become similar to non-offenders in their propensity to commit crime (e.g., Bushway et al. 2011). This idea of redemption and reintegration is important as it can help to determine policy such as how long DNA samples or old criminal records are kept for. This work needs to be extended into sexual offending, and the risk or hazard of future offending for different lengths of arrest or conviction-free periods needs to be determined and compared with non-offenders. Such results should inform the length of time sexual offenders are placed on sex offender registers.

Apart from these specific developments, it is important to realize that societal factors as well as personal and individual factors will affect sexual offending and recidivism. Indeed, it can be argued that changes in society rather than individual upbringing and attitudes have affected the long-term trends in sexual offending and changes in rates of recidivism. The focus of research into sexual recidivism in the last 20 years has focused on the individual; the next 20 years needs to focus more on society and community effects.

Related Entries

- ▶ [Criminal Careers](#)
- ▶ [Risk Assessment, Classification, and Prediction](#)
- ▶ [Risk factors for Adolescent Sexual Offending](#)
- ▶ [Sex Offender Treatment](#)
- ▶ [Specialization and Sexual Offending](#)

Recommended Reading and References

- Barbaree HE, Blanchard R, Langton CM (2003) The development of sexual aggression through the life span: the effect of age on sexual arousal and recidivism among sex offenders. In: Prentky RA, Janus ES, Seto MC (eds) *Sexually coercive behavior: understanding and management*, vol 989. New York: Academy of Sciences, New York, pp 59–71
- Barbaree HE, Langton CM, Blanchard R, Cantor JM (2009) Aging versus stable enduring traits as explanatory constructs in sex offender recidivism: partitioning actuarial prediction into conceptually meaningful components. *Crim Justice Behav* 36(5):443–465
- Beggs SM, Grace RC (2010) Assessment of dynamic risk factors: an independent validation study of the violence risk scale: sexual offender version. *Sex Abuse J Res Treat* 22(2):234–251
- Bushway SD, Nieuwbeerta P, Blokland A (2011) The predictive value of criminal background checks: do age and criminal history affect time to redemption? *Criminology* 49(1):27–60
- Caldwell MF (2007) Sexual offense adjudication and sexual recidivism among juvenile offenders. *Sex Abuse J Res Treat* 19(2):107–113
- Cann J, Falshaw L, Friendship C (2004) Sexual offenders discharged from prison in England and Wales: a 21-year reconviction study. *Leg Criminol Psychol* 9(1):1–10
- Cortoni F, Hanson RK, Coache ME (2010) The recidivism rates of female sexual offenders are low: a meta-analysis. *Sex Abuse J Res Treat* 22(4):387–401
- Doren DM (2006) What do we know about the effect of aging on recidivism risk for sexual offenders? *Sex Abuse J Res Treat* 18(2):137–157
- Falshaw L, Bates A, Patel V, Corbett C, Friendship C (2003) Assessing reconviction, reoffending and recidivism in a sample of UK sexual offenders. *Leg Criminol Psychol* 8(2):207–215
- Freeman NJ, Sandler JC (2008) Female and male sex offenders – a comparison of recidivism patterns and risk factors. *J Interpers Violence* 23(10):1394–1413
- Grove WM, Meehl PE (1996) Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical-statistical controversy. *Psychol Public Policy Law* 2:293–323
- Hanson RK (2002) Recidivism and age – follow-up data from 4,673 sexual offenders. *J Interpers Violence* 17(10):1046–1062
- Hanson RK, Bussière MT (1998) Predicting relapse: a meta-analysis of sexual offender recidivism studies. *J Consult Clin Psychol* 66(2):348–362
- Harris GT, Rice ME (2007) Adjusting actuarial violence risk assessments based on aging or the passage of time. *Crim Justice Behav* 34:297–313
- Harris DA, Knight RA, Smallbone S, Dennison S (2011) Postrelease specialization and versatility in sexual offenders referred for civil commitment. *Sex Abuse J Res Treat* 23(2):243–259
- Lussier P (2005) The criminal activity of sexual offenders in adulthood: revisiting the specialization debate. *Sex Abuse J Res Treat* 17:269–292
- Lussier P, Tzoumakis S, Cale J, Amirault J (2010) Criminal trajectories of adult sexual aggressors and the age effect: examining the dynamic aspect of offending in adulthood. *Int Crim Justice Rev* 20:147–168
- Maltz MD (1984) *Recidivism*. Academic, Orlando
- Miethe TD, Olson J, Mitchell O (2006) Specialization and persistence in the arrest histories of sex offenders: a comparative analysis of alternative measures and offense types. *J Res Crime Delinq* 43(3):204–229
- Milner JS, Campbell JC (1995) Prediction issues for practitioners. In: Campbell JC (ed) *Assessing*

- dangerousness: violence by sexual offenders, batterers, and child abusers. Sage, Thousand Oaks
- Nisbet IA, Wilson PH, Smallbone SW (2004) A prospective longitudinal study of sexual recidivism among adolescent sex offenders. *Sex Abuse J Res Treat* 16(3):223–234
- Prentky RA, Lee AFS, Knight RA, Cerce D (1997) Recidivism rates among child molesters and rapists: a methodological analysis. *Law Hum Behav* 21(6):635–659
- Quinsey VL, Rice ME, Harris GT (1995) Actuarial prediction of sexual recidivism. *J Interpers Violence* 10(1):85–105
- Sandler JC, Freeman NJ (2009) Female sex offender recidivism: a large-scale empirical analysis. *Sex Abuse J Res Treat* 21(4):455–473
- Serin RC, Mailloux DL, Malcolm PB (2001) Psychopathy, deviant sexual arousal, and recidivism among sexual offenders. *J Interpers Violence* 16(3):234–246
- Soothill K (2010) Sex offender recidivism. *Crime Justice* 39(1):145–211
- Soothill K, Francis B, Sanderson B, Ackerley E (2000) Sex offenders: specialists, generalists – or both? *Br J Criminol* 40:56–67
- Thornton D (2006) Age and sexual recidivism: a variable connection. *Sex Abuse J Res Treat* 18(2):123–135
- Vandiver DM (2006) A prospective analysis of juvenile male sex offenders: characteristics and recidivism rates as adults. *J Interpers Violence* 21(5):673–688
- Worling JR, Langstrom N (2003) Assessment of criminal recidivism risk with adolescents who have offended sexually: a review. *Trauma Violence Abuse* 4(4):341–362

Sexual Reconviction

- ▶ [Sexual Recidivism](#)

Sexual Reoffending

- ▶ [Sexual Recidivism](#)

Shame in Criminological Theory

Nathan Harris
Regulatory Institutions Network, Australian National University, Acton, ACT, Australia

Overview

The end of the 1980s marked a rediscovery of shaming in criminology. Criminological theories,

and most particularly John Braithwaite's theory of reintegrative shaming, reintroduced the concept of shaming into the criminological debate. At around the same time, criminal justice systems experimented with new interventions, such as restorative justice and shaming penalties, which both draw on notions of shaming, albeit in very different ways. While these developments are principally concerned with actions that are taken by authorities, or broader communities, towards offenders, they also highlight the importance of shame as an emotion. Use of the word shaming implies that the corresponding emotion is in some way critical, but what shame is and what its benefits are for social control are less clearly articulated. The application of shaming within criminal justice is not without controversy, and the merits of shame as an emotion have provoked considerable debate. The concept of shame management and an ethical identity conception of shame have both been proposed to clarify the nature of the emotion as well as its implications for criminology.

The Rediscovery of Shaming in Criminology

The relevance of shaming to the regulation of crime has a long history. Shaming is central to accounts of social control in anthropological descriptions of Polynesian and Asian societies as well as analyses of European criminal justice practices in earlier centuries. Equally well documented is a move away from shaming practices in European-based criminal justice systems during the century and a half until the 1970s. However, in 1980s and 1990s, interest in shaming underwent something of a revival, most particularly through the publication of reintegrative shaming theory, the rise of restorative justice, and an interest in the judicial use of shaming punishments.

Reintegrative Shaming Theory

A focal point for the revival of interest in shaming was publication of John Braithwaite's (1989) book *Crime, Shame and Reintegration*. In this

book it is argued that institutions of criminal justice as well as criminological theory have underestimated the importance of social disapproval. Braithwaite argues that to understand crime rates, we need to look beyond official mechanisms, such as penalties that are imposed by criminal justice systems, to the degree to which societies express disapproval of crimes. Strong social norms against criminality, which arise through community activism, are seen as critical to low crime rates because they engender a culture in which crime is unthinkable because people come to see it as abhorrent. The concept that is central to Braithwaite analysis is shaming, which he defines as "... all societal processes of expressing social disapproval which have the intention or effect of invoking remorse in the person being shamed and/or condemnation by others who become aware of the shaming" (Braithwaite 1989, p. 100). An important characteristic of this definition is that it does not limit itself to demeaning or humiliating forms of disapproval but seeks to encompass the full spectrum of ways in which disapproval might be expressed.

The fundamental distinction the theory makes is between stigmatization and reintegration. Stigmatization occurs when disapproval is directed at the person as well as at the offensive behavior, when the person is not treated with respect, when there is no ceremony to decertify the individual's deviant status, and where deviance is allowed to become a master status trait. As with labeling theories, it is predicted that stigmatization of offenders leads to greater re-offending. Being charged with a crime, found guilty of it in a court, and then sanctioned imposes a deviant identity on an individual because it ceremonially changes the position of the person within society and has important social implications, such as reduced employment opportunities. This critique of criminal justice asserts that once imposed, a deviant identity becomes a self-fulfilling prophecy and is almost irreversible: marginalization reduces the individual's access to legitimate opportunities

while increasing perceptions of injustice and the attractiveness of supportive subcultures.

However, in identifying reintegrative shaming as an alternative, Braithwaite diverges from the labeling tradition by rejecting the idea that stigmatization is an inevitable product of social disapproval. Reintegration can be seen to have occurred when shaming is respectful, distinguishes between the person and their actions, concludes with forgiveness or decertification of deviance, and does not allow them to take on a negative master status trait. One context in which this often occurs is in family life and the disciplining of children, where research shows that authoritative approaches are more effective than either permissiveness or authoritarianism. Another example Braithwaite cites is Japanese society, which is both high in shaming and high in reintegrative traditions and which has a remarkably low crime rate.

In arguing for the positive effects of reintegrative shaming, Braithwaite highlights two mechanisms. One of these is that reintegrative shaming is an effective deterrent, particularly when it comes from those who the individual is close to, because it poses a threat to relationships that are valued. Yet, reintegrative shaming also transcends the rational actor model of deterrence. The second mechanism, which Braithwaite suggests is more important, is that reintegrative shaming communicates that certain behaviors are morally wrong and thus builds internalized controls or conscience. Braithwaite (1989, p. 72) argues:

Shaming is more pregnant with symbolic content than punishment. Punishment is a denial of confidence in the morality of the offender by reducing norm compliance to a crude cost-benefit calculation; shaming can be a reaffirmation of the morality of the offender by expressing personal disappointment that the offender should do something so out of character.

Restorative Justice

Shaming has also been an important concept in the development of restorative justice programs, which exploded in popularity during the 1990s, and is now found in criminal justice, child protection, school, and prison systems in many

parts of the world. Restorative justice is an alternative to the criminal justice system that redefines the goals of justice as well as the way in which it is carried out. A defining principle of restorative justice is that an offense creates an obligation for offenders to repair the harm that has been caused (Zehr 1990). Unlike the principles of traditional justice that emphasize the importance of consistent and proportional punishment, the aim of restoration focuses attention on apology, reparation, and reconciliation. While Braithwaite's concept of shaming can be applied to many different kinds of interventions, restorative justice quickly came to be seen as the principle way of implementing reintegrative shaming. The broad goals of restorative justice as well as the practices associated with it are consistent with reintegrating offenders, and there is evidence that offenders perceive it as more reintegrative.

The dynamics of restorative justice interventions, such as family group conferences, victim offender mediation, or healing circles, are rich contexts for shaming. Family group conferences, for example, involve semiformal meetings between the offender(s), people who are close to them, the victim(s), and their supporters. The focus of a conference is on finding out what happened and how the incident has affected all of the parties as well as coming to an agreement about what needs to be done to repair the harms that are identified. As a consequence, they involve communities in the kinds of conversations about the negative consequences of crime that Braithwaite argues are critical to developing individual conscience and commitment to the law. Empirical observations suggest that shame dynamics do play an important role in conferences and that well-run programs have the potential to assist in resolving these feelings (Retzinger and Scheff 1996).

Judicial Shaming

Finally, the explicit use of shaming by courts has also seen the rise of "shaming" practices that are completely contrary to the restorative approaches

discussed above. Recent examples have occurred, particularly in American criminal justice, where shaming has been used in the court system as a deterrent or punishment for convicted offenders. Offenders have been ordered to complete "shame sentences" relevant to the crime they commit instead of spending time in jail. Shoplifters have been ordered to stand out the front of shops holding signs declaring that they stole, drink drivers are ordered to attach "DUI" stickers to their cars, while those convicted of soliciting sex are ordered to sweep the streets. An explicit aim of this kind of shaming is to humiliate offenders (see Kahan 1996). The significance that feelings of shame hold as a deterrent is exploited by these approaches, but the shaming they impart is explicitly stigmatizing. Another area of American criminal justice which reputedly incorporates shaming are boot camps where offenders are subject to military-style discipline but are also publicly confronted with their offense.

Concerns About Shaming

While awareness of shaming has increased, so too have concerns about the explicit use of shaming to control or respond to crime. Shaming punishments, in particular, have been seen by some as a regressive step which demeans the dignity of offenders while failing to protect basic human rights or allow for rehabilitation. Massaro (1997) argues that this "modern" kind of shaming is one that outcasts certain segments of society in a way that does not protect the individual and undermines the dignity of the whole community. In addition to arguing against the decency of this approach, she argues that the complexity of the emotion of shame is such that courts are ill-equipped to employ shaming and that the effect on offenders would be difficult to predict. Martha Nussbaum (2004) identifies five arguments in the literature against the use of shaming punishments: that they are an offense against human dignity, that they are a form of mob justice, that they are unreliable, that they don't hold the deterrent potential that

they are supposed to, and that they are potentially net widening.

While it is not surprising that questions have been raised about these overt forms of humiliation, the appropriateness of shaming within more reintegrative forums such as restorative justice has also been questioned. Maxwell and Morris (2002) and others have argued that overt disapproval is not an aim of restorative practices, suggesting instead that they are oriented towards exploring the consequences that an offense has on its victims, with the aim of provoking empathy. They argue that shaming is a dangerous proposition in restorative justice because even with the best of intentions, shaming might be interpreted by offenders as stigmatizing. Shaming a young offender may exacerbate problems rather than prevent re-offending, particularly if offenses have been committed as a consequence of low self-esteem, which has occurred as a consequence of an absence of emotional support or a difficult past.

This critique of shaming is based in part on doubt as to whether shame is a positive emotion for offenders to feel. A number of scholars have argued that the more important mechanism in restorative justice is the eliciting of remorse, which occurs as a consequence of the offender coming to understand the impact that their actions had on the victims. Empathy, as understood by Maxwell and Morris, aids in this process. Shame, on the other hand, is said to be a dangerous emotion to invoke in offenders because it is a threat to the offender's sense of self-worth and is potentially destructive. These questions reflect a broader debate about the virtues of shame as an emotion, in which there is a clear division between scholars who are pessimistic about the role the emotion plays and those who are more optimistic.

Conceptions of Shame in Criminology

As just illustrated, various ideas about shaming, both positive and negative, are based on assumptions, often implicit, about the nature

of the emotion that shaming invokes. This raises the following questions: what is shame, and what are its characteristics? While it is not possible to provide a neat typology because of the disparate manner in which theoretical approaches to the emotion have advanced, three broad characterizations of shame have been identified in the literature (Harris 2001). Each of these characterizations is reflected within criminological research.

Shame as a Social Threat

The first of these conceptions of shame conceives of the emotion as a response to social threat, which is precipitated by the individual's perception that they have been rejected or disapproved of in some way. This conception of shame is apparent in early anthropological perspectives which describe shame cultures as those that rely for social control on the sensitivity of individuals to negative perceptions of others, rather than through the development of conscience. This idea has been elaborated in various ways in contemporary research. While these approaches have varied in their explanations of why people are sensitive to social evaluation, they all emphasize the need to be accepted by others either because the need to have strong personal ties is a basic human motive or because there is an evolutionary need to maintain status or because shame is related to the person's perception of his or her own self-worth. An important characteristic of this conception is that it describes shame as exterior to individuals and as constraining. The individual feels shame as a result of another's decision to reject. If others do not reject in the face of the same actions, no shame is felt. Shame, or the fear of shame, is described as a powerful motivation for the individual to continually monitor and work on personal relationships and to comply with social expectations at a broader level.

As a consequence a fear of shame has been seen as a strong motivator for law-abiding behavior, and this assumption is explicitly made in early anthropological research on shame cultures. More recently, criminologists have drawn on this understanding of shame to

argue that informal social sanctions represent a significant deterrent to crime (e.g., Grasmick and Bursik 1990). A number of empirical studies, which place shame within a rational choice perspective, suggest that expectations of feeling shame are associated with lower self-reported projections of offending and in some cases that the effect is comparable with, or greater than, official sanctions.

Shame as Personal Failure

A second way in which shame is described in the literature is as a response to perceptions of personal failure. This is based upon the proposition that shame occurs when an individual perceives that they have failed to live up to an ideal or standard that they uphold and that the consequence of this is the perception that the “whole” self is a failure. This proposition has been explained using a number of theoretical frameworks including psychoanalysis, attribution theory, and affect theory. Shame is contrasted with guilt because the focus of attention is the self rather than an act or omission. Unlike the social threat conception described above, perceptions of failure are not necessarily prompted by disapproval, but can occur in isolation and in relation to personal ideals.

Research that explores the impact of a disposition to feel shame as defined by personal failure has been applied to criminology. June Tangney and her colleagues, in particular, have argued that a disposition to feel shame is far less adaptive than a disposition to feel guilt because shame involves an overwhelming negative evaluation of the self that prevents individuals from responding positively. An extensive program of research shows that individuals who are shame prone are more likely to feel anger and hostility, are less likely to feel empathy for others, and are more likely to suffer from psychopathology (see Tangney and Dearing 2002). When applied to inmate populations, shame proneness is correlated with substance abuse and does not seem to play the same protective role that guilt proneness does (Tangney et al. 2011).

Shame as Ethical Threat

The third conception of shame cuts across these two literatures by incorporating the notion that shame occurs when wrongdoing is recognized by the individual and their community. Shame in this view is connected, unlike embarrassment, with serious transgression as well as the idea of fault. The individual feels shame for having intentionally committed a wrong. This is implicit in William’s (1993) description of the precondition for shame being one in which a respected other, defined in ethical terms, would think badly of us. Taylor (1985) argues that shame is tied to the loss of self-respect, which defines what the individual feels is tolerable and what is not. Thus, unlike the social threat conception, shame is seen as occurring in response to the violation of internalized values, and the discomfort associated with the emotion concerns perceptions of the self. However, in emphasizing the ethical character of the emotion, this approach acknowledges the degree to which individual rely on others in forming their beliefs about what is right and wrong. Judgments about what is shameful are not externally imposed or arrived at in isolation, but they are socially negotiated.

This conception of shame is most clearly represented in criminological theory in Braithwaite’s theory of reintegrative shaming. Reintegrative shaming according to Braithwaite is “conceived as a tool to allure and inveigle the citizen to attend to the moral claims of the criminal law, to coax and caress compliance, to reason and remonstrate with him over the harmfulness of his conduct” (Braithwaite 1989, p. 9). This suggests that shaming is important because of its educative value in developing or reinforcing beliefs about what is wrong. While the theory suggests that shaming can have a deterrent effect, as an informal sanction that threatens the loss of respect by valued other, this is considered secondary to its moralizing qualities. Braithwaite argues that the primary reason individuals do not commit crime is because they have commitments to shared moral norms and social institutions. Punishment is irrelevant to most people because committing

serious crime is simply unthinkable to them. The socialization of children in families and schools about moral norms leads to a broad consensus about what acts should be crimes. In social contexts where there is a broad consensus that particular behavior is wrong and individuals are interdependent because of strong communities, those behaviors will be shameful. Individuals will feel ashamed for violating these values because, consistent with an ethical conception of shame, they have subscribed to these values.

Another approach that has likewise focused on shame's moral qualities is Wikström's (2004) situational action theory of crime, which argues that the emotion is a protective factor in preventing offending. Shame in this framework reflects the individual's commitment to do the right thing, which in turn influences their perception of the choices available in a given context. A number of studies have shown that juveniles who report that they would feel shame in front of others (e.g., friends) if they committed a crime also reported lower levels of delinquency.

Shame Management: The Different Faces of Shame

Evidence that feeling bad about one's actions can have both positive and negative consequences has turned attention to understanding why shame is a constructive emotion in some situations but is counterproductive in others. Why do we hope that some individuals feel shame for offenses they commit, yet also experience unease at the idea of imposing shame within criminal justice? A long tradition of research on shame emotions has explored variation in how individuals experience the emotion, and this, like more contemporary research on dispositions, has recently been drawn on to explore the notion of shame management (Ahmed et al. 2001). This theoretical perspective suggests that when confronted with feeling ashamed for their actions, individuals can manage or respond to the emotion in different ways and that this has important implication for criminal justice institutions.

Evidence of differences in shame experiences was first captured in the seminal work of psychiatrist Helen Block Lewis (1971). In her research with patients, Lewis identified three different forms of shame. The first, "acknowledged shame," involves the recognition that one feels shame and awareness of the feeling associated with it. "Overt-unidentified" shame describes the experience of feeling the negative emotion associated with shame but not recognizing it as shame and thus mislabeling it. "Bypassed" shame involves an awareness that an event may be shameful and doubt about how others see the self, but the emotion is bypassed leaving the person with "... an insoluble, plaguing dilemma of guilt thought which will not be solved" (Lewis 1971, p. 134).

One of the important findings from this work for understanding the implications of shame is that unacknowledged forms of shame are associated with feeling of anger and hostility towards others. Scheff and Retzinger (1991) extended Lewis's analysis by arguing that shame is a signal that the bond between the individual and others is threatened. When feelings of hurt associated with rejection are not acknowledged by the individual, as is the case in unacknowledged forms of shame (bypassed and overt-unidentified), then this emotion becomes redirected as anger towards the self and others. According to Scheff and Retzinger, this is the cause of humiliated fury and helps to explain not just individual anger but also conflict between nations. In both Lewis's and Scheff's accounts of shame, it is evident that when shame is not acknowledged, it manifests itself in an unhealthy reaction.

Eliza Ahmed and her colleagues (2001) have described the various manifestations of shame through the concept of shame management. This captures the notion that when confronted with a shame-inducing situation, individuals can manage the negative feelings in a variety of ways and that this is influenced by both individual characteristics and the social context. Acknowledged shame occurs when the individual accepts that they are responsible and thus acknowledges the emotion. It is argued

that when shame is acknowledged, the person is more likely to make amends, feels less anger towards others, and is more likely to discharge the negative feelings. In contrast, unacknowledged shame, which Ahmed describes as displaced shame, occurs when the person does not accept that they are responsible. Failure to resolve the emotion, because of the tension between the disapproval of others and this denial of responsibility, results in shame being displaced into anger towards others.

There is growing empirical evidence that shame management predicts both bullying and criminal behaviors. Ahmed's own research in Australia and Bangladesh shows that children who are bullies are more likely to displace shame compared to children who haven't bullied, who are more likely to acknowledge shame feelings (Ahmed and Braithwaite 2006; Ahmed et al. 2001). These results have been partially supported in a study by Ttofi and Farrington (2008) which showed that displaced shame was a predictor of bullying, but did not demonstrate the expected relationship between shame acknowledgement and lower bullying. A similar result was found by Murphy and Harris (2007) in the context of white-collar crime. In this study, shame displacement predicted recidivism, but the relationship between shame acknowledgement and recidivism was mediated through a measure of remorse.

This research on shame management has significant implications for reintegrative shaming theory and has prompted a revision of the theory (Ahmed et al. 2001). While the revision does not alter the theory's prediction that reintegrative shaming reduces offending (while stigmatic shaming increases offending), it does clarify why this is the case as well as the role that shame plays. The original formulation of the theory implies that the benefit of reintegrative shaming is that it leads to greater feelings of shame in offenders. However, the implication of shame management is that reintegrative shaming reduces offending in most cases because it allows offenders to manage feeling of shame more constructively. Reintegration is more likely to result in offenders acknowledging shame, feeling

remorseful for what has happened, and making amends. Stigmatization, on the other hand, is more likely to result in offenders displacing shame and feeling anger towards others. Thus it would seem, somewhat ironically, that the benefit of reintegrative shaming is that it allows offenders to resolve and diminish any shame that they feel.

The implication for criminal justice procedures that have drawn on reintegrative shaming theory, like restorative justice, is that the kinds of shaming or disapproval that are effective in those setting do not involve explicit attempts to shame offenders (Retzinger and Scheff 1996). Overt forms of disapproval, such as a focus on participants expressing their dislike of the behavior, may undermine the offender's ability to acknowledge and resolve feelings of shame. This may be exacerbated in those situations where it is culturally inappropriate to overtly disapprove of another or where offenders have already acknowledged wrongdoing. Processes like family group conferences instead focus on discussion of the consequences of the offenses for all parties. By focusing on how people have been hurt, they avoid stigmatization of the offender and allow them to express remorse and to make amends. Both of these behaviors, as well as acknowledgement and forgiveness by others, are important mechanisms for resolving shame.

An Ethical-Identity Conception of Shame

A critique of the empirical research on shame is that it fails to adequately explain the complex relationship between the individual and the social contexts in which shame occurs, either conceptualizing shame as a response to values that are extrinsic to the person (social threat conception) or having little to say about the social context at all (personal failure conception). Neither of these conceptions adequately accounts for repeated observations that shame is both intimately tied to identity and sensitive to disapproval of others.

To better explain the social context in which shame occurs, as reflected in the empirical research conducted on shame management, an ethical-identity theory of shame has been proposed, which draws on both the ethical conception of shame discussed earlier as well as insights from social psychology (Harris 2011).

The central claim of this approach is that the defining characteristic of shame is that it occurs when we experience a threat to our ethical identity. The precondition for this to occur is an awareness that we violated an ethical value that we subscribe to: we realize that we have behaved in a way that we feel is wrong or at least have significant doubts about how acceptable our behavior was. Even though shame is experienced in reference to internalized values, it is also sensitive to the opinions of others, firstly in making us aware that we might have violated a value that is important to our sense of self, but also because the opinions of others (those whose views we respect) contribute to our interpretation of our behavior.

Research on social influence and conformity suggests that a simple dichotomy between our own and others' values is too crude. A long history of research in social science demonstrates that values, attitudes, and beliefs held by individuals are influenced by others (Turner et al. 1987). Others' opinions are important because, as is illustrated by social identity theory, our values, attitudes, and beliefs, all part of our identities, are often shared with others. We expect to agree with those people whom we see as similar to ourselves, and it is disconcerting when we do not. The reason why disapproval results in shame is because it acts as a form of social validation, either reinforcing the belief that what we did was shameful or undermining the assumption that it was not.

Shame is linked to identity because violating one of our values only leads to shame if it undermines our sense of who we are. This threat to our ethical identity occurs because there is a reciprocal relationship between our values and

who we think we are. Holding certain values is at the heart of personal or social identities because identities are defined in large part by sets of beliefs (e.g., "being nurturing" would be an important value if "mother" was an important identity). It follows that when we become aware that we have acted contrary to our values, our identity is called into question. The painful feelings of self-awareness, anger at oneself, and confusion that are associated with shame occur because the contradiction between our values and our behavior cannot be easily reconciled.

An important characteristic of shame is that it motivates us to resolve this dissonance between our actions and how we think of ourselves, and we are able to do so in a variety of ways. We can resolve the threat to our ethical identity by diminishing the significance of our behavior (either by concluding that there was a good excuse or reason for what we did or by making up for it in some way); we can respond by perceiving ourselves as having an alternative identity that is consistent with our behavior and in doing so rejecting any disapproval, or we may perceive ourselves as defective because we have failed to live up to our values.

It is hypothesized that the way in which we resolve this threat to ethical identity will also be influenced by the social validation we receive from others. This explains why individuals react differently to reintegrative shaming and to stigmatizing shaming. Disapproval of our behavior that is reintegrative, when it comes from those whom we respect, can reaffirm a positive ethical identity and thus encourage us to apologize and repair what we've done so as put our misdeeds within a larger positive story about oneself (Maruna 2001). Stigmatization makes this less likely because the message from others is that there is something defective about who we are. In this situation we are more likely to defend our ethical identity by deciding we had a valid reason for our behavior, we might feel very bad about ourselves, or we might even decide that we have fundamentally different values to those who disapprove of us and embrace a deviant identity.

Conclusion

An important issue that has been at the center of emerging research on shame is whether it should be understood as a productive and useful emotion that allows offenders to reconsider their behavior or whether it is a dangerous and unhelpful emotion that may even promote greater offending. This question has been complicated by the various notions about what shame is and the various ways in which it has been incorporated into criminological research. However, it does seem clear that the answer is unlikely to be simply one or the other. This is unsurprising once we accept that shame is integrally connected to our belief systems, to our identities, and to our relationships with those around us. Research on shame management and an ethical-identity conception of the emotion of shame have sought to delineate the various ways in which individuals can respond to shame and have examined the consequences of these responses for both the individual and society.

Recommended Reading and References

- Ahmed E, Braithwaite V (2006) Forgiveness, reconciliation, and shame: three key variables in reducing school bullying. *J Soc Issue* 62(2):347
- Ahmed E, Harris N, Braithwaite J, Braithwaite V (2001) Shame management through reintegration, Cambridge studies in criminology. Cambridge University Press, Melbourne
- Braithwaite J (1989) *Crime, shame and reintegration*. Cambridge University Press, Cambridge, UK
- Grasmick HG, Bursik RJ (1990) Conscience, significant others, and rational choice: extending the deterrence model. *Law Soc Rev* 24(3):837–861
- Harris N (2001) Shaming and shame: Regulating shame drink-driving. In Ahmed E, Harris N, Braithwaite J, Braithwaite V (eds) *Shame management through reintegration*. Cambridge University Press, Cambridge
- Harris N (2011) Shame, ethical identity and justice interventions: lessons from research on the psychology of influence. In: Karstedt S, Loader I, Strang H (eds) *Emotions, crime and justice*. Hart Publishing, Oxford, pp 193–209
- Kahan DM (1996) What do alternative sanctions mean? *Univ Chicago Law Rev* 63:591–653
- Lewis HB (1971) *Shame and guilt in neurosis*. International Universities Press, New York
- Maruna S (2001) *Making good: how ex-convicts reform and rebuild their lives*. American Psychological Association, Washington, DC
- Massaro TM (1997) The meanings of shame: Implications for legal reform. *Psychol Public Policy Law* 3(4):645–704
- Maxwell G, Morris A (2002) The role of shame, guilt, and remorse in restorative justice processes for young people. In: Weitekamp E, Kerner H-J (eds) *Restorative justice: theoretical foundations*. Willan Publishing, Cullompton, pp 267–284
- Murphy K, Harris N (2007) Shaming, shame and recidivism: a test of reintegrative shaming theory in the white-collar crime context. *Br J Criminol* 47(6):900–917
- Nussbaum MC (2004) *Hiding from humanity: disgust, shame, and the law*. Princeton University Press, Princeton
- Retzinger SM, Scheff TJ (1996) Strategy for community conferences: emotions and social bonds. In: Galaway B, Hudson J (eds) *Restorative justice: international perspectives*. Criminal Justice Press, New York
- Scheff TJ, Retzinger SM (1991) *Emotions and violence: shame and rage in destructive conflicts*, Lexington books series on social theory. Lexington Books/D. C. Heath and Company, Lexington
- Tangney JP, Dearing RL (2002) *Shame and guilt*. The Guilford Press, New York
- Tangney JP, Stuewig J, Mashek D, Hastings M (2011) Assessing jail inmates' proneness to shame and guilt. *Crim Justice Behav* 38(7):710–734. doi:10.1177/0093854811405762
- Taylor G (1985) *Pride, shame and guilt: emotions of self-assessment*. Oxford University Press, Oxford
- Ttofi MM, Farrington DP (2008) Reintegrative shaming theory, moral emotions and bullying. *Aggress Behav* 34(4):352–368. doi:10.1002/ab.20257
- Turner JC, Hogg MA, Oakes PJ, Reicher SD, Wetherell MS (1987) *Rediscovering the social group: a self-categorization theory*. Basil Blackwell, New York
- Wikström P-O (2004) Crime as alternative: towards a cross-level situational action theory of crime causation. In: McCord J (ed) *Beyond empiricism*. Transaction Publishers, New Brunswick
- Williams B (1993) *Shame and necessity*. University of California Press, Berkeley
- Zehr H (1990) *Changing lenses: a new focus for criminal justice*. Herald Press, Scottsdale

Shaming in Asian Societies

Hong Lu, Wyatt Merritt and Mari Sakiyama
University of Nevada, Las Vegas, Las Vegas,
NV, USA

Synonyms

[Restorative justice](#)

Overview

Shaming plays a critical role in crime and punishment. Within the context of criminological theories, there are two different styles of shaming: stigmatizing and reintegrative. Shaming that stigmatizes is known as labeling. Through labeling theory, it has been widely recognized within criminological circles that having a negative criminal label (a stigma) contributes to the formation of a deviant self-image, which pushes offenders into criminal subcultures, facilitates the internalization of the deviant identity, and leads to secondary deviance (Becker 1973; Lemert 1951). In contrast, reintegrative shaming combines punishment with compassion and aims to rehabilitate and reintegrate individual offenders and restore broken relationships. While both types of approaches to crime and punishment have been practiced throughout human history, only relatively recently (since the 1980s) has restorative justice and reintegrative shaming caught the attention of scholars and policy makers in Western countries such as in the United States.

This entry first defines and describes shaming and its different forms. It then examines unique social conditions that are conducive to reintegrative shaming. Research has shown that unique social conditions, such as a high interdependency among individuals and communitarianism within a society, facilitate the use of reintegrative shaming (Braithwaite 1989). Asian countries, such as China and Japan, have been widely regarded as communitarian societies, and shaming has played an important role both in informal and formal social settings. To help illustrate the characteristics and nature of reintegrative shaming practiced in Asia, crime prevention and control programs, such as *bang-jiao* in China and *bizai shobun* in Japan, are assessed within the context of the individualism and communitarianism debate.

Shame and Reintegration

Within the context of criminology, shame refers to the disapproval of deviant acts and/or criminal

wrongdoings by others and often in public. Shaming can take different forms. It can be as subtle as a frown, a slight shake of the head, or gossip. It can also be as direct as a verbal confrontation, a media broadcast, or an official pronouncement by a judge (Braithwaite 1989). Shame serves as an effective social control on two levels: (1) shame deters future criminal behavior through internally building conscience (guilt induction) and (2) shame also deters future criminal behavior by public disapproval through external referents (e.g., family members, teachers, neighbors, police officers) that set boundaries and reaffirms rules.

Developmental psychologists sometimes make the distinction between shame and guilt by depicting shame as a reaction to criticisms by others such as parents and neighbors, whereas guilt is induced internally after a wrongdoing (Braithwaite 1989). Others regard shaming and guilt induction as inextricably part of the same social process. They both imply certain moral expectations of the individual and the community.

Scholars argue that a culture that has a high set of moral expectations and reinforces these moral expectations through shaming will have a more effective social control than a culture that seeks control through strict law and punishment (Braithwaite 1989). This is because shaming implies that citizens have the responsibility to express public disapproval of behaviors that harm the community. The potency of this form of self-policing lies in its swiftness (in time) and intimacy (both in physical distance and social distance).

A new form of shaming – online shaming – takes shaming to the virtual world. Witnesses of an immoral/deviant incident, often out of frustration and outrage, may take pictures and post them online. For example, pictures of a South Korean woman who refused to clean up after her dog defecated on a train were posted online. This resulted in her being “shamed” by bloggers from all over the world. This newly emerged shaming obviously grew from modern technology but also out of the growing diversity, mobility, and fear of disorder and crime shared in

modern societies. Online shaming essentially serves the function of peer surveillance, or lateral surveillance, and may help deter deviant behaviors and reaffirm moral boundaries (Skoric et al. 2010).

There are two types of shaming: reintegrative and stigmatizing. A reintegrative type of shaming has two essential components: (1) shame (public expressions of community disapproval) and (2) reintegration (efforts to reaccept the offender into the community). Community disapproval can range from a mild rebuke to a strong disapproval within a formal ceremony. Also, efforts at reintegration can range from a simple smile expressing forgiveness and love to a formal ceremony to decertify the offender as a deviant. Shaming is more potent when expressed and administered by significant others – parents, friends, and neighbors. This is because significant others tended to have a continuing and meaningful relationship with the deviants; thus, their opinions bear more importance to the deviants than would a stranger's or an institution's (i.e., the police, court officials). In addition, shaming is more effective when there is a goal of reintegration and restoration, rather than stigmatization and rejection. The best place to witness reintegrative shaming is the loving family. Within a loving family, disapproval is expressed while bonds of respect are maintained. Condemnation is made toward the deviant act, not the deviant family member. More important is the forgiveness and genuine gesture of reacceptance that leads to reintegration back into the family (Braithwaite 1989).

The stigmatizing type of shaming is disintegrative. It typically involves attaching a deviant label to the offender. Labeling is stigmatizing because it focuses on the public degradation process, the labeling of the deviant, yet with little attention on delabeling and reintegration. The deviant person, once labeled and stigmatized, is likely to become an outcast and pushed into the criminal subculture. The deviant behavior becomes their master status. A master status is a characteristic(s) that defines an individual. For example, a person could be a good athlete or a good student, but if they are caught for engaging

in delinquent behavior and are labeled, this label trumps what they were characterized as being before. Consequently, it precipitates the outcast to internalize a criminal identity and reject any attempts at shaming and reintegration. This self-fulfilling prophecy happens as a response to shaming that dramatizes the offenders' evil.

Criminologists of divergent perspectives (e.g., rational choice, social learning, labeling) now recognize that most criminals do not reject criminal law outright. Instead, they rationalize their deviant actions by shifting the blame of their deviance onto others (e.g., "it's the victims' fault," "everybody else is doing it," "I did not mean it"). The ability to rationalize their deviant behavior allows the deviants to temporarily suspend their commitment to the law and insulates them from shame. In this process of rationalization, criminal subcultures provide the much-needed social support for the deviant behavior, particularly when legitimate opportunities are blocked due to negative labels. Within the world of criminal subcultures, deviant actions are rationalized as a defensible lifestyle – against the injustice and stigma often felt inflicted upon the outcasts. Deviance thus becomes a way of life. The more attempts that are made to reform the outcasts, the worse it becomes (Braithwaite 1989; Becker 1973; Lemert 1951).

Some societies are prone to reintegration, whereas other societies prefer punishment as a way of social control. In general, societies characterized as communitarian, with reciprocal relationships among its citizens, are more likely to shame reintegratively. Shaming within this context is likely to be more potent, resulting in lower crime rates, because disapproval is unlikely to be rejected. In contrast, societies characterized as individualistic, with its citizens interacting only out of necessity and convenience, are more likely to resort to punishment. Shaming in this context is less effective because stigmatization and punishment are likely to cut off the deviants' interdependencies with their mainstream social relations and push them into the criminal subcultural world.

In addition, shaming is culture specific. Similar public expressions of disapproval of a deviant

act may be perceived differently. In the United States, for example, the same outward disapproval of a delinquent act in an Asian culture setting may be received differently than it would in an African American culture setting. This is based on the concept that the ideals and attitudes held by different cultures are not the same in regards to legal justice, norm violations, and the extent of the impact by external forces (Braithwaite 1989). Even within the same cultural context, different social organizations (i.e., family, neighborhood) may deliver various degrees/kinds of shaming. For example, a study on Chinese urban residents found that family members generally use shaming and reintegration highly when a family member deviated because acts of deviance reflect badly on the rest of the family and their ancestors. With the economic development since the 1980s, Chinese neighborhoods are transforming to be become more transient. The study found that residential mobility affected shaming practices in the neighborhood (making shaming less reintegrative) much more significantly than shaming practices within the family (Lu et al. 2002).

In sum, reintegrative shaming is likely to label only the deviant acts, whereas stigmatizing shaming is likely to label the actor. Reintegrative shaming is more effective at crime control than stigmatizing shaming, because reintegrative shaming minimizes the risks of pushing deviants into criminal subcultures, whereas stigmatizing shaming increases these risks. In addition, deviants who are embedded in relationships that are overwhelmingly characterized by social approval are more likely to respond positively to shame. In contrast, deviants, who are not in these types of relationships, are more likely to reject shaming and make it criminogenic rather than crime inhibiting.

Shame and Communitarianism

The effectiveness of shaming depends upon two interrelated social conditions: interdependency and communitarianism. Interdependency refers to the interrelationship among individuals. It is

an individual level variable. Individuals with more interdependencies are more susceptible to shaming; thus, they are less likely to commit crime.

There is an overwhelming consensus among criminologists regarding factors critical to an individual's interdependent relationships. These factors include, but are not limited to, age, gender, marital status, education, and employment status. For example, studies have found that the most important correlate of interdependency is stage in the life cycle. Individuals are mostly likely to become a deviant between 15 and 25 years of age. The two life-defining events for deviants to go straight are getting married and having children (Braithwaite 1989).

Communitarianism is the antithesis of individualism. It is the combination of a dense network of individual interdependencies with strong cultural commitments to mutuality of obligation. There are three basic elements to communitarianism: (1) interdependencies densely enmeshed in all spheres of social life, (2) interdependencies characterized by mutual obligation and trust, and (3) interdependencies built upon group loyalty (Braithwaite 1989). Thus, while numerous interdependencies are a necessary condition for communitarianism, mere interactions among group members are not sufficient for the group/society to be characterized as communitarian.

In communitarian societies, communities are typically defined by social ties, such as in family or neighborhood, as well as other significant relationships such as fellow students and coworkers. These communities are held together by mutual dependency and the need for cooperative endeavor (typically in a strict hierarchical order based on an individual's group status and/or social roles) to function. Countries and/or tribes in many parts of the world, such as Asia and Africa, can be characterized as communitarian. Most Western societies may be identified as more individualistic rather than communitarian. As the economy progresses, societies become more urbanized, mobile, diverse, and individualized. While there may be an abundant amount of interdependent relationships in highly developed areas (i.e., Western societies), because of

increasing economic activities (i.e., landlords and tenants, bank tellers and customers) and other activities accompanying the economic growths (i.e., the police and criminals), these interdependencies may be more out of necessities and convenience, opposed to mutual obligation and trust. Within this context of high individualism, shaming by communities becomes particularly challenging. This is because “As a society becomes more role-differentiated [as in a modern society], the potential for effective shaming increases in important ways, but so does the potential for stigmatization that cuts off effective shaming” (Braithwaite 1993, p. 15).

Shaming can be administered by private individuals (i.e., communitarian-based shaming), and it can also be administered by the state (i.e., state-sanctioned shaming). In most developed Western countries, shaming is handled by the state via the criminal-justice system. In contrast, communitarian based shaming is dominant in many developing nations where legal systems are not fully developed and accessible.

Studies have found that communitarian-based shaming is more effective than state-sanctioned shaming in two aspects. First, shaming delivered by people of high interdependency and within a communitarian society (i.e., family, friends) is more potent than when it is delivered by an impersonal state because of the constant contact the intimates have with the individual. Second, shaming by significant others is more likely to be expressed in ways that is reintegrative, compared to the shaming by an impersonal state. Offenders are less likely to be stigmatized and adopt an outcast master status. In this context, a communitarian society resembles an enlarged loving family, where reintegrative shaming works best.

Reintegrative Shaming and Its Effectiveness

Since its publication in 1989, the theory of reintegrative shaming has aroused great interest among scholars and policy makers. Several studies tested this theory in a variety of settings.

A study using adolescents (ages from 14 to 17) from a single urban area in a southwestern state of the United States examined the relationship between the adolescents’ perceptions of their parents’ sanctioning methods (reintegrative vs. stigmatizing) and their reports of predatory delinquency (Hay 2001). Findings of this study are largely consistent with Braithwaite’s theory of reintegrative shaming in that the level of interdependency of the adolescent with the parents and the use of reintegration by the parent were statistically significant. Parental sanctions are a way to reinforce a close relationship within families that have strong interdependent relationships between parents and children.

Reintegrative shaming theory has been used as a valuable interpretive framework in some experimental studies. For example, a study on a drug court in a southwest state of the United States revealed that drug courts were perceived to follow more of the reintegrative shaming type of model than a traditional court that offered more of an assembly line type of service. However, drug court participants had a higher recidivism risk than nondrug court participants. A compelling explanation for the failure of drug courts is because the presumed comparability with RST in process and structure is not necessarily true. Drug courts may be more stigmatizing than general courts based on field observation data (i.e., the initial and second appearances in courts resembled more of a public degradation ceremony than reintegrative shaming). In addition, disapproval expressed by the judge may not be potent. Judges represent an impersonal state and may not be regarded by offenders as someone respectable or important. Furthermore, judges do not continue to have relationships with the drug users after seeing them in court (Miethe et al. 2000). In sum, the RST can be used to guide programs that are intended to restore offenders back into the community and can help those who manage and evaluate the programs to make them more successful.

Several studies examined family group conferences in Australia using reintegrative shaming theory. For example, a study examined basic conditions for a successful restorative/

reintegrative family conference and identified a list of key factors. First is separating the event and the offender. While the event may be defined as irresponsible, wrong, or criminal, the offender is ought to be supported and viewed as a whole person. Second, those who administer reintegrative shaming must represent all the interests involved (i.e., offender, victim, and the law) to mete out a fair and effective solution to the problem (i.e., the final goal is reintegration and restoration). Third, apology, repentance, and forgiveness must occur in order to terminate the separation of the offender and victim and to enforce reintegration rituals (Braithwaite and Mugford 1994). Another study on restorative justice conferences suggests that these types of conferences tended to have a greater and positive psychological impact on participants' view on the legitimacy of the law (Tyler et al. 2007).

Reintegrative shaming theory has also been applied to a study on nursing homes' compliance to regulatory rules in Australia. The study showed that the nursing homes that had inspection teams who had a reintegrative shaming philosophy (high disapproval and high reintegration scores) showed greater improvement in regulatory compliance, compared to inspection teams having other philosophies (i.e., tolerance [low on disapproval and high on reintegration] or stigmatizing [high on disapproval and low on reintegration]). In addition, the more interdependency between regulators and the nursing home managers, the more compliant the nursing home managers are toward the regulations (Makkai and Braithwaite 1994).

Despite the vastly different social conditions in Western and Eastern countries, scholars have also attempted to examine the validity of the claims made by reintegrative shaming theory in Asian countries. Using a sample of 1,725 adolescents (ages 11–17) in China, one study attempted to identify the general effect of reintegrative shaming theory and the predictive effect of delinquency disapproval (shaming) and forgiveness (reintegration) on delinquency involvement. The study found that parental shaming, parental forgiveness, and peer shaming had reduced the involvement of predatory offenses. However,

when the interactive effect of shaming and forgiveness was introduced in the multivariate analysis, reintegrative shaming theory (represented by the interaction term of shaming and reintegration) did not appear to be a significant predictor of predatory offense involvement (Zhang and Zhang 2004).

The study in China, described above, generated somewhat mixed results about the theory. Nevertheless, it is generally believed that Asian cultures, such as in China and Japan, which are influenced by Confucianism and centered on family, frequently use shaming and guilt induction as means of informal and formal social control (Braithwaite 1989). This widespread use of shaming within the family model has been regarded as contributing to the low crime rates in these countries. In the next section, two unique crime prevention and control programs, *bang-jiao* in China and *bizai shobun* in Japan, are described in order to illustrate how the principles of reintegrative shaming work in these unique sociocultural contexts.

Reintegrative Shaming in Asia

Compared to individualism shown in Western countries, many of the countries in Asia are based on ideas of communitarianism. Due to strong Confucian influence, Southeast Asia in particular (e.g., China, Japan, Korea, Taiwan) holds high regards for social status, hierarchy, order, peace, and harmony. These values help promote conformity and submission to group interests. In contrast, the values fundamental to Western ideas of individualism – individual rights and freedom – are only placed in a relative context in these Asian countries.

As a result, a communitarian society prefers informal social control to formal social control. Collective citizen action is considered an effective strategy for maintaining peace and harmony, solving disputes, controlling crime, and reducing the fear of crime. To foster an effective informal social control, citizens are taught these social responsibilities through socialization, which is to behave appropriately according to group/social

norms and rules. This takes place first and foremost within the family, then at school, in the neighborhood, and in the workplace. In addition, citizens are expected to help others observe the rules of ritual propriety and help educate and correct behaviors that deviate from the norms. This network of informal social control is largely operated on reciprocal interdependencies with mutual respect. The *bang-jiao* program widely practiced throughout urban neighborhoods in China represents such an informal social control model.

Bang-jiao and Reintegrative Shaming in China

One of the major differences between social control in China and Western countries is that the Chinese attempt to control both the behavior and minds of the people. In contrast to the Western concept of original sin, Confucianism assumes the goodness in people. Crime is thus viewed as the result of the environmental influence. To offset the negative influence of the environment and help individuals get in touch with the good of their inner selves, Confucius and his followers believed in the importance of the rule of *li* (moral code). This was because *li* was viewed as essential in fostering an internalization of the basic moral principles, promoting voluntary compliance to the rule, and being virtuous and good. Confucius also believed in the power of education in shaping people's thoughts and behaviors. Moral awakening through thought education is considered the primary stabilizer of society (Chen 2002).

In this process of moral education, shaming is regarded as playing a critical role – a role that calls attention to the damage the deviant/criminal has done to the victim/community and a moral imperative to correct the wrongs and amend the broken relationships caused by the deviant/criminal act. To the Chinese, shaming can generate both negative (e.g., stigma) and positive (e.g., deterrence, rehabilitation) results, depending on the strategies (Chen 2002).

One strategy is early intervention. Nipping crime in the bud is common wisdom rooted in Chinese philosophy. Early intervention only

requires a small dosage of shaming and small educational efforts could result in full reintegration. Even though early intervention may run the risks of intrusion into other people's lives, it is taken for granted and viewed not as meddling, but caring.

Another strategy is popular participation. Chinese social control is from the bottom up, not the top down. The Chinese prefer to handle their own crime and delinquency problems in their community rather than leaving them to the professionals. The popular participation approach ensures a swifter and more effective crime control service because those who deliver the service presumably have more intimate knowledge about the deviant person as well as the act and have more stakes in the quality of service.

The total approach is yet another strategy used in crime prevention and control. As China moves toward legalization and professionalization, greater emphasis is given to law and the legal professionals in addressing major legal issues such as crime. Nevertheless, legal professionals, such as the police and judicial officers, are expected to use all sorts of means (e.g., legal, administrative, and social) to prevent crime and reintegrate offenders. *Bang-jiao* in urban neighborhoods of China represents one such program.

Bang-jiao literally means, assisting, helping, guiding, and directing offenders, especially juvenile offenders (Zhang et al. 1996). It is a community-based program that utilizes remedial and preventative measures for controlling crime. Even though no particular group of individuals are excluded from the objects of *bang-jiao*, *bang-jiao* typically handles predelinquents, delinquents, deviants, and offenders who have committed minor offenses and are without a prior record. These groups of people are targeted, because they are the mostly likely to be “helped” (*bang*) and “educated” (*jiao*) successfully (13–28 years of age). Most *bang-jiao* programs are situated in urban neighborhoods, and they consist of parents, relatives, neighbors, teachers, coworkers, resident committee members, and/or local police who work as a team in carrying out *bang-jiao*.

Even though *bang-jiao* is community based and represents an extra legal measure to crime

prevention and control, it still has some basic principles guiding its program. According to the Juvenile Delinquency Prevention Law (1991), *bang-jiao* must be conducted based on the following principles: (1) fairness and equality, (2) practical guidance, and (3) response to genuine repentance with love, emotional support, and sincere acceptance of deviants back into the community (Zhang et al. 1996).

The characteristics and process of a typical *bang-jiao* intervention can be illustrated with the following example. A teenage boy is sent to a work-study school for rehabilitation because of repeated thefts. After recommendations from his *bang-jiao* team and work-study school principals, he was transferred to a normal school. Every day he came home after dark, because he was too ashamed to be seen by his neighbors. The boy's mother finally talked to the *bang-jiao* team members. The team members visited the neighbors and found out that the neighbors wanted to communicate with the boy, not ridicule him, but did not know how. The *bang-jiao* team discussed how to overcome these barriers with the neighbors and decided to hold a community event in which they knew the boy was interested in and would participate in (community painting exhibition). The boy was invited and at the event he was highly praised for his work. This experience changed the boy, and he started being involved with community events and being a part of the community (Braithwaite 2002).

From a theoretical standpoint, reintegrative shaming facilitates this process of *bang-jiao*. Although little empirical research has examined the effectiveness of *bang-jiao* in deterring crimes and reducing recidivism rates, one study based on self-report data suggests there is a significant negative relationship based on reports by inmates living in communities with a *bang-jiao* program (Zhang et al. 1996). More specifically, the inmates who lived in neighborhoods with an active *bang-jiao* program reported a significantly lower likelihood of recidivism compared to those who lived in neighborhoods without an active *bang-jiao* program.

Even though crime rates, including juvenile delinquency rates, have surged since the

economic reforms in the 1980s in China, the majority of juvenile delinquents were able to successfully have their deviant label removed and be reintegrated back into the community. This was supported with the low recidivism rate of between 8 % and 15 % annually in recent years (Chen 2002).

Bizai Shobun and Reintegrative Shaming in Japan

Similar to the Chinese view of the relationship between the individual and society, the Japanese see individuals not as an isolated entity, but as part of the network. There is a kind of social web that binds individuals into a collective. Within this largely communitarian society, punishment is not geared toward retribution, but rehabilitation and restoration.

Due to the importance of status and family to an individual, shaming has different meanings in the Japanese culture. Its potency can be different as well, particularly when compared with Western countries.

When assessing punishment, one of the criteria used in Western culture is severity of punishment (proportionality of punishment to crime), typically measured by the length of prison sentences and/or the amount of monetary fines imposed on the convicted offenders. In Japan, however, the focus is different. Contrary to isolation and punishment of the individual wrongdoer, the Japanese prefer restitution and restoration. Numerous terms, such as benevolent, paternalistic, and familiar, have been used to describe this reintegrative shaming practice unique to the Japanese. Among these different terms, an individualized decision-making model seems to best capture the Japanese crime control and prevention practice (Foote 1992).

The Japanese criminal-justice system is benevolent in that its goal is to achieve reformation and reintegration of offenders into society through lenient sanctions tailored to the offender's particular circumstances (Foote 1992). This benevolent-paternalism model emphasizes love and mutual respect so that the best interest of the offender and his/her rehabilitation is the focus of any decisions. The success of this family model

depends on officers' use of widespread power of discretion. It not only permits but also expects public officials to use discretion and take a preventive approach to crime and disorder. The major factors officers consider include family circumstances, employment status, and other types of support mechanisms available to the offender and, to a lesser extent, satisfaction of the victim.

Central to the Japanese practice of restoration and reintegrative shaming are concepts of apology, repentance, and forgiveness. To apologize, the wrongdoer must first confess to the crime. The importance of confession in the Japanese criminal-justice system can hardly be overstated. Numerous scholars of the Japanese legal system concur that confession, apology, and begging for forgiveness are essential elements at virtually every stage of the criminal process in Japan. Confession and admission of guilt are not used as part of the plea bargain, but are regarded as a gesture of complete submission to authorities in Japan. Confession, accompanied by sincere apology and remorsefulness, is viewed as having both the probative and correctional value. It aids the police, the judge, and the correctional officers to better achieve the goals of holding criminal offenders accountable for their wrongdoings and reforming them into law-abiding citizens.

The ability of legal officials to exercise widespread discretionary power lies in the great faith that is bestowed on public officials. Legal officials enjoy great autonomy and widespread power sometimes beyond the legal spheres in Japan. When dealing with an offender, the classification scheme of legal (e.g., offense severity, prior criminal record) vs. extralegal (e.g., age, gender, employment) factors commonly used in Western countries may not be relevant in Japan. Instead, means of punishment/treatment will be considered favorably if they are least disruptive to the deviants'/criminals' life, fit individual circumstances, and most likely lead to the successful reintegration back into the community.

Reintegrative shaming as a punishment philosophy has been institutionalized in Japan. All three authorities in the criminal-justice system are given widespread discretion when disposing of cases. For example, the police have the

authority not to report minor offenses (*bizai shobun*) (Code of Criminal Procedure, art. 246), the prosecutors have the authority to suspend prosecution (Code of Criminal Procedure), and the courts have the authority to suspend execution of sentences (Criminal Code, art. 25).

Below the practice of *bizai shobun* is summarized to show how the police, in exercising their discretionary power, translate the philosophy of reintegrative shaming into practice.

The Japanese police enjoy widespread discretion when performing their crime prevention and law enforcement duties. For example, the Japanese local police have the authority to conduct residential surveys twice a year for the purpose of getting to know the residents, registering new comers, and preventing future crimes. The Japanese police officers may stop and question anyone in public without reasonable suspicion or probable cause. They also have the authority to interrogate a criminal suspect for a lengthy period of time without a warrant. The relative "crime-free society," combined with high citizen respect and cooperation in Japan, make the police even more powerful and free in using the tools at their disposal when dealing with criminal suspects. The Japanese police have a "quasi-judicial" function, especially when dealing with simple cases or cases involving petty offenses. One such power is discretion to drop trivial cases from further investigation and charges.

Bizai shobun is rationalized based on two principles: (1) when involving minor offenses, the police shall not put victims under undue psychological stress with an unnecessarily punitive punishment of offenders and (2) to enhance efficiency of the police work, police should focus on more serious crimes and use informal means in dealing with the less serious crimes (Arakawa 1987).

To maintain jurisdictional consistency, the police power of *bizai shobun* is granted by the prosecutor. Typically the supervising prosecutor's office will issue an order for specific types of trivial offenses to be dropped from further processing by the police. These trivial offenses commonly include assault, theft, fraud, embezzlement, and gambling. For example, in the early 1990s there were instructions for theft cases to be

dropped under three conditions: (1) if the value of stolen goods did not exceed 10,000 yen (\$85); (2) if the offender had a fixed residency, which excluded transients (particularly foreigners); and (3) if the offender repented (Johnson 2002).

It was estimated that up to 40 % of criminal suspects arrested by the police are released without any charges. Many of these cases were resolved through negotiation, apology, compensation, and forgiveness between the offender and the victim. Furthermore, victims, their family, and their community are critical in the police decision to employ *bizai shobun*, because they must be willing to accept the offender's apology and to forgive him/her. This is to ensure the successful reintegration of the offender back into the community.

Limitations of Reintegrative Shaming

One of the most fundamental criticisms of reintegrative shaming theory is its incomplete conceptualization of sanctions (Hay 2001). Braithwaite depicts sanctions as shaming, either reintegrative or stigmatizing, and makes the assumption that all deviance is detected and reacted to with some type of corrective action. However, deleterious reactions (no reaction) to deviance are common. Even if existing, some reactions to deviance may only intend to evoke fear of further punishment, rather than trying to evoke remorse/shame. Reintegrative shaming theory is thus regarded as incomplete, because shaming is not the only form of corrective reaction.

The concept of reintegrative shaming has the obvious strengths of maintaining continuous and interdependent relationships. However, shaming in communitarian societies (e.g., Confucian culture) may become too powerful and overwhelming, thus, having the risk of destroying robust individuation within secure social bonds. An example of the danger associated with engulfment of the individual is the high suicide rate in Japan and, to a certain degree, in China (Braithwaite 1989)

Reintegrative shaming theory is also limited in scope to personal and property offenses where there is a clear consensus regarding their moral wrongfulness (Hay 2001). In cases where societal

consensus is unclear, such as white-collar crime, drug offenses, and public order offenses (e.g., prostitution, gambling), shaming may not take place (either not expressed by the public or dismissed by the wrongdoer), let alone reintegration.

Besides these conceptual challenges, reintegrative shaming theory faces difficulties of empirical validation. Some of the key concepts such as shame and reintegration are difficult to operationalize and measure with empirical data (Zhang 1995; Zhang and Zhang 2004). In addition, major moderating and mediating factors that may indirectly affect the potency of reintegrative shaming have yet to be established. For example, a study found that procedural justice (i.e., a sense of being treated fairly) affected how well offenders responded to efforts of reintegrative shaming (Tyler et al. 2007).

These issues of measurement are magnified particularly in cross-cultural settings where definitions and perceptions of shame may diverge markedly. A successful shaming program in one type of society (e.g., communitarian) may not work well in another type of society (e.g., individualistic).

Last, but not the least, shaming may run the risk of civil vigilantism. It is especially noteworthy of its potentially massive invasions of personal privacy in the online environment. An example involves the growing popularity of the "Human-Flesh Search" (*renrou sousuo*) in China where individuals' private records were dug out and posted on the website by Internet users for a variety of reasons such as personal revenge and resentment to corrupt public officials and celebrity. This counters the very intent of maintaining a civil and orderly society (Skoric et al. 2010). Within the criminal-justice system, reintegrative shaming programs may run the risk of widening the net of social control and blurring the line between moral and legal issues.

Conclusions

Reintegrative shaming, as practiced in China and Japan, place great faith on public officials – their

personal morality, skills, and judgment – in making important decisions about law and order. By using extralegal means in dealing with essentially a legal matter, these criminal-justice systems runs the risk of corruption and undermining procedural transparency and the predictability and fairness of the decisions. Despite of these potential limitations, reintegrative shaming provides a plausible, alternative means for crime control and prevention. Reintegrative shaming has been demonstrated theoretically, and to a certain degree, empirically, to be more effective and efficient than the traditional means of social control.

Recommended Reading and References

- Arakawa M (1987) Diversion and criminal law in Japan (ディバージョンと刑法に関する一考察 – 警察における微罪処分を中心として). *Law Polit (法と政治)* 38:421–466
- Becker H (1973) *Outsiders*. Free Press, New York
- Braithwaite J (1989) *Crime, shame, and reintegration*. Cambridge University Press, New York
- Braithwaite J (1993) Shame and modernity. *Br J Criminol* 33:1–18
- Braithwaite J (2002) *Restorative justice and responsive regulation*. Oxford University Press, New York
- Braithwaite J, Mugford S (1994) Conditions of successful reintegration ceremonies. *Br J Criminol* 34:139–171
- Chen X (2002) Social control in China: applications of the labeling theory and the reintegrative shaming theory. *Int J Offender Ther Comp Criminol* 46:45–63
- Foote DH (1992) The benevolent paternalism of Japanese criminal justice. *Calif Law Rev* 80:317–335
- Hay C (2001) An exploratory test of Braithwaite's reintegrative shaming theory. *J Res Crime Delinq* 38:132–153
- Johnson DH (2002) *The Japanese way of justice: prosecuting crime in Japan*. Oxford University Press, New York
- Lemert E (1951) *Social pathology*. McGraw-Hill, New York
- Lu H, Zhang L, Miethe TD (2002) Interdependency, communitarianism and reintegrative shaming in China. *Soc Sci J* 39:189–201
- Makkai T, Braithwaite J (1994) Reintegrative shaming and compliance with regulatory standards. *Criminology* 32:361–385
- Miethe T, Lu H, Reese E (2000) Reintegrative shaming and recidivism risks in drug court: explanations from some unexpected findings. *Crime Delinq* 46:522–541
- Skoric MM, Ping J, Chua E, Liew MA, Wong KH, Jue P (2010) Online shaming in the Asian context: community empowerment or civic vigilantism? *Surveill Soc* 8:181–199
- Tyler TR, Sherman L, Strang H, Barnes GC, Woods D (2007) Reintegrative shaming, procedural justice, and recidivism: the engagement of offenders' psychological mechanisms in the Canberra RISE drinking and driving experiment. *Law Soc Rev* 41:553–586
- Zhang SX (1995) Measuring shaming in an ethnic context. *Br J Criminol* 35:248–262
- Zhang L, Zhang S (2004) Reintegrative shaming and predatory delinquency. *J Res Crime Delinq* 41:433–452
- Zhang L, Zhou D, Messner SF, Liska AE, Krohn MD, Liu J, Lu Z (1996) Crime prevention in a communitarian society: Bang-Jiao and Tiao-Jie in the People's Republic of China. *Justice Q* 13:199–222

Shift Work

- ▶ [Officer Safety, Health, and Wellness](#)

Shock Incarceration

- ▶ [History of Boot Camps](#)

Shoplifting

Jason Davis
Criminal Justice, Clayton State University,
Morrow, GA, USA

Overview

Shoplifting refers to the theft of retail merchandise during hours of operation (Clarke 2003; Hayes and Cardone 2006). This form of larceny is viewed as a minor offense that does not generate significant public outrage or fear. Furthermore, it lacks the sensational appeal that necessitates extensive media coverage or immediate legislative responses (Cromwell et al. 2006; Klemke 1992). Notwithstanding these perceptions, retail theft imposes significant financial and social harms. Each year, shoplifting

costs companies billions of dollars in lost revenues or merchandise. Beyond these losses, however, retail theft entails additional financial costs. Companies allocate millions of dollars each year for security personnel and measures to help deter shoplifting (Tonglet 2002; Gills et al. 1999; Cox et al. 1993). In some cases, retailers pay higher insurance premiums because they experience high rates of theft (Burrows 1998; Finklea 2011). Shoplifting can also adversely affect honest patrons. To offset financial losses, companies levy a “consumer tax” in which merchandise prices of commonly pilfered items are substantially increased (Cox et al. 1993; Schwartz and Wood 1991). Customers can also be inconvenienced by cumbersome security measures such as cable locks, key controlled glass cabinets, or other equipment that limits accessibility to merchandise (Tonglet 2002; Cox et al. 1993; Dawson 1993). Finally, state governments and local communities lose tax revenues that could be generated from stolen merchandise.

Beyond the financial tolls, shoplifting can prompt safety and domestic security problems. Unsuspecting customers or employees can be injured by fleeing or aggressive shoplifters attempting to avoid apprehension (Axelrod and Elkind 1976). A 2011 survey found that 13 % of retail crime apprehensions involved some level of physical assault or battery (National Retail Foundation 2011). When perishable goods such as infant formula and over-the-counter drugs are shoplifted, fraudulently repackaged, and sold on the open market, customers can unknowingly consume outdated or spoiled products (Finklea 2011; National Retail Foundation 2011). Finally, shoplifting by professional or organized retail theft (ORT) rings has been linked to money laundering, organized criminal syndicates, and the funding of terrorist activity (National Retail Federation 2011; Finklea 2011).

This entry provides a general overview of shoplifting. Specifically, it discusses the financial losses associated with shoplifting and how these figures likely underestimate the true costs of shoplifting. Next, a typology of shoplifters will

be offered with primary attention given to amateur and professional shoplifters. A discussion of common shoplifting targets and techniques will follow as well as an examination of the theoretical explanations used to help understand why offenders engage in shoplifting. Finally, this entry will review common antishoplifting measures utilized within the retail industry and summarize their effectiveness.

Financial Estimates of Shoplifting

No other property crime including robbery, burglary, or auto theft can match the financial impact of shoplifting. Estimates have consistently shown that retailers lose anywhere between \$10 and \$50 billion dollars annually to shoplifting (National Association for Shoplifting Prevention 2012; National Retail Foundation 2011; Centre for Retail Research 2011; Clarke 2003). A recent study found that shoplifting cost US retailers approximately \$12.1 billion (National Retail Federation 2011). In 2010, a global retail theft survey discovered that shoplifting incidents cost 1,187 retailers representing 43 countries \$51.5 billion (Centre for Retail Research 2011). While the financial costs attributed to shoplifting are staggering, it is generally understood that these figures grossly underestimate the true extent of the problem (Burrows 1999; Dabney et al. 2004; Clarke 2003).

Dilemmas Associated with Shoplifting Estimates

Underreporting is a major problem associated with shoplifting data as only one-third of all thefts get reported to the police (Dabney et al. 2004; Clarke 2003). When shoplifting is officially reported, few incidents result in a formal record or prosecution (Klemke 1992; Dabney et al. 2004; Clarke 2003). Instead, law enforcement is more inclined to issue a warning or have the charges downgraded to a misdemeanor or violation

offense (Dabney et al. 2004; Clarke 2003). Many retail companies are reluctant to formally report shoplifting and thus contribute to underreporting dilemmas (Burrows 1999). Part of the hesitancy may stem from legal obstacles. Stores may be unaware of shoplifting statutes, or they may be dissatisfied with the weak laws or outcomes that insufficiently punish offenders (Axelrod and Elkind 1976; Burrows 1999). In lieu of formal responses, companies more commonly seek civil restitution from the offenders. Companies may also be reticent about disclosing adverse shoplifting information (Burrows 1999). Publicity about shoplifting could alert potential offenders about a store's security vulnerabilities, alarm shareholders about potential crime problems, or signal that costly security measures are ineffective (Burrows 1999; Axelrod and Elkind 1976). Even if retailers suspect customers of shoplifting, they are apprehensive about accosting them because of a fear of litigation resulting from false arrest or imprisonment (Axelrod and Elkind 1976). Some stores are equally worried about the risk of embarrassment and loss of reputation that false arrests could yield. Collectively, there appears to be reluctance to report, classify, or respond to shoplifting as a serious felony offense.

Beyond underreporting dilemmas, sources of shoplifting information including police records, victimization or self-report surveys, and industry level data are problematic because they fail to directly measure shoplifting (Burrows 1998). For instance, shoplifting is coalesced into the larger category of larceny-theft in the Uniform Crime Report, and therefore, it is impossible to discern what percentage of thefts can be attributed to shoplifting. The National Crime Victimization Survey (NCVS) neglects shoplifting altogether since it focuses on individual victimization, not crimes committed against organizations (Burrows 1998). Store-generated data suffers similar measurement limitations. Since only a small percentage of offenders are detected at the time of the offense, store security managers or personnel must "guesstimate" what percentage of loss can be directly attributed to shoplifting

(Dabney et al. 2004). Through the use of routine audits, companies measure the disparity between merchandise physically present in a particular store and financial statements regarding sold merchandise (DiLonardo 1997). Such audits determine levels of *inventory shrinkage* or shortages in inventory levels. The greater the amount of unaccounted merchandise, the more revenue lost. With the exception of employee theft, shoplifting is generally considered the second most common source of inventory loss (Hayes and Cardone 2006). In most cases, estimates maintain that roughly 33–43 % of all retail losses are caused by shoplifting (National Retail Foundation 2011; Centre for Retail Research 2011). Nevertheless, without direct observation, it is not feasible to parse out what percentage of inventory shrink is actually shoplifting related (Dabney et al. 2004). Indeed, a study relying on surveillance cameras observed more shoplifting incidents than were formally reported to law enforcement (Dabney et al. 2004). The authors concluded that shoplifting incidents from a single retail drug store approximated the total number of all larceny-theft cases reported to police in the metropolitan Atlanta area over the course of 1 year. Such findings call into question the reliability and validity of both formal and store-generated estimates regarding the prevalence and financial impact of shoplifting (Dabney et al. 2004).

Shoplifting Typology

Research has consistently identified two categories of shoplifting offenders: (a) amateur or opportunistic shoplifters and (b) professional shoplifters that generally operate as part of an organized retail crime (ORC) ring (Clarke 2003). To a lesser extent, studies have also acknowledged the existence of compulsive shoplifters known as kleptomaniacs (Blanco et al. 2008).

Amateur Shoplifters

The majority of shoplifters are considered amateurs who steal relatively inexpensive merchandise for the purpose of personal use or consumption (Cameron 1964; Cox et al. 1990;

Clarke 2003). In many cases, amateur shoplifters lack a clear motive for stealing as many offenders possess the resources to legitimately purchase merchandise or they may not have an immediate need or desire for the products they steal. Accordingly, amateur shoplifting is typically viewed as an opportunistic crime in which offenders steal goods when they perceive relatively easy opportunities to do so. Some amateurs envision shoplifting more as a challenge and thus steal to see if they can get away with it. Amateur shoplifting also tends to be impulsive and infrequent. Their acts of theft are generally unplanned, and many steal only a few times a year (Cameron 1964). Overall, amateurs are not committed to a criminal lifestyle (Cameron 1964; Clarke 2003). Formal detection or apprehension is generally enough to deter future offending (Cameron 1964).

Amateur shoplifting is not confined to a small subgroup of offenders (Cox et al. 1990). Instead, a sizeable number of people have shoplifted at some point in their life (Cameron 1964; Cromwell and Thurman 2003). According to the National Association for Shoplifting Prevention (2012), an estimated 1 out of 11 people or 27 million people have shoplifted. Not surprisingly, there is not a prototypical profile of a shoplifter. Amateur shoplifters are a demographically diverse group. Various studies have concluded that men and women are equally likely to steal as are young and old and people of various ethnic backgrounds and social classes (Klemke 1992; Dabney et al. 2004; Blanco et al. 2008; Asquith and Bristow 2000; Dawson 1993; Tonglet 2002; Cromwell and Thurman 2003; Cromwell et al. 2006). While the influence of demographic characteristics has been inconclusive, behavioral clues may be stronger predictors of shoplifting. Shoppers who leave a store without purchasing an item, scan the store for security measures, or tamper with products are more likely to steal (Dabney et al. 2004).

Organized Retail Crime (ORC): Professional Shoplifting

Unlike amateurs, professional shoplifters are more inclined to steal merchandise for the

purposes of reselling stolen goods for a profit (Cameron 1964; National Retail Federation 2011; Finklea 2011). This category of shoplifters tends to be more organized and skilled, steal merchandise more frequently, and target more expensive items. It is estimated that roughly one-quarter of shoplifting cases are committed by professional shoplifters (Finklea 2011). Professional thefts are generally part of a larger criminal network known as organized retail crime (ORC). At its core, ORC networks involve a relationship between *boosters* and *fences*. Boosters steal merchandise and sell these goods to fences for approximately 10–30 % of the retail value of a stolen product (Cameron 1964; Finklea 2011). For their part, fences purchase, conceal, or transport stolen merchandise. Fences use a variety of business such as small “mom and pop” community establishments, pawn shops, flea markets or swap meets, convenience stores, and online auctions to dispose stolen goods (Cromwell and Olson 2006; Finklea 2011). In some cases, the fences may represent organized crime or terrorist operations. Fences typically supply boosters with “fence sheets” that specify desired merchandise, the value or profit margin for stolen goods, and stores that commonly carry preferred items (Cromwell and Olson 2006). Some mid- or higher-level fences ostensibly operate legitimate warehouses that in reality remove security tags, store labels, repackage goods, or alter expiration dates from stolen merchandise (Finklea 2011; Cromwell and Olson 2006). The “cleaned” stolen goods are amalgamated with legitimately purchased merchandise and reintroduced on the open market. In addition to professional fences, there is the existence of part-time fences who occasionally buy stolen merchandise primarily for personal consumption. In other instances, barter exchange professional services or drugs for stolen merchandise. These nonprofessional fences do not purchase a high volume of stolen merchandise nor do they rely on fencing as their primary source of income (Cromwell and Olson 2006). The advent of online internet auction sites such as eBay has created another outlet for professional shoplifters to dispose stolen goods.

Known as *e-fencing*, this form of disposing stolen merchandise is potentially more lucrative as professional shoplifters can earn approximately 70 cents on the dollar for stolen goods (Finklea 2011).

Kleptomaniacs

While amateurs and professionals are the ubiquitous shoplifters, there is a smaller category of offenders known as kleptomaniacs who are unable to resist the temptation to steal and generally steal items they do not have a conceivable use for (Blanco et al. 2008). Historically, the term kleptomaniac was used to describe females who had an uncontrollable urge to pilfer (Hayes and Cardone 2006). More recently, kleptomania is considered a psychological disorder in which a person experiences intense feeling of anxiety prior to stealing and a sense of gratification or release after stealing (Blanco et al. 2008). It is generally believed that kleptomaniacs also suffer from other psychological disorders such as impulsivity, mood disorders, or obsessive-compulsive behaviors that compel them to steal.

Common Shoplifting Targets and Techniques

Common Shoplifting Targets

A general sample of items generally shoplifted include but are not limited to music, games, electronics, apparel, batteries, tobacco products, infant formula, and over-the-counter (OTC) medicines (Clarke 1999; Clarke 2003; Hayes and Cardone 2006; Finklea 2011). Clarke's (1999) CRAVED model focuses on six attributes that make certain targets vulnerable. In particular, items that are *concealable, removable, available, valuable, enjoyable, and disposable* are deemed *hot products*. On the whole, shoplifters commonly target merchandise that can be easily concealed and avoid stealing merchandise in large quantities or items too large to conceal for fear of attracting too much attention (Carroll and Weaver 1986). Both amateur and professional shoplifters will target

widely available merchandise especially new attractive products such as cell phones, computers, video games, and apparel. While professional shoplifters typically steal merchandise that will garner greater profit margins, novices including adolescent offenders may be driven to steal merchandise that holds personal worth or cultural value. Finally, the CRAVED model predicts that shoplifters, particularly professional shoplifters, are more inclined to steal merchandise that is easy to sell or convert into cash.

Common Shoplifting Techniques

In its simplest form, shoplifting involves the illegal concealment of merchandise in one's clothing or hand. While some brazen shoplifters may simply grab goods and leave the store quickly, both amateur and professional shoplifters utilize a variety of methods to illegally pilfer retail merchandise. Research has noted that many shoplifters work in teams with some members serving as lookouts to alert the thieves about potential dangers (Hayes and Cardone 2006). Listed below are select shoplifting techniques commonly identified by researchers and practitioners (Hayes and Cardone 2006; Clarke 2003; Finklea 2011; DiLonardo 1997):

- **Booster boxes or bags:** Shoplifters use a special foil-lined dummy box or bag that has false bottoms or openings to hide stolen merchandise. The foil lining is designed to weaken electronic article surveillance (EAS) tag signals so that they do not activate or trigger the alarm upon departure from a store. In other situations, shoplifters will conceal stolen goods within a legitimately purchased boxes or bags including televisions boxes, computer boxes, sporting goods merchandise, suitcases, dog food bags, or safes.
- **Diversions:** Working in groups, one or more shoplifters will create some type of diversion (pretending to slip and fall, faking an illness or injury, dropping or breaking merchandise, knocking down a display cabinet or shelves, or triggering a fire alarm). While store employees are distracted, other shoplifters will steal merchandise.

- **Fitting room:** Shoplifters will take several items or garments into a fitting room where they will wear the stolen clothing or conceal items in a bag or purse.
- **Clothing or crotchwalking** This shoplifting method occurs when an offender wears large baggy clothing to hide merchandise. In some cases, shoplifters wear special clothing (i.e., coats, shirts, pants) that contains multiple hidden pockets to help conceal merchandise. Some female offenders will wear large maternity clothing to give the appearance of being pregnant while concealing merchandise.
- **Receipt manipulation:** In some instances, shoplifters will make a legitimate purchase and take the item to their vehicle. Thereafter, they will return to the store and steal the same item listed on the receipt. If stopped by employees or store detectives, the shoplifter can present the receipt as proof of a legitimate purchase. Days later, the shoplifter may return the legitimately purchased item for a full refund. In other cases, shoplifters may use discarded receipts found in the parking lot, trash bins, or store floor to steal the items listed on the receipt. The disposed receipts can also be used for refund purposes. The shoplifter may pick up the items listed on the dispensed receipt and seek a cash refund.
- **Smash and grab:** This form of theft occurs when a group of shoplifters typically steal a car or sports utility vehicle (SUV) and drive it through a storefront window. Once the vehicle has smashed through the store, shoplifters quickly grab and pack the car with as much merchandise as possible before driving away.
- **Price tag switching:** This refers to the fraudulent removal of price tags or bar codes from expensive merchandise and replacing them with cheaper price codes.

Theories of Shoplifting

Empirical studies have concluded that there is not a solitary factor that causes a person to engage in shoplifting (Cromwell et al. 2006; Tonglet 2002).

Instead, shoplifting behaviors are shaped by a range of factors. Many view shopping as a crime of opportunity, while other perspectives assert that offenders view shoplifting as a trivial offense that does not cause harm. Some frameworks maintain that shoplifting is related to age-specific factors. Finally, research has argued that psychological needs or disorders may be a determinant of shoplifting behaviors.

Opportunity Perspectives

Some of the prevailing theories of shoplifting including routine activity theory (RAT) and opportunity structure for crime can be classified as theories of opportunity. Partially rooted in rational choice principles, these opportunity theories assert that would-be offenders take advantage of perceived easy opportunities to offend after careful consideration of the potential risks and rewards associated with a particular act (Clarke 1997). Opportunity theories also focus on the importance of the crime setting (Clarke 2003). In the context of shoplifting, opportunity theories are interested in understanding how the retail environment contributes to the prevalence of shoplifting. Shoplifters will seek stores with weak or ineffective security measures because they believe the benefits of shoplifting (procuring something for free) outweigh the risk of apprehension (Gills et al. 1999). Indeed, many shoplifters perceive retail pilfering as a low-risk offense that requires minimal effort, abundance of opportunities, and substantial gains (Tonglet 2002).

Routine activity theory maintains that crime occurs when a *motivated offender* comes into contact with *suitable targets* in the absence of *capable guardianship* (Cohen and Felson 1979). The convergence of these three factors in time and space represents a crime triangle. Routine activity theory is ideally suited for understanding shoplifting. Retail environments bring potential motivated offenders into contact with suitable targets including small, portable, and desirable goods that can be easily concealed. To the extent that retail environments lack adequate surveillance or security measures, motivated

offenders will take advantage of the lack of guardianship and shoplift valuable targets.

Touted as an integrated theory that incorporates elements of environmental criminology, rational choice, lifestyle, and routine activity theory, Clarke's (1997) opportunity structure theory asserts that the physical environment influences opportunities for crime. Immediate situational or environmental factors dictate the supply of *targets* (objects), *victims* (vulnerable individuals), and *crime facilitators* (weapons) necessary for offending. To this end, retail environments play an important role in creating or reducing shoplifting opportunities. Through personal experiences, peer associations, or the media, would-be offenders learn the risks, effort, and rewards associated with specific stores, and they will target those with the greatest vulnerabilities or relatively easy opportunities to steal. In this context, retailers must manipulate the environment and implement target-hardening devices to decrease opportunities for offending. Research indicates that shoplifters do in fact evaluate store characteristics when deciding whether or not to steal (Carroll and Weaver 1986). In particular, experienced shoplifters assess store layouts for security devices, counter heights, or behavior of store staff in order to identify potential risks. The presence of these factors is enough to deter most expert shoplifters.

Techniques of Neutralization

Another explanation for why people shoplift is techniques of neutralization theory (Sykes and Matza 1957). This perspective asserts that most offenders hold conventional norms or attitudes, and they must rely on rationalizations to help neutralize the guilt and shame associated with offending. Potential offenders use a priori excuses that allow them to drift into criminality while temporarily suspending their moral convictions. In its original conception, Sykes and Matza (1957) identified five different techniques of neutralization that offenders use to justify criminal behavior including *denial of responsibility* (offending is beyond person's control), *denial of injury* (offending behavior is trivial and does not really harm anyone), *denial*

of victim (victim deserved to have a crime committed against them), *condemnation of the condemners* (if authority figures engage in unethical behaviors with impunity, then others should be allowed to engage in similar acts), and *appeal to higher loyalties* (commitment to peers or peer pressure caused offending behavior). Recent research has identified at least seven other neutralizations including *defense of necessity* (offending is a matter of survival for basic or essential needs), *metaphor of the ledger* (crime offsets unfair conditions or can settle a grievance), *everyone else is doing it* (certain crimes are common), *denial of the necessity of the law* (laws are unfair), and *sense of entitlement* (offender should enjoy certain privileges) (Cromwell and Thurman 2003).

Studies have found that shoplifters typically use one or more justifications to help neutralize their illegal pilfering (Tonglet 2002; Cromwell and Thurman 2003; Clarke 2003). Many offenders hold proshoplifting attitudes and perceive pilfering as a relatively trivial offense that does not significantly harm the overall sustainability of profitable retailers (*denial of injury*). Some shoplifters imply that they had little control over their actions (Cromwell and Thurman 2003). This *denial of responsibility* neutralization would be consistent with shoplifters who suffer some psychological compulsion or addictive behavior to engage in theft. In other instances, offender attitudes are congruent with a *denial of victim* perspective. Many people view large organizations as wealthy and somewhat ruthless entities (Schwartz and Wood 1991; Clarke 2003; Dawson 1993). Accordingly, shoplifters are less likely to experience feeling of guilt, remorse, or shame about stealing from impersonal stores. Such antagonistic attitudes about retailers may also reflect a *metaphor of the ledger* justification in which they view shoplifting as a sense of entitlement or compensation for unfair treatment or price markups by indifferent companies (Schwartz and Wood 1991). Other shoplifters may utilize a *defense of necessity* justification in which offenders assert they had to steal in order to survive or because they did not possess

the resources to make a legitimate purchase (Cox et al. 1990; Cox et al. 1993; Cromwell and Thurman 2003). A recent study identified two other neutralizations including *justification by comparison* and *postponement* (Cromwell and Thurman 2003). Offenders using the former rationalization maintain that if they were not shoplifting, they would be committing more serious offenses, while those exploiting the latter neutralization momentarily delay any thoughts about the offending behaviors. In essence, people that use postponement do not think about their crime.

Age-Specific Reasons

Motivations to shoplift can also be age specific. Studies of juvenile shoplifters have discovered that adolescents steal for thrill or novelty reasons (Cox et al. 1990; Cox et al. 1993). In addition, social desirability is a motivation for adolescent shoplifters. Juveniles steal popular merchandise including electronics or clothing in order to enhance their status among peers. In some cases, juveniles steal items that they are too young to legally purchase or too embarrassed to buy such as cigarettes, alcohol, or condoms (Cox et al. 1990). Interestingly, there is little evidence to suggest that juvenile shoplifters are directly affected by peer pressure, dares, or status attainment (Cox et al. 1993). There is also evidence that juvenile shoplifters may not be deterred by antishoplifting security measures or formal apprehension (Klemke 1992). Apprehension by parents, store personnel, or the police can actually increase the likelihood of future shoplifting for adolescent shoplifters presumably because deviant labels propel the offender, particularly female shoplifters, to commit future acts of pilfering (Klemke 1992).

Psychological Factors

Some argue that shoplifting represents coping mechanism to address stressful life events or psychological problems including depression, anxiety, bulimia, anorexia, substance addiction, financial hardship, or relationship problems (Lamontagne et al. 2000; Blanco et al. 2008; Schwartz and Wood 1991). For instance,

bulimics may steal food to address eating disorders, while people who suffer from anorexia may steal clothing or beauty products to enhance their physical appearance (Schwartz and Wood 1991). Stress can also contribute to shoplifting behaviors. There is evidence that people who are experiencing stressful life events like marital problems, employment or financial problems, or health issues shoplift as a coping mechanism (Schwartz and Wood 1991). These psychological perspectives assert that shoplifting generates excitement and produces an adrenaline rush or high that temporarily subdue feelings of anger, frustration, or depression (Blanco et al. 2008). Shoplifters may also suffer some sort of psychiatric disorder generally related to pathological gambling, bipolar disease, nicotine dependency, or alcohol abuse. In support of psychological perspectives, research has noted that shoplifters are indeed more likely to seek mental health treatment compared to nonshoplifters (Blanco et al. 2008).

Common Antishoplifting Measures and Their Effectiveness

In accordance with the basic tenets of opportunity theories, many antishoplifting measures focus on manipulation of retail environment or target-hardening measures. Clarke's (1997) situational crime prevention model proposes 16 different techniques designed to make crime opportunities less attractive. The various techniques are designed to (a) increase the effort associated with offending, (b) increase the perceived risks of crime, (c) reduce the anticipated rewards of criminality, and (d) remove the excuses to offend. Some of the crime reducing-measures include target-hardening devices, formal and natural surveillance, access control, entry and exit screening, ink security tags, property identification, and warning signs. Several of these techniques are well established within the retail industry. Stores commonly use target-hardening devices like locks, safes, cabinets, mechanical cables, security tags, or benefit denial tags. Most retailers employ various

forms of formal surveillance such as store security officers, employees, closed-circuit television (CCTVs), or dressing room attendants. Stores display dummy products or empty boxes as a means of removing the benefits of shoplifting. Finally, companies use warning signs about the prosecution of apprehended offenders or signs that indicate a store is under formal surveillance. Collectively, these security measures emphasize the importance of physical store layout, staff and security personnel, and target-hardening devices that help monitor, detect, and apprehend potential shoplifters.

Physical Store Layout

In many ways, the physical layout of a store can facilitate the commission of shoplifting (Carroll and Weaver 1986; Burrows 1998). Most store layouts openly display merchandise and encourage self-service. In addition, the size of most stores makes it difficult to effectively monitor all merchandise and customers simultaneously. Stores typically have multiple exit points, display merchandise near exits, or have high shelves that obstruct viewing angles. Therefore, retail stores provide easy access to valuable and concealable goods in the absence of effective surveillance (Clarke 2003; Tonglet 2002; Cox et al. 1993).

Accordingly, retailers employ a variety of techniques intended to manipulate the store environment, increase the level of surveillance, and reduce the opportunities to steal. For instance, stores create wide aisles to help alleviate congestion and create clear sight lines for employees to easily monitor customer behavior. Stores use convex mirrors and adequate lighting that enables staff to monitor the store from multiple angles. Merchants commonly monitor entry or exit points with employees, detectives, CCTVs, or audible alarm antennas. Smaller items are placed near checkout counters for increased visibility. Finally, stores like to maintain neatly ordered or arranged shelves so that employees can easily notice irregularities (Clarke 2003). While the presence of various security measures is necessary, retailers face the challenge of balancing a customer-friendly

environment with an overly intrusive security atmosphere. The inclusion of too many cumbersome, intrusive, or highly visible crime prevention tools can make legitimate customers feel uncomfortable or alienated (Dawson 1993).

Staff and Security Personnel

Research indicates that employees and security personnel are the most important factor in reducing shoplifting (Gills et al. 1999; Clarke 2003; Hollinger and Dabney 1994; Dawson 1993). When sales associates immediately approach and greet customers, make eye contact with them, or consistently offer assistance, it sends a signal to would-be shoplifters that they have been noticed and are being monitored. Indeed, shoplifters report that they are less likely to steal if they believe they are being closely scrutinized or watched by alert employees or store detectives (Gills et al. 1999). Not surprisingly, shoplifters prefer environments in which workers are disinterested in providing quality customer service. They also prefer stores that are routinely understaffed so that employees are too busy to monitor customer behaviors. Furthermore, stores that have high turnover rates or rely heavily on part-time workers undermine effective shoplifting surveillance efforts (Hollinger and Dabney 1994). In these situations, workers are less committed to the company, lack security training, and have limited experience or awareness of shoplifting clues. To be sure, company commitment to equitable treatment of its employees and low turnovers significantly lower levels of inventory loss (Hollinger and Dabney 1994).

However, misuse or abusive employee surveillance can lead to lawsuits. Since untrained employees may misinterpret customer behaviors, they may inadvertently confront and detain innocent customers (Axelrod and Elkind 1976). In addition, employee racial biases or prejudices can affect surveillance efforts. For example, several retail companies including Eddie Bauer, Lord & Taylor, and The Children's Place have settled civil suits amid accusations that store security personnel unfairly monitored, accosted, or apprehended ethnic customers.

These high-profile cases have drawn attention to a phenomena of *consumer racial profiling* (CRP) in which employees and store detectives target shoppers, primarily African Americans, on the basis of race and ethnicity, not necessarily suspicious customer behavior (Gabbidon and Higgins 2007; Asquith and Bristow 2000). Several African Americans report that they have been victims of retail racism in terms of poor service, excessive monitoring, or being looked upon with suspicion by employees or security personnel (Gabbidon and Higgins 2007). Such practices overlook white customers who are equally as prone to shoplift as other customers.

Target-Hardening Devices: Electronic Article Surveillance (EAS) Tags

Introduced in the late 1960s, EAS tags are antishoplifting protection systems that have been widely adopted by the apparel, music, and general retail industries. In general, EAS tags are detachable devices affixed or pinned to merchandise that must be removed or deactivated by a sales associate at the time of purchase. Transmitter stands strategically placed near exit doors are designed to trigger an audible or silent alarm if a customer passes through the transmitter with an active EAS tag. Not only do EAS tags provide physical obstacles to shoplifting, they provide a psychological deterrent as potential offenders are reluctant to risk activating an alarm that alerts store security or employees (DiLonardo 1997; Dawson 1993). A multiyear study of several retail apparel stores found that stores that used EAS tags experienced significant declines in inventory shrinkage, while shortages among retailers without EAS tags actually increased by 30 % during the same time period (DiLonardo 1997). Overall, shoplifting and inventory shrinkage rates have been successfully reduced by 35–75 % after the introduction of EAS systems (DiLonardo 1997).

Other studies, however, have found that EAS tags do not deter shoplifters. Many shoplifters note that EAS tags are easy to remove with pliers or other tools (Gills et al. 1999). In some instances, shoplifters are able to use specially designed foiled bags that stunt the

ability of transmitters to sound an audible alarm. More brazen shoplifters simply walk out of the store even if the alarm is sounded. The errant functioning of EAS tags can also contribute to their ineffectiveness. Stores that experience a high number of false alarms triggered by system malfunction or sales associates that fail to remove or deactivate sensor tags risk the potential of offending customers and losing their patronage (Dawson 1993). Legitimate customers are likely to feel embarrassed, annoyed, or upset by inadvertent EAS alarms. Perhaps more importantly, customers may be subject to unwarranted suspicion or scrutiny from security personnel and other customers. Beyond the potential of tarnishing a retailer's image and reputation, false alarms can levy financial costs in the form litigation and punitive damages (Dawson 1993). Another consequence of errant EAS tags is related to the fact that customers, staff, and security may, over time, become desensitized to the alarms.

Despite the best efforts of retailers, evidence suggests few antishoplifting measures including cameras, security tags, mirrors, or prosecution warning signs effectively deter the motivation of pilfers (Gills et al. 1999; DiLonardo 1997). Closed-circuit televisions (CCTVs) are considered ineffective by many shoplifters because they cannot capture quality indisputable images (Gills et al. 1999). Furthermore, shoplifters maintain that most stores have natural blind spots that make it impossible for cameras to monitor. Warning signs that imply all shoplifters will be prosecuted if detected similarly have little impact on the motivation to steal. Shoplifters typically believe the most stores will not go through the trouble of going to court or legal proceeding to recoup relatively inexpensive merchandise (Gills et al. 1999). While some shoplifters acknowledge that ink tags are more challenging to defeat, offenders will still target dye-tagged merchandise primarily because they lack an audible alarm. Once shoplifters effectively remove the merchandise from stores, they have plenty of time to carefully remove the tag (Gills et al. 1999). Even if shoplifters are deterred, the effect is temporary as most will

attempt to discover new means to defeat security impediments (Gills et al. 1999). Since the chances of being detected and apprehended are relatively low and the possibility of stiff criminal sanctioning is equally low, shoplifters associate minimal risks with this type of crime and therefore are not easily deterred (Gills et al. 1999; Clarke 2003; Burrows 1988).

Related Entries

- ▶ [CCTV and Crime Prevention](#)
- ▶ [Crimes of Globalization](#)
- ▶ [Effectiveness of Situational Crime Prevention](#)
- ▶ [Fencing/Receiving Stolen Goods](#)
- ▶ [Situational Crime Prevention](#)

Recommended Reading and References

- Asquith JAL, Bristow DN (2000) To catch a thief: a pedagogical study of retail shoplifting. *J Educ Bus* 75:271–276, May/June
- Axelrod AD, Elkind T (1976) Merchants' response to shoplifting: an empirical study. *Stanf Law Rev* 28:589–612
- Blanco C, Grant J, Petry NM, Simpson HB, Alegria A, Liu S-M, Hasin D (2008) Prevalence and correlates of shoplifting in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *Am J Psychiatry* 165:905–913
- Cameron MO (1964) *The Booster and the snitch: department store shoplifting*. Free Press, New York
- Carroll J, Weaver F (1986) 'Shoplifters' perceptions of crime opportunities: a process-tracing study. In: Cornish DB, Clarke RVG (eds) *The reasoning criminal: rational choice perspectives on offending*. Springer, New York
- Centre for Retail Research (2011) *Global retail theft barometer 2011*. http://www.retailresearch.org/grtb_globaltrends.php. Retrieved 29 Nov 2011
- Clarke RV (1997) *Situational crime prevention*, 2nd edn. Harrow and Heston, Albany
- Clarke RV (1999) Hot products: understanding, anticipating and reducing demand for stolen goods. *Police Research Series Paper 112*. Home Office, London
- Clarke RV (2003) "Shoplifting." Problem-oriented guides for police problem-specific guides series, no. 11. U.S Department of Justice, Office of Community Oriented Policing Services, Washington, DC
- Cohen LE, Felson M (1979) Social change and crime rate trends: a routine activity approach. *Am Sociol Rev* 44:588–608
- Cox D, Cox AD, Moschis GP (1990) When consumer behavior goes bad: an investigation of adolescent shoplifting. *J Consum Res* 17:149–159
- Cox D, Cox AD, Anderson RD, Moschis GP (1993) Social influences on adolescent shoplifting – theory, evidence, and implications for the retail industry. *J Retail* 69:234–246
- Cromwell P, Olson JN (2006) Fencing: avenues for redistribution of stolen property. In: Cromwell P (ed) *In their own words: criminals on crime*, 4th edn. Roxbury, Los Angeles, pp 199–209
- Cromwell P, Thurman Q (2003) The devil made me do it: use of neutralization by shoplifters. *Deviant Behav* 24:535–550
- Cromwell P, Parker L, Mobley S (2006) The five finger discount. In: Cromwell P (ed) *In their own words: criminals on crime*, 4th edn. Roxbury, Los Angeles, pp 113–125
- Dabney DA, Hollinger RC, Dugan L (2004) Who actually steals? A study of covertly observed shoplifters. *Justice Q* 21:693–728
- Dawson S (1993) Consumer responses to electronic article surveillance alarms. *J Retail* 69:353–362
- DiLonardo RL (1997) The economic benefit of electronic article surveillance. In: Clarke RV (ed) *Situational crime prevention*, 2nd edn. Harrow and Heston, Albany
- Finklea KM (2011) *Organized retail crime*. Congressional Research Service. CRS Report for Congress, pp 1–25. www.crs.gov
- Gabbidon SL, Higgins GE (2007) Consumer racial profiling and perceived victimization: a phone survey of Philadelphia area residents. *Am J Crim Justice* 32:1–11
- Gills M, Bilby C, Turbin V (1999) Retail security: understanding what deters shop thieves. *J Secur Adm* 22:29–39
- Hayes R, Cardone C (2006) Shop theft. In: Gill M (ed) *The handbook of security*. Palgrave Macmillan, New York
- Hollinger RC, Dabney D (1994) Reducing shrinkage in the retail store: it's not just a job for the loss prevention department. *Secur J* 5:2–10
- Klemke LW (1992) *The sociology of shoplifting: boosters and snitches today*. Praeger, Westport
- Lamontagne Y, Boyer R, Héту C, Lacerte-Lamontagne C (2000) Anxiety, significant losses, depression, and irrational beliefs in first-offence shoplifters. *Can J Psychiatry* 45:63–66
- National Association for Shoplifting Prevention (2012) *Shoplifting statistics*. <http://www.shoplifting-prevention.org/main.asp>. Retrieved 2 Apr 2012
- National Retail Foundation (2011) *2011 organized retail crime survey*. Washington, DC. http://nrf.com/modules.php?name=News&op=viewlive&sp_id=1132&parent_id=950&peer_rev=0&nrf_or=0. Retrieved 28 March 2012
- Schwartz S, Wood HV (1991) Clinical assessment and intervention with shoplifters. *Soc Work* 36:234–238
- Sykes G, Matza D (1957) Techniques of neutralization: a theory of delinquency. *Am Sociol Rev* 22:664–670
- Tonglet M (2002) Consumer misbehaviour: an exploratory study of shoplifting. *J Consum Behav* 1:336–354

Signal Crimes Perspective

► [Reassurance Policing and Signal Crimes](#)

Simulation

► [Agent-Based Assessments of Criminological Theory](#)

Simulation as a Tool for Police Planning

Vasco Furtado
Graduate Program in Applied Informatics,
University of Fortaleza (Unifor), Fortaleza,
Ceará, Brazil

Synonyms

[Crime simulation](#); [MABS](#); [MACS](#)

Overview

Crime relies, directly or indirectly, upon an array of factors, ranging from the levels of concentration of wealth to the physical organization of the urban center under consideration. Modeling the highly interconnected nature of this social system has recently attracted attention in computer science. As experiments in this domain cannot be performed without high risks, because they result on loss of human lives, simulation models have been chosen as supporting tools for this process. Multiagent systems (MAS) primarily study the behavior of autonomous and organized groups of software agents with the purpose of providing solutions to complex problems that could not be achieved by each individual agent alone. Multiagent-based simulation systems have been successfully adopted because the inherent characteristics of the agents (e.g., autonomy,

sociability, and pro-activity) facilitate the construction of more dynamic models, thus contrasting with conventional computer simulation approaches.

Other different approaches, dubbed in general as bio-inspired approaches, have recently been investigated in conjunction with MAS for crime modeling. Collective intelligence as that demonstrated by swarms and the evolutionary approach are two of the most prominent concepts and have been used in the design of crime models. Here, an overview of these concepts from a decision support perspective is done in order to describe how they have been applied for the development of multiagent-based crime simulation (MACS). Two particularly relevant issues in the MACS context are discussed: model calibration and model evaluation.

Fundamentals

According to Russell and Norvig (1995), an *agent* is a physical or abstract entity that can be viewed as perceiving its environment through sensors and acting on the environment through actuators. Ferber (Ferber 1999) considers a multiagent system (MAS) as comprising (i) an environment; (ii) a set of passive and active objects (agents); (iii) an assembly of relations, which link objects to each other; and (iv) operations making it possible for the agents to perceive, produce, consume, transform, and manipulate objects. One can distinguish three levels of organization of agents (Rocher and Sherif 1972): micro-social, groups, and global societies. The micro-social level concerns the interactions between agents and the various forms they relate to one another. The level of groups and of societies refers to the dynamics of intermediate structures like organization and cities, respectively. Multiagent systems have been successfully adopted in conjunction with simulation models, which are generally referred to as multiagent-based simulation (MABS) systems. According to Gilbert and Conte (1995), MABS are especially appropriate when one has to deal with interdisciplinary problem domains. Such an

approach, which is bottom-up in nature, is also appropriate for the study of social and urban problems, since social or urban environments are dynamic, nonlinear, and composed of a great number of variables and entities. The main objectives behind the construction of MABS systems are the following:

- To test hypotheses related to the emergence of macro-level behavior from interactions occurring at micro-social levels
- To build theories that can contribute to a better understanding of sociological, psychological, and ethnological phenomena
- To integrate partial theories coming from different disciplines (e.g., sociology, cognitive psychology, and ethology) in a common theoretical framework

A particular kind of simulation, called geosimulation, addresses an urban phenomena simulation model with a multiagent micro-social approach to simulate discrete, dynamic, and event-oriented systems (Benenson and Torrens 2004). In geosimulated models, simulated urban phenomena are considered a result of the collective dynamic interaction among animate and inanimate entities that compose the map representation. The Geographic Information System (GIS) is responsible for providing the “data ware” in geosimulations.

The study of agent self-organization and related concepts, such as emergence, is another important concept within micro-social MAS. The basic idea is that societies of agents demonstrate intelligent behavior at the collective level out of simple rules at the individual level. Moreover, these individual rules often do not explain the behavior that is attained at the collective level. Swarm intelligence is characterized (i) by strictly local communication, (ii) by the formation of emergent spatial-temporal structures, and (iii) by the agent’s making stochastic decisions based on the local information available. One of the branches of swarm intelligence is ant-colony optimization (ACO), proposed by Bonabeau et al. (1999). ACO is a meta-heuristic model for solving combinatorial problems that can typically be represented as graphs. ACO gets inspiration from other areas of science, in this case, biological

sciences having the special feature of adapting well to dynamic settings.

In a nutshell, ACO works by allowing agents (ants) to explore a search space, but it requires these ants to leave feedback information about locations with good solutions on the space itself. Agents are then attracted by the feedback left in the environment – the larger the amount of information (pheromone), the more attractive the agents find the position in the environment. In order to avoid early convergence to local optima, the approach assumes that the information left is volatile and impermanent; if no other activity occurs, a piece of information left in the environment “expires” or disappears over a certain period of time. Although never explored as a model for criminal behavior, ACO’s characteristics appeared from the start to be an ideal fit to this purpose, as we will discuss later. Briefly, it allows for investigating variations of a learning process with or without a social factor.

Genetic algorithms (GAs) (Holland 1975) are general-purpose search and optimization algorithms that comply with the Darwinian natural selection principle and with some principles of population genetics to efficiently design (quasi-) optimal solutions to complicated computational and engineering problems. Such meta-heuristics maintain a population of chromosomes, which represent plausible solutions to the target problem and evolve over time through a process of competition and controlled variation. The more adapted an individual is to its environment (i.e., the solution is to the problem), the more likely such individual will be exploited for generating novel individuals. In order to distinguish between adapted and non-adapted individuals, a score function (known as fitness function) should be properly specified beforehand in a manner as to reflect the main restrictions imposed by the problem.

Micro-Social Level for Crime Simulation

Despite the existence of several works on MABS that investigate aggregated features of crime by means of macro-simulation, tools for police

planning have essentially focused on the micro-social level. These latter follow a conceptual framework defined upon the following features: the environment into which the agent is inserted, the agents (their perceptions, productions, transformations, and objects manipulation), and the interaction between them.

The environment is a space, which generally has a volume (Ferber 1999). The agents are situated in the environment or interact with objects inserted into it. In MACS, the environment can be real or artificial. Typically, when the designer wants to reproduce the real environment, Geographic Information Systems (GIS) using digitalized maps of a geographic area for representing, for instance, streets (Groff 2008), are used. Artificial environments reproduce the main features of a geographic area via abstractions such as grids (Brantingham and Tita 2008) or graphs. Despite the claim of Elffers and Van Baal (2008) that the use of real representation in MACS is not so relevant as the quality of the model, there is a trend in this direction mainly because it facilitates evaluation by visual comparisons. Another way to model the environment is by means of cellular automata (CA) typically to represent static objects. CA is a discrete model of regular grids in which the state of a cell at time t is a function of the status of a finite number of neighborhood cells at time $t-1$. Examples of environments modeled as a CA can be found in Liu et al. (2005).

The related work on MACS has typically been based on the routine activity theory (RAT) (Cohen and Felson 1979), which states that in order for a criminal act to take place, three elements must coexist: a motivated offender; a suitable target, either an object or a person that can be attacked; and the absence of capable guardians in charge of preventive actions. Therefore, most of the micro-social crime models are based on these three kinds of agents: criminals, guardians, and targets.

On a micro-social level, the criminal agent tries to commit crimes or to move. Each criminal is endowed with a limited view of the environment, measured in terms of a radius following a predefined unit of measure. Criminals have one

or more points of departure that we are going to call "gateways." Such points of departure represent places where criminals are likely to start out, e.g., their residences, metro stations, and bus stops, before committing crimes (Wright and Decker 1994). Target selection is typically probabilistic, based on factors such as target vulnerability, distance between the criminal and the target, and the criminal's experience. The decision whether to commit a crime or not is made based on the existence of guardians within the radius of the criminal's sight.

The criminal behavior can be modeled with a learning component that exploits the agent's experience as well as with information coming from other criminal agents. The success rate of individual agents can be computed as the ratio of the number of successful crimes to the overall number of crimes attempted in their lifetime, as in Furtado et al. (2008), or based on preferences of criminals computed from data mining and based on discrete choice theory (Xue and Brown 2006). Criminals form communities wherein hints are shared. Due to the interconnection of the communities, such hints could be relayed to other criminals in other communities, and the rate at which this happens depends directly on the topology of the network of communities.

Usually there is a set of guardians (police or not) available, each one associated with a target area. A guardian can have a route of length n , which is defined as a set, and each component of which is a triple composed by the target area, the interval of time the guardian remains at the target, and the daily period (patrol shift) the routes refer to. Guardians can demonstrate deterministic or stochastic behavior. A deterministic guardian will always move to the same target area and at the same pace predefined as an input parameter. Police guardians are modeled following real data, since these data are available and known by police institutions. However, finding good routes can be an important goal in order to understand the impact of police patrol on crime prevention or reduction. In Reis et al. (2006), a genetic algorithm system called GAPatrol is devoted to the specification of effective police patrol route strategies for coping with criminal activities

happening in a given artificial urban environment, which, in turn, mimics a real demographic region of interest. The approach underlying GAPatrol allows for the automatic uncovering of hot spots and routes of surveillance, which, in real life, are usually discovered by hand with the help of statistical and/or specialized mapping techniques.

Notice that, while many simulation studies are aiming at understanding and analyzing the role of various assumptions about the behavior of the respective agents as a function of the ever-changing constellation of other agents and, hence, have a theoretical emphasis (cf. Birks & Elffers, Gerritsen & Klein, this volume), we like to stress here the use of simulations as a means to study the resulting pattern of offender agents and victim agents as a function of various strategies that guardian agents, especially police agents, could have. This use of simulation modeling exploits the method for investigating various possible police strategies on the simulated offence pattern and, hence, has a more applied character.

The locations to be chosen by criminals are referred to as targets, which can be differentiated with respect to their mobility. Commercial/entertainment establishment such as drugstores, banks, gas stations, lottery houses, and malls are fixed, while mobile targets are, e.g., citizens in movement. In Brantingham and Tita (2008), citizen movement is modeled according to the Levy probability distribution, while Furtado et al. (2008) and Liang et al. (2001) have concentrated their study on crimes against property, hence modeling only fixed target.

Targets have a state of vulnerability that can be either active or inactive. A vulnerable target means that it is perceivable by a criminal. Otherwise, it would not take part in the set of high-priority choices of that criminal. In this case, each target must have a probability of being vulnerable, which can follow a distribution based on past real crime data for the associated target type or based on the preference of criminals (Xue and Brown 2006). In Liu et al. (2005), a tension factor was introduced in the model by measuring the impact of crime events on human beings. After a crime in a region, the tension increases, while

the vulnerability decreases. The attractiveness of a target can vary depending on cost and reward factors related to the selection of the target such as the location, income, and race composition of the area based on census data. In Furtado et al. (2008), an exponential temporal distribution is used and varies on a daytime basis. For each period and type of target, a value for a configurable parameter, λ , must be determined at the start of the simulation in order to define the pace of occurrence of crimes. For instance, in the evening, drugstore robberies may follow a distribution based on a given value for λ ; whereas, in daylight periods, the crime temporal distribution might shift, achieving values four times higher for λ . At any simulation tick, at least one target is made vulnerable in accordance with the temporal distribution associated with its related type. The state of vulnerability is essential as a parameter to control the pace of crimes per type as happens in real life. However, one of the limitations of using values for input parameters from historical real-data analyses, e.g., the pace of crimes (the λ parameter), is that the simulation model will not be capable of identifying a change in the pace of crime occurrences if that change occurs during the simulation time. This problem becomes more significant as the simulation time increases.

Modeling the Interaction Between Agents

Direct interaction means that agents communicate with each other by means of message exchange and/or because they are part of a community or society. In order to be part of a society, it is imperative for an individual to establish social links with other peers. Different forms of interaction among the same individuals, even considering small groups, may take place simultaneously and may vary at different paces through time. One usual means to represent and analyze the (evolving) social structure underlying an organization of individuals is by resorting to the concept of social networks. Roughly speaking, a social network alludes to any formal,

graph-based structure where individuals are represented by nodes and the social relationships that unite them are represented by links (ties) between those nodes. The topology of a social network is an important issue to be considered in the analysis thereof, as it helps to determine the network's usefulness (from the viewpoint of the individuals that participate in the network).

The social interaction and learning aspects that underlie criminal activities were investigated in Sutherland's seminal work (Sutherland 1974) in which the differential association theory was proposed. This theory advocates that interaction with others who are delinquent increases the likelihood of someone becoming and remaining a delinquent. That is, peers can play a crucial role in the development of values and beliefs favorable to law violation. In this theory, Sutherland elaborates nine postulates, out of which two are particularly relevant from a perspective of direct agent interaction:

- Criminal behavior is learnable and can be especially learned through the interactions one establishes with other persons, typically through a verbal communication process.
- The main part of the learning of criminal behavior occurs within intimate personal groups.

Another important result coming from works investigating social network models within the context of criminology is that social networks are a natural way to explain the concentration of crimes per area. Crime data analyzed from different regions, and even countries, usually reflect the fact that there are huge spatial (but also temporal) variations in the crime rate between different cities and between different regions in a city. In this regard, Glaeser et al. (1996) show that less than 30 % of the spatial variation of crime (both inter- and intracity) can be explained by differences in local attributes. The remaining 70 % can be explained by social interactions, which means that the agents' decisions about crime are somewhat positively correlated. The authors also show that the impact of social relations is greater in thefts, burglaries, assaults, and robberies (i.e., crimes against property) than in homicides. It is worth mentioning the work of Bosse et al. (2007),

which created a model to simulate social learning of adolescence-limited criminal behavior, and the work of Furtado et al. (2008), which designed a model in which social networks are used to model criminal communication. In the latter, the authors showed that the goal of the crime model of generating a power-law spatial distribution of crimes was correlated to the communication aspect modeled via the social network.

Indirect interaction is modeled by means of objects or variables sharing, typically represented in the environment. Other kinds of non-intentional forms of communication that are sent by diffusion or propagation into the environment, like signals, are also used. Ferber (1999) alerts to the limitation that the lack of semantics of signals can provoke. Since the signal is propagated in the environment, all the agents living there can perceive it. A cry of a citizen can be perceived by a guardian as a call for help as well as used by a criminal as a discovery of a potential prey. One important feature of these signals is that their intensity decreases with the distance from the source and with time. The concepts of tension (Liu et al. 2005) and conductivity (Dray et al. 2008) previously mentioned are examples of indirect interaction.

An example of a hybrid approach that uses direct and indirect communication is that of Furtado et al. (2008). Here, communication between the agents was proposed from the concepts of ant-based optimization augmented with a social network. In this model, criminals prefer to commit crime in locations known to be vulnerable, with high payoff, etc. In other words, their choice considers their preference and knowledge about the crime points. The link here to ACO is that, according to this approach, ants always choose their next location in the environment (the place they move toward) biased by a mechanism (the pheromone marks) of indirect interaction. Another indirect communication strategy that ACO offers is that it includes concepts intrinsically related to the notion of "collective." In ACO, ants perform their local search tasks without dictating the whole colony's behavior, which, in turn, is recognized as an emerging

result coming from all these local activities. Each criminal has three possible actions: commit a crime, not commit a crime, and move to a certain location. In order to reach a decision whether to commit a crime or not, criminals make use of a probabilistic approach which is adapted from the context of ant-based swarm systems (Dorigo and Stützle 2004).

Key Issues

In MABS, the calibration of the model is a crucial step of the design process (Malleon, this volume). These models are typically characterized by the existence of a lot of parameters, which together determine the general behavior of the model. The development and the parameter setting of MABS models can be long and tedious if there is no accurate and systematic manner to explore the parameter space.

Genetic algorithm has been studied as an alternative for parameter calibration in MABS. The basic idea is to consider the tuning process as an optimization problem. The optimization function for the MABS would be the distance between the artificial model and the real system. In MACS, the calibration of certain parameters in which the complexity is high deserves special attention, e.g., the place from where each criminal starts out to commit crimes. Examples of these initial locations, here called gateways, are bus stops, metro stations, and slums. Usually there is no real data or theoretical model to help one configure those gateways in a crime simulation model. Moreover, it is a combinatorial optimization problem, i.e., the problem of assigning criminals to gateways. More formally, let $G = \{G_i, i = 1, \dots, N_g\}$ be the set of gateways and $C = \{C_j, j = 1, \dots, N_c\}$ be the set of criminals under consideration. The goal is to allocate each C_j to a G_i in a way that a quality measure F , somehow related to the aim of the simulation model, is maximized. In this allocation process, any gateway can be assigned to a criminal and all criminals must be allotted to one, and only one, gateway. Besides, more than one C_j can be assigned to a given G_i (i.e., we have not imposed

any limit over the number of criminals assigned to a gateway). Since this assignment problem is combinatorial in nature, the number of feasible gateway configurations is an exponential function of the number of possible gateways.

As important as the calibration of guardians in gateways is the calibration of the number of criminals. There is no real data to help on that and even estimations are very tough to do. Also, this number can change during the simulation or remain constant. Typically in many models, the number of criminals is constant during a simulation. The absence of mechanisms that could implement variability in the number of criminals, such as arrests, can be justified when the ultimate goal is to find good police patrol routes. Considering constant the number of criminals means that crime reduction is only attained by preventing a potential criminal from acting. We could say that, by doing so, there is a preparation of the model to work in the worst of the scenarios, i.e., no way to reduce the number of potential criminals. Thus, the preventive police-planning goal is to define strategies to cover the urban space in a way that could prevent crimes from occurring; events that change the number of criminals are irrelevant.

Social sciences have struggled with this topic due to the difficulty to conduct experiments in controlled environments. MABS have emerged as a tool for social analysis in a way similar to natural controlled experiments. However, model evaluation is one of the biggest challenges of MABS and MACS in particular. A typology for validation must consider the following aspects. The first aspect regards constructing validity to account for the difference between the real world and the rendering of the simulated environment. The challenge here is to design a model representing an approximation that won't be too detailed or similar to the real world because, in this case, the model loses its pragmatic value, and rather than testing theories, it only enumerates what happens under a specific and limited set of conditions. On the other hand, the model must fit the purpose for which it has been created without variables that would bring excessive and unnecessary complexity. The second aspect refers to

the internal validity of the model, also called verification. Basically, it refers here to the reliability of the software in generating a determined result from the inputs and processing function for which it was designed. Minor bugs and ill-defined implementations can be responsible for results that falsify the experiments. The third aspect refers to the external validity, meaning how reliable generalizations of the model are for populations larger than the samples. Statistical conclusion validity is another important aspect to be considered in terms of MACS evaluation. Typically, MACS are stochastic and their variation and unpredictability pose problems in the establishment of statistical validation. The identification of regularity that deviates from chance is essential in that context. This must be done across simulation runs in order to be convincing. Finally, they describe the need for models to be evaluated as to their empirical validity via comparisons with real data. In geosimulation, for instance, visual comparisons between hot spots generated from simulations are plotted on maps for comparison with real-data hot spots. Several studies conducted in different countries and different contexts have shown that crime is not uniformly distributed in space and that some victims or targets have a much greater risk of victimization than others (Pease 1998). The temporal dimension also presents a nonuniform distribution with different types of crime having different rhythms (their periodicity) as well as time (their rate of occurrence). In Johnson et al. (2007), an analysis showing how crime clusters in space and time was provided. Following victimization at one's home, those nearby experience an elevated risk of victimization, which decays as time elapses. Another strategy for validating simulation models is to find patterns in real data that indicate crime distribution instead of only relating to the exact numerical values. Doing so, it is possible to compare the results of the simulation with expected distribution of events.

Notice that the use of simulation for investigating possible police strategies is, to some extent, less hampered by the type of problems we have discussed in the present section. In such applications we simply vary the parameters

governing the guardian (police) agents, in a way that they represent the strategies to be tested. Actual parameters about the police process (e.g., number of officers available at particular times, priorities) may be available from the police force for whom the simulation is run, thus limiting the variability and hence the number of simulations to be run.

Future Directions

Simulation of criminal activities in urban environments is an asset to decision-makers seeking to find preventive measures. Law enforcement authorities need to understand the behavior of criminals and their response in order to establish safety measures and policies. A conceptual framework for micro-social multiagent-based crime simulation was described involving concepts from Computer Science and AI in particular. Special attention was given to calibration and evaluation aspects, since they constitute open issues demanding further investigations and techniques.

Future investigations in the calibration field are in fact advance in terms of decision support systems, since some parameters like a police patrol route and criminal-gateway composition can shed light on non-understood aspects of preventive policing. In this context, such investigations are expected to provide satisfactory answers to questions like: How far from the optimal patrol routing strategies are those that are actually adopted by human police managers? How complex do such optimal patrolling routes need to be in terms of their total lengths and urban area coverage?

Related Entries

- ▶ [Agent-based Modeling for Understanding Patterns of Crime](#)
- ▶ [Agent-based Models to Predict Crime at Places](#)
- ▶ [Crime Mapping](#)
- ▶ [Offender Decision Making and Behavioral Economics](#)
- ▶ [Predictive Policing](#)

Recommended Reading and References

- Benenson I, Torrens PM (2004) Geosimulation: object-based modeling of urban phenomena. *Comput Environ Urban Syst* 28(1/2):1–8
- Bonabeau E, Dorigo M, Herculaz G (1999) Swarm intelligence: from natural to artificial systems. Santa Fe Institute Studies in the Sciences of Complexity Series. Oxford Press
- Bosse T, Gerritsen C, Treur J (2007) Cognitive and social simulation of criminal behavior: the intermittent explosive disorder case. *AAMAS* 58:367–374
- Brantingham P, Brantingham P (1979) Environment, routine, and situation: toward a pattern theory of crime. In: Clark R, Felson M (eds) *Routine activity and rational choice*, vol 5. Transaction books, pp 259–294
- Calvez B, Hutzler G (2005) Automatic tuning of agent-based models using genetic algorithms. Fourth international joint conference on autonomous agents & multi agent system, Netherlands
- Cohen L, Felson M (1979) Social change and crime rate trends: a routine approach. *Am Sociol Rev* 44:588–608
- Dorigo M, Stützle T (2004) Ant colony optimization. The MIT Press, Cambridge
- Elffers H, Van Baal P (2008) Realistic spatial backcloth is not that important in agent based simulation research: an illustration from simulating perceptual deterrence. In: Liu L, Eck J (eds) *Artificial crime analysis systems: using computer simulations and geographic information systems*, pp 19–34
- Ferber J (1999) *Multi-agent systems: an introduction to distributed artificial intelligence*. Addison-Wesley
- Furtado V, Melo A, Coelho AL, Menezes R (2008) Simulating crime against properties using swarm intelligence and social networks. In: Liu L, Eck J (eds) *Artificial crime analysis systems: using computer simulations and geographic information systems*, pp 300–318
- Gilbert N, Conte R (1995) *Artificial societies: the computer simulation of social life*. UCL Press, London
- Glaeser E, Sacerdote B, Scheinkman J (1996) Crime and social interactions. *Q J Econ* 111(2):507–548, MIT Press
- Groff E (2008) Characterizing the spatio-temporal aspects of routine activities and the geographic distribution of street robbery. In: Liu L, Eck J (eds) *Artificial crime analysis systems: using computer simulations and geographic information systems*, pp 226–251
- Holland JH (1975) *Adaptation in natural and artificial systems*. University of Michigan Press
- Johnson S, Bernasco W, Bowers K, Elfers H, Ratcliffe J, Rengert G, Townsley M (2007) Space-time patterns of risk: across national assessment of residential burglary victimization. *J Quant Criminol* 32(3):201–219
- Liang J, Liu L, Eck J (2001) Simulating crimes and crime patterns using cellular automata and GIS. UCGIS student award papers 2001. Retrieved from <http://www.ucgis.org/f2oppor.html>
- Liu L, Wang X, Eck J, Liang J (2005) Simulating crime events and crime patterns in a RA/CA Model. In: Wang F (ed) *GIS and crime analysis*. Idea Group, pp 197–213
- Pease K (1998) Repeat victimization: taking stock. Crime detection and prevention series paper 90. Home Office, London
- Reis D, Melo A, Coelho ALV, Furtado V (2006) GAPatrol: an evolutionary multiagent approach for the automatic definition of hotspots and patrol routes. In: Sichman JS, Coelho H, Oliveira S (eds) *Proceedings of IBERAMIA/SBIA 2006, Lecture Notes in Artificial Intelligence (LNAI) 4140*, pp 118–127
- Rocher G, Sherif P (1972) *A general introduction to sociology: a theoretical perspective*. MacMillan Canada
- Russell S, Norvig P (1995) *Artificial intelligence: a modern approach*. Prentice Hall Series in AI, New Jersey
- Scott JP (2000) *Social network analysis: a handbook*. Sage
- Snook B (2004) Individual differences in distance travelled by serial burglars. *J Invest Psychol Offender Profiling* 1:53–66
- Sutherland E (1974) *Principles of criminology*, 4th edn. J. B. Lippincott, Philadelphia
- Wright R, Decker S (1994) *Choosing the target. Burglars on the job: street life and residential break-ins*. Northeastern University Press, Boston
- Xue Y, Brown D (2006) Spatial analysis with preference specification of latent decision makers for criminal event prediction. *Decis Support Syst* 41(3):560–573

Situational Action Theory

Per-Olof H. Wikström
Institute of Criminology, University
of Cambridge, Cambridge, UK

Synonyms

[Analytical criminology](#); [Causal processes](#); [Crime](#); [Deterrence](#); [Morality](#); [Moral norms](#); [Self-control](#)

Overview

This chapter introduces Situational Action Theory (SAT) and its key arguments:

Crimes are moral actions. There are many different types of actions that may constitute

a crime, but what they all have in common is that they break a rule of conduct (stated in law) about what is the right or wrong thing to do (or not to do). What distinguishes crime is thus not that they are a particular kind of action but that they are acts that breach rules of conduct (stated in law) and, therefore, need to be studied and explained as such.

People are the source of their actions, but the causes of their actions are situational. To explain the causes of acts of crime we need to understand the situational factors and processes that move people to break rules of conduct (stated in law). People are different and so are also the settings (environments) to which they are exposed. Acts of crime (particular acts of crime) occur when particular kinds of people (person propensities) encounter particular kinds of settings (environmental inducements) creating specific kinds of interactions that make people see certain action alternatives and make certain choices.

Social and developmental factors and processes are best studied and explained as causes of the causes. To explain the role of social and developmental factors and processes in crime causation, we need to understand how they (as causes of the causes) affect people's crime propensities and how they impact settings (environments) levels of criminogeneity and how they influence the spatiotemporal convergence of crime-prone people and criminogenic settings (creating the interactions that may result in acts of crime or certain types of acts of crime).

Situational Action Theory: Background and Main Aims

Situational Action Theory (SAT) aims to explain *why* crime happens and the role of social and developmental factors and processes in crime causation (e.g., Wikström 2006, 2010; Wikström et al. 2012, pp. 3–43). It was specifically designed to overcome some recurrent problems in criminological theorizing such as the unclear

definition of crime (i.e., the ambiguity about what a theory of crime aims to explain), the lack of a satisfactory action theory (i.e., the poor understanding of what it is that moves people to engage in acts of crime), and the poor integration of levels of explanation (e.g., Wikström 2010, pp. 212–216; Wikström et al. 2012, pp. 3–10). The argument is not that all criminological theories fail on all these points, but it is just that they generally fail on one or more of these points.

SAT seeks to integrate key insights from main person and environment-oriented criminological theories and research, and relevant social and behavioral science theories and research more generally, within the framework of an adequate action theory. SAT advocates an analytical criminology, a criminology that focuses on the identification and detailing of the key mechanisms involved in crime causation rather than the production of lists of statistically relevant correlates and predictors (often referred to as risk factors) of which most, at best, are likely to be only markers or symptoms and, therefore, lack any causal relevance (e.g., Wikström 2006, pp. 66–69, 2011, pp. 53–60).

Basic Assumptions About Human Nature and Social Order

Situational Action Theory is based on the key assumptions that humans are essentially *rule-guide actors* and that social order is fundamentally based on people's adherence to shared rules of conduct. People vary in their personal morals (personal rules of conduct) and settings vary in their moral norms (shared rules of conduct). A personal moral rule is held and enforced (through the process of self-control) by the actor, while a moral norm is held and enforced (through the process of deterrence) by (significant) others. Explaining human action (such as acts of crime) requires understanding how people's actions are influenced by the interplay between personal morals and perceived moral norms of the setting (environment), a problem that has been largely ignored in criminological theory.

SAT assert that humans have *agency* (powers to make things happen) but that they exercise agency (i.e., express their desires (needs) and their commitments and respond to frictions) within the constraints of rule-guided choice. The theory acknowledges that there are elements of predictability (habit) and “free will” (deliberation) in people’s choices. Explaining human action (such as acts of crime) requires an understanding of the role of agency (how it works) within rule-guided choice, a problem that has seldom (if at all) been treated in criminological theory.

Crime as Moral Actions

Criminological theories rarely specify or clearly analyze what it is they aim to explain. This is an important common omission since an explanation has to explain *something*. Without clearly defining what it is that is to be explained, it is difficult to unambiguously identify putative causes and suggest plausible causal processes that may produce the effect under study (e.g., acts of crime). A cause has to be a cause of something, and what that something is determines possible causes (and relevant causal processes).

Situational Action Theory argues that crime is best analyzed as moral actions. SAT defines morality as value-based rules of conduct about what is the right or wrong thing to do (or not to do) in a particular circumstance. *Crime* is an act that breaks a rule of conduct stated in law (rules that may be quite general or quite specific). What defines crime is thus not any particular type of action but the fact that carrying out a particular action (or refraining from carrying out an action) in a particular circumstance is regarded as breaching a rule of conduct (stated in the law).

Any particular action can, in principle, be defined as a crime, and there are variations over historic time and between places (e.g., countries) in what kinds of action are regarded as crime. Moreover, specific actions, like hitting or even shooting another person, may be considered crimes in some circumstances but not in others.

The advantage of conceptualizing crime as breaches of rules of conduct (stated in law) is that it makes a general theory of crime possible by focusing on what all kinds of crime, in all places, at all times, have in common, namely, the rule-breaking. What is to be explained by a theory of crime causation is thus why people (follow and) breach rules of conduct stated in law. A theory of crime causation may be regarded as a sub-theory of a more general theory of moral action. The law is just one of many sets of rules of conduct that guide people’s action (e.g., Ehrlich [1936] 2008). The law is no different from other sets of rules of conduct; in fact, the law may be regarded as a special case of rules of conduct more generally. Explaining why people follow and break the rules of law is, in principle, no different from explaining why people follow and break rules of conduct more generally.

Analyzing crime as moral action does not imply a “moralistic” approach (no judgement whether existing laws are good or bad, just or unjust, needs to be made). However, SAT does not assume moral relativism (i.e., that all moral rules are equally likely to emerge). Common moral rules may be grounded in human nature and the problem of creating social order (see further Wikström et al. 2012, pp. 13–14). The explanation of why we have particular moral norms (and laws) and why people follow or breach moral norms (and laws) is different analytical problems. Importantly, people may breach a moral rule of conduct (stated in law) because they disagree with or care less about the rule (or their ability to exercise self-control is not strong enough to make them adhere to their own personal morals when faced with a temptation or provocation).

The Situational Model: Kinds of People in Kinds of Settings

People are different and so are the settings to which they are exposed. When a particular kind of person is exposed to a particular kind of setting, a particular situation (perception-choice process) arises that initiates and guide his or her

actions in relation to the motivations he or she may experience. The elements of the situational model are the *person* (his or her relevant propensities), the *setting* (its relevant inducements), the *situation* (the perception-choice process that arises from the exposure of a particular person to a particular setting), and *action* (bodily movements such as speaking, walking, or hitting). A setting is defined as the part of the environment (objects, persons, events) that is directly accessible to the person through his or her senses (including any media present).

SAT proposes that variations between people in their *crime propensity* (i.e., the tendency to see, and choose, particular crimes as an action alternative) is essentially a question of their *law-relevant personal morality* (the extent to which their personal morality corresponds to the various rules of conduct stated in the law) and their ability to exercise self-control (which depends both on dispositional characteristics such as executive functions, and momentary influences such as intoxication and levels of stress – see Wikström and Trieber 2007).

People will vary in their propensity to see a particular kind of crime as an action alternative not only depending on their personal moral rules but also depending on their moral emotions. *Moral emotions* (shame and guilt) attached to violating a particular rule of conduct may be regarded as a measure of the strength with which a person holds that particular rule of conduct. For example, while many people may think it is wrong to steal something from another person, some may feel very strongly about this while others may not. Those who feel less strongly about stealing from others may be regarded as having a higher propensity to perceive such action as an option.

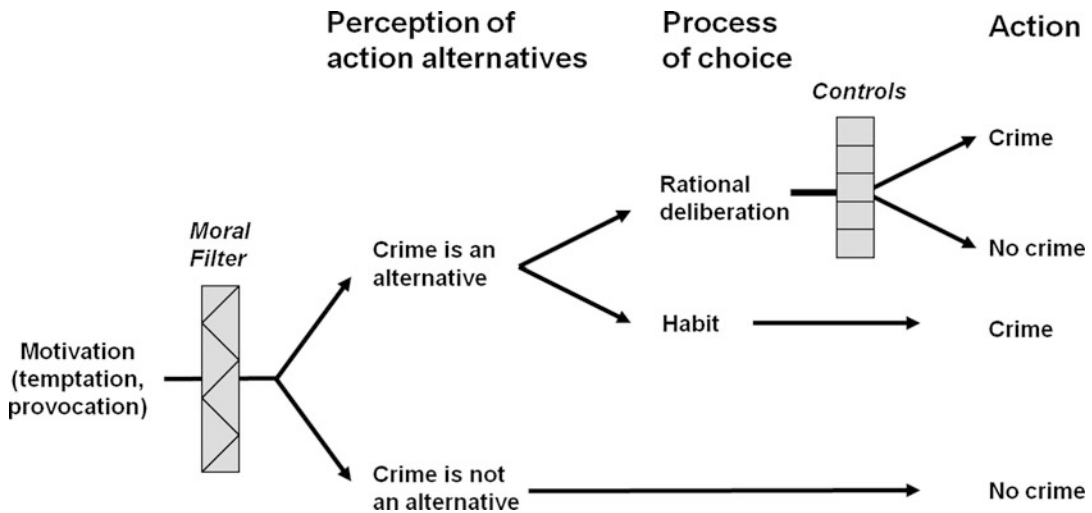
People vary in their *ability to exercise self-control* (i.e., their ability to act in accordance with their personal morality when it conflicts with a perceived moral norm of a setting). The ability to exercise self-control exerts its influence through the process of choice and is only important in the explanation of crime when a person deliberate over several potent action alternatives, of which at least one involve an act of crime. The

concept of self-control in SAT is different from that of self-control in Gottfredson and Hirschi's (1990) General Theory of crime (see further, e.g., Wikström 2010, pp. 228–234). Crucially SAT makes a clear distinction between self-control as situational process (the process by which a person manage conflicting rule guidance) and people's abilities to exercise self-control.

SAT further proposes that the *criminogeneity of a setting* depends on its (perceived) moral norms (the extent to which it encourage or discourage the breaking of particular laws in relation to the opportunities a setting provides and the frictions it creates) and their level of enforcement through the process of deterrence (note that if a moral norm encourages the breaking of a particular law, a high degree of its enforcement will be criminogenic). *Deterrence* is defined as “the process by which the (perceived) enforcement of a setting's (perceived) moral norms, by creating concern or fear of consequences, succeeds in making a person adhere to the moral norms of the setting even though they conflict with his or her personal moral rules.”

Key Situational Factors and the Situational Process

The cornerstone of Situational Action Theory is that people's actions ultimately is a consequence of how they perceive their action alternatives and, on that basis, makes their choices when confronted with the peculiarities of a setting. The perception-choice process is crucial for understanding a person's actions. *Perception* (the selective information we get from our senses) is what links a person to his or her environment, and *choice* (the formation of an intention to act in one way or another) is what links a person to his or her actions (see further Wikström 2006, pp. 76–84). In contrast to most choice-based theories of action, which focus on how people choose among predetermined alternatives, Situational Action Theory stresses the importance of why people perceive certain action alternatives (and not others) in the first place. Perception of action alternatives thus



Situational Action Theory, Fig. 1 The situational process and key situational factors (applied to the explanation of crime) (Source: Wikström 2011)

plays a more fundamental role in explaining actions, such as acts of crime, than the process of choice (which is secondary to perception of action alternatives). The key stages of, and the key situational factors in, the perception-choice process, according to SAT, is illustrated in Fig. 1.

Motivation (defined as goal directed attention) is a situational factor that initiates the action process and is an outcome of the interaction between the person (preferences, commitments, sensitivities) and the setting (opportunities, frictions). According to SAT there are two main kinds of motivators; (1) temptations, which are either the outcome of the interaction between (a) a person's desires (wants, needs) and opportunities to satisfy a desire (want, need) or (b) the outcome of the interaction between a person's commitments and opportunities to fulfill a commitment (opportunities to satisfy a desire or to fulfill a commitment may be legal or illegal), and (2) provocations, which occur when a friction (an unwanted external interference) causes anger or annoyance towards the perceived source of the friction or a substitute. People vary in their sensitivities to particular kinds of frictions (as a consequence of their cognitive-emotive functioning and life-history experiences).

Motivation is a necessary but not sufficient factor in the explanation of crime. What particular action alternatives a person perceives in relation to a particular motivation (and whether that includes an act of crime) is dependent on the moral filter. The *moral filter* is defined as "the moral rule-induced selective perception of action alternatives in relation to a particular motivation" and is an outcome of a person's moral engagement with the perceived moral norms of a setting in relation to a specific motivation. If crime is not among the action alternatives a person perceives, there will be no crime and the process of choice is irrelevant as an explanation of why he or she refrained from committing an act of crime; he or she simply did not see crime as an option.

If an act of crime is among the perceived action alternatives the process of choice will determine whether or not a person will commit (or attempt) an act of crime. SAT distinguishes between two main kinds of processes of choice: habitual (automated) or deliberative processes of choice (in prolonged action sequences the action guidance may drift between habitual and deliberative influences). The important difference between the two processes is that in case of actions (such as acts of crime) committed out of habit, the actor only sees *one* potent alternative in

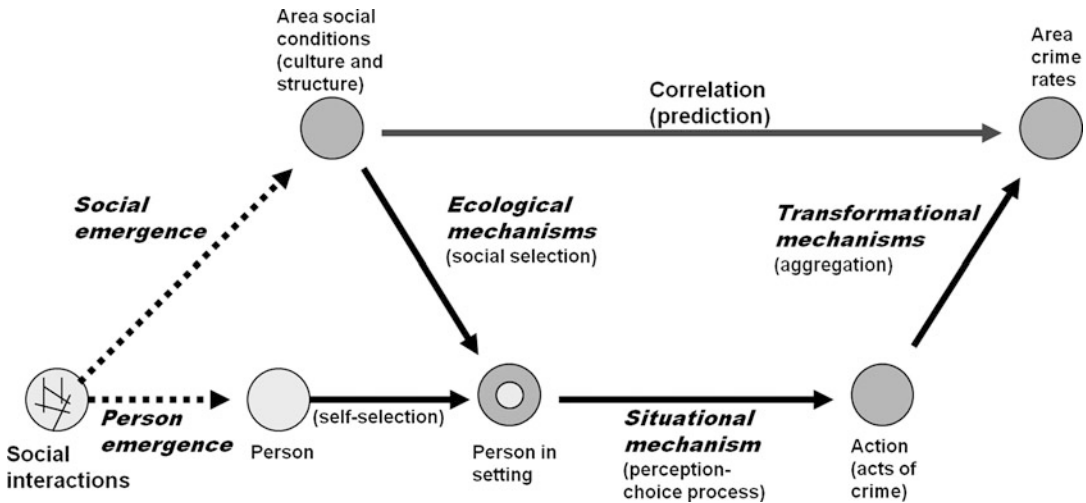
response to a motivation (temptation or provocation) and act accordingly, while in the case of deliberation there is no predetermined response so the actor has to weigh pros and cons of the different perceived alternatives (a process that may be more or less elaborate dependent on the importance of the choice). There is plenty of evidence for the existence of a dual process of human reasoning of this kind (see, e.g., Evans and Frankish 2009). The neuropsychological basis of SAT's framework for habitual and deliberate decision making, linking these processes to areas of the brain's prefrontal cortex which have been associated with separate intuitive and cognitive functions, with particular implications for the role of self-control and emotions in action decisions, has been specifically explored by Treiber (2011).

When people act out of *habit*, they essentially react (in a stimulus-response fashion) to environmental cues, automatically applying experienced based moral rules of conduct to the peculiarities of a setting, while *deliberation* involves taking moral rules of conduct into consideration when actively choosing between perceived action alternatives. Habitual action is oriented towards the past, while deliberate action is oriented towards the future. When people act habitually, they routinely apply past experiences to guide current action (i.e., they do what they normally do in the particular circumstance without giving it much thought); when they act deliberately, they try to anticipate future consequences of perceived action alternatives and choose the best course of action. When people act out of habit (only see one causally effective alternative; although they may be loosely aware "in the back of their minds" that there are other alternatives, no such alternative is actively considered), rationality does not come into play (since there is only one potent alternative and, therefore, no weighing of pros and cons of different options) and their actions may even be irrational, that is, people may act in ways they would not consider in their best interest had they deliberated. When people deliberate they aim to be rational (to choose the best option as they see it), but, importantly, this does not necessarily involve an aim to maximize personal advantage

(whether maximizing personal advantage is seen as the best option is a question of the actor's moral judgement). Habitual responses are most likely when people operate in familiar circumstances with congruent rule guidance (i.e., when there is a high level of correspondence between personal morals and perceived moral norms of the setting), while deliberate responses are most likely when people are in unfamiliar circumstances and/or there is a conflicting rule-guidance.

Only in choice processes when people deliberate do they exercise "free will" (although it is a "free will" within the constraints of perceived action alternatives) and are they subject to the influence of their ability to exercise self-control (internal control) or respond to deterrence cues (external controls). If there is no conflicting rule guidance, there is nothing to control, and, hence, in these cases (inner and outer) controls lack relevance for the action taken. A person's ability to exercise self-control is a relevant factor in crime causation when his or her personal moral rules discourage and the moral norms of the setting encourage an act of crime in response to a motivation. In this case, the extent to which a person can act in accordance with his or her personal morals (i.e., refrain from crime) is dependent on the strength of his or her ability to exercise self-control. Deterrence is a relevant factor in crime causation when a person's personal morals encourage and the moral norms of a setting discourage an act of crime in response to a motivation. In this case, whether or not a person will carry out an act of crime is dependent on the (perceived) efficacy of the enforcement of the moral norms of the setting.

SAT offer a general explanation of key situational factors and processes involved in crime causation. What may differ in *the explanation of different kinds of acts of crime* is not the process (the perception-choice process) leading up to the action, but the input to the process, that is, the content of the moral context (the action-relevant moral norms) and a person's morality (the action-relevant moral values and emotions) which drive the process. For all kinds of crime to happen (e.g., shoplifting, insider trading, rape, or roadside



Situational Action Theory, Fig. 2 The social and situational models of SAT integrated (Source: Wikström 2011)

bombing), the actor has to perceive the particular action as a viable alternative and choose to carry out the act. However, the specific moral background (relevant personal morals and moral norms of the setting) which guide whether, for example, an act of shoplifting is perceived as an action alternative may differ from that which guides whether an act of rape is perceived as an action alternative.

The Social Model

Situational Action Theory insists that the causes of crime are situational and that social and developmental factors and processes in crime causation are best studied and analysed as the *causes of the causes* (of acts of crime). SAT advocates a mechanistic explanation of human action (such as acts of crime). The theory is based on four key propositions (of which the first two refer to the situational model and the subsequent two to the social model):

1. *Action is ultimately an outcome of a perception-choice process.*
2. *This perception-choice process is initiated and guided by relevant aspects of the person-environment interaction.*
3. *Processes of social and self-selection place kinds of people in kinds of settings (creating particular kinds of interactions).*

4. *What kinds of people and what kinds of environments (settings) are present in a jurisdiction is the result of historical processes of person and social emergence.*

Figure 2 illustrates how the social and situational models are integrated. The key content of the situational model has already been discussed, and the remaining part of this paper will be devoted to a brief outline of the role of processes of emergence and selection in crime causation.

The social model of SAT focuses on the role of historical processes of emergence in the creation of criminogenic environments (social emergence) and crime prone people (person emergence) and contemporaneous processes of self- and social selection that bring together crime prone people and criminogenic settings (creating the situations to which people may respond to motivators by committing acts of crime). The concept of *emergence* refers to how something becomes as it is (e.g., Bunge 2003). For example, how people acquire a certain crime propensity (personal emergence) or environments acquire a certain criminogeneity (social emergence) as an outcome of processes of social interactions. The concept of *selection* refers to the contemporaneous socio-ecological processes responsible for introducing particular kinds of people to particular kinds of settings (and thus creating the situations to which people's actions are a response).

SAT proposes that *psychosocial processes* of moral education and cognitive nurturing are of central interest in the explanation of why people develop specific and different crime propensities (i.e., tendencies to see and choose particular crimes as an action alternative) (see further Wikström et al. 2012, pp. 31–32). SAT further proposes that *socio-ecological processes* (e.g., processes of segregation and their social consequences) become of particular interest in the explanation of why particular kinds of moral contexts emerge in particular places at particular times (see further Wikström et al. 2012, pp. 2–37).

Why certain kinds of people end up in certain kinds of setting depends on processes of social and self-selection and their interaction. *Social selection* refers to the social forces (dependent on systems of formal and informal rules and differential distribution of personal and institutional resources in a particular jurisdiction) that encourage or compel, or discourage or bar, particular kinds of people from taking part in particular kinds of time and place-based activities. *Self-selection* refers to the preference-based choices people make to attend particular time and place-based activities within the constraints of the forces of social selection. What particular preferences people have developed may be seen as an outcome of their life-history experiences. Depending on the circumstances, social or self-selection can be more influential in explaining why a particular person takes part in a particular setting (see further Wikström et al. 2012, pp. 37–41).

Analyzing relevant processes of emergence and selection and how they relate help us understand (i) why people develop certain kinds of crime propensities (and why a population in a particular jurisdiction come to have a certain distribution of crime propensities and why some jurisdictions have more crime prone people than others), (ii) why certain places become criminogenic environments (and why there is a certain prevalence and kind of criminogenic environments in a jurisdiction and why some jurisdictions have more criminogenic

environments than others), and (iii) why certain contemporaneous self- and social selection processes create particular kinds of crime hot spots (and why the prevalence and kind of crime hot spots vary between different jurisdictions).

Related Entries

- ▶ [Criminology of Place](#)
- ▶ [Rational Choice Theory](#)
- ▶ [Routine Activities Approach](#)
- ▶ [Situational Crime Prevention](#)
- ▶ [Social Control and Self-control Through the Life Course](#)

Recommended Reading and References

- Bunge M (2003) Emergence and convergence: quantitative novelty and the unity of knowledge. University of Toronto Press, Toronto
- Ehrlich E (1936) 2008. Fundamental principles of the sociology of law. New Brunswick: Transaction Publishers
- Evans J, Frankish K (2009) In two minds: dual processes and beyond. Oxford University Press, Oxford
- Gottfredson M, Hirschi T (1990) A general theory of crime. Stanford University Press, Stanford
- Treiber K (2011) The neuroscientific basis of situational action theory. In: Beaver KM, Walsh A (eds) The Ashgate research companion to biosocial theories of crime. Ashgate Publishing Company, Burlington
- Wikström P-O (2006) Individuals, settings and acts of crime. Situational mechanisms and the explanation of crime. In: Wikström P-O, Sampson RJ (eds) The explanation of crime: context, mechanisms and development. Cambridge University Press, Cambridge
- Wikström P-O (2010) Explaining crime as moral action. In: Hitlin S, Vaysey S (eds) Handbook of the sociology of morality. Springer, New York
- Wikström P-O (2011) Does everything matter? Addressing the problem of causation and explanation in the study of crime. In: McGloin J, Sullivan CJ, Kennedy LW (eds) When crime appears: the role of emergence. Routledge, London
- Wikström P-O, Trieber K (2007) The role of self-control in crime causation. Beyond Gottfredson and Hirschi's general theory of crime. Eur J Criminol 4:237–264
- Wikström P-O, Oberwittler D, Trieber K, Hardie B (2012) Breaking rules. The social and situational dynamics of young people's urban crime. Oxford University Press, Oxford

Situational Approaches to Terrorism

Graeme R. Newman

School of Criminal Justice, University at Albany,
State University of New York, Albany, NY, USA

Overview

The situational approach to terrorism evolved from situational crime prevention, a well-established evidence-based approach to preventing and reducing crime. It argues that terrorists make choices that are limited by their perception of the opportunities afforded to them to carry out their mission. The situational approach explains *how* terrorists carry out their missions in contrast to *why* they do so. There are four pillars of terrorist opportunity: targets, weapons, availability of tools for doing attacks, and local conditions that facilitate terrorist operations. By manipulating these opportunities, interventions are directed at increasing the effort of mounting an attack, increasing the risk of carrying out an attack, reducing the rewards of attacks, reducing provocations to terrorists, and removing excuses terrorists use to justify their attacks. Because the situational approach focuses primarily on prevention, it places great emphasis on planning, which makes it a natural supplement to disaster response planning and to problem-oriented policing that seeks to solve crime and disorder problems. It is thus particularly useful to local police as a guide for developing local counterterrorism strategies.

Introduction

Situational approaches to terrorism derive from several decades of research, theory, and practice in the field of situational crime prevention. Pioneered by Ronald V. Clarke in 1988 in his seminal paper on coal gas suicides, situational crime prevention has evolved in many respects outside of mainstream criminology,

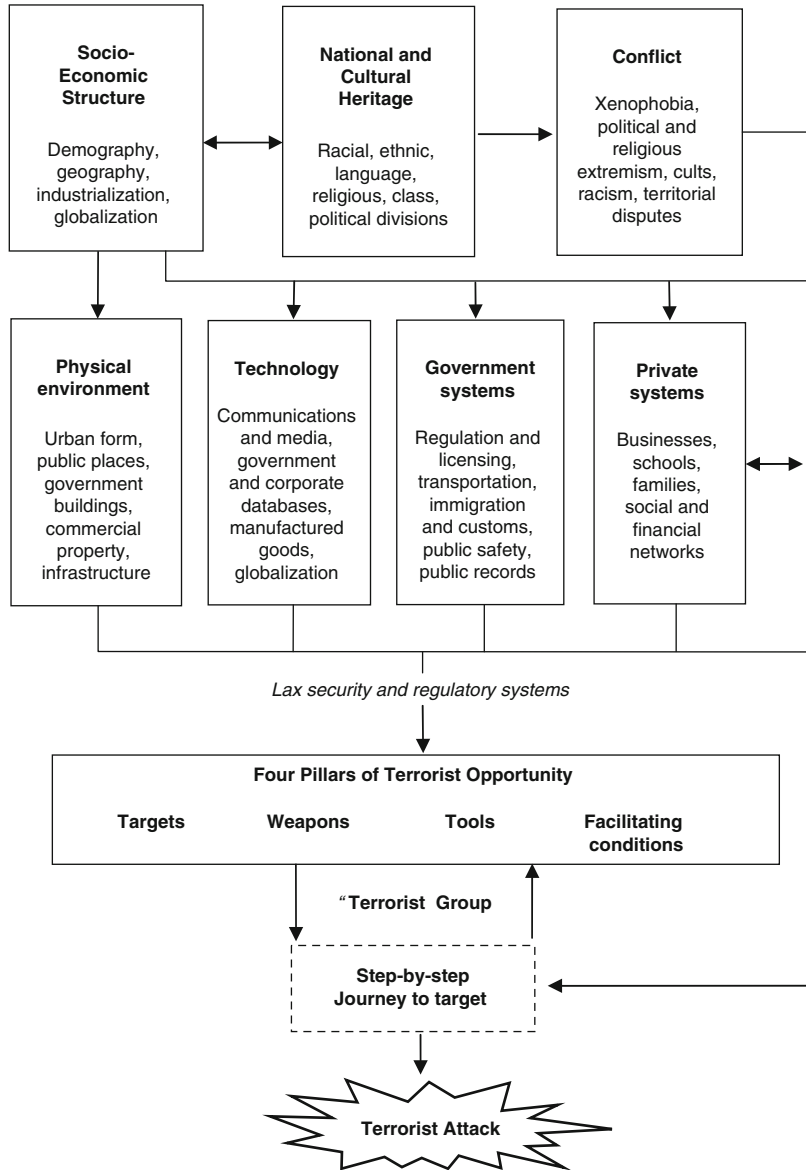
which, since its inception in the nineteenth century, has been preoccupied with the root causes of crime, the search for the “born” or “made” criminal. Dominated by political science, the approach to terrorism has paralleled that of criminology: the search for the root causes of terrorism, with an overwhelming focus on its ideological, psychocultural, economic, and political causes. In contrast, the situational approach takes up where mainstream explanatory models leave off: it begins with a pinpoint focus on the specific situations in which terrorism occurs and, depending on the type of terrorism, works backwards to the causes. It does not completely eschew root causes; it views them as distant, background factors which for the most part are impermeable to counterterrorism interventions (see Fig. 1).

This conceptual distinction between situational crime prevention and other approaches in social science drives it to ask questions that others do not: how did the offender get to the point of being able to carry out his crime? What immediate factors in the social and physical environment produced a situation in which the crime was possible? Situational crime prevention asks not *why* but *how* an offender carries out his tasks. It is concerned with *motives* rather than *motivation*. It is concerned with *intervention* rather than *causation*. It takes seriously the old adage “prevention is better than cure.”

Situational crime prevention is often attacked by those who hold to a contrasting view: “You can’t cure an illness (disease, crime, terrorism) by treating its symptoms.” The symptoms may disappear, but then reappear in some other form because the root causes have not been diagnosed. Counterterrorist experts complain, “If we protect this government building, they’ll attack a shopping mall.” This seems like a devastating blow to the situational crime prevention approach, but it is not as will be seen later in this entry. First, it is necessary to address the basic, intractable problem in the study of the causes and response to terrorism: its definition.

Situational Approaches to Terrorism,

Fig. 1 The terrorist opportunity structure (Adapted from *Outsmarting the Terrorists*, p. 8)



The Situational Definition of Terrorism

The definition of terrorism is famously elusive. The United Nations has tried for many years to agree on a definition and failed. The problem is political, legal, and conceptual. The political difficulty is reflected in the popular saying, "one man's terrorist is another man's freedom fighter." The conceptual problem is tied up with the legal (which essentially binds the conceptual to the political). For example, serious political

disagreements arose in the United States when the Obama administration chose to define the terrorist attacks on the World Trade Center as criminal acts and therefore subject to the regular due process of US criminal law. Opponents of that view argued that they were not criminal acts but acts of war against the USA, so that other legal procedures should follow. Others claimed that the attacks were not acts of war or criminal acts but something else – the position eventually favored by the Bush administration and resulted

in setting up the prison in Guantanamo Bay, a kind of legal “nowhere land.”

The situational approach finds these debates as mostly irrelevant to its enterprise. In fact, as noted earlier, situational crime prevention began with a study of suicide, an act arguably not always a crime, depending on the jurisdiction and circumstances of its commission. As it evolved, situational crime prevention developed more and more a focus on the specificity of the crime which often cuts across the traditional definitions of crime. This had important ramifications in respect to how crime was measured and recorded. Simply recording and reporting the number of burglaries, for example, in a police jurisdiction, were found to be useless for research or for prevention. It was necessary to collect very specific information about very specific types of burglaries: single-family residences or apartment blocks? Suburban or urban neighborhoods? Commercial or business districts? Adjacent to parks or woods? and so on. Generally speaking, it is the definition of the situation that is important for both research and prevention purposes, rather than the definition of the crime. So for the purposes of situational crime prevention, it does not really matter what one calls it, a crime, terrorist attack, or even an accident; it is the analysis of the situation in which these events occur that is of prime importance. As a result, Clarke and Newman in their book *Outsmarting the Terrorists* sidestep the issue by defining terrorism simply as “crime with a political motive.”

The Situational Approach to Terrorism

Apart from the importance of specificity already discussed, there are three basic operating principles that play an important part in any analysis of terrorism from the situational perspective. These are opportunity, rational choice, and intervention.

1. *Opportunity*. This principle simply states that in order for a terrorist attack to be accomplished, the environmental and social conditions must be such that they offer advantage to the terrorist for carrying out the attack. For example, before there were cars, car theft

was not possible. Before there were computers, software piracy was not possible. It is true that “theft” and “piracy” in the generic sense have always been possible, but it is clear that to speak of preventing or even analyzing such crimes, it makes little sense without examining their specific attributes. Furthermore, the social and physical environment may be structured in such a way that easy access to the targets of crime or terrorism is available. For example, early computers or cars were never designed in anticipation of software piracy or car theft. The mosaic of factors that make various crimes or terrorist attacks possible is called the opportunity structure as shown in Fig. 1.

2. *Rational choice*. Depending on what the terrorist wants to achieve by his attack (i.e., his motives), the terrorist will carry out his enterprise according to a rational choice from available options (i.e., part of the opportunity structure). The observer may consider what the terrorist is doing to be “irrational” such as the terrorist who kills himself in order to successfully complete an attack. This is called “limited” rationality as expounded by Cornish and Clarke in their groundbreaking work *The Reasoning Criminal*. Failure to understand the way terrorists think contributed to the short-sighted US counterterrorist policy which assumed that terrorists would not hijack an airplane and willingly kill themselves in the process – the unanticipated methodology of the 9/11 attacks.
3. *Intervention*. Situational crime prevention always begins from the point of how to disrupt the sequence of events that result from the offender’s decision making. This involves generally two procedures for revealing points of weakness in the offender’s or terrorist’s action sequence. First, a step-by-step analysis is conducted to find out exactly the sequence of choices made by the terrorist in his journey to the target. For example, in the case of a suicide bomber, we must follow his journey from induction into the pool of potential bombers, his training, the selection of the target, and finally his journey to the target (see Fig. 2).

Situational Approaches to Terrorism,

Fig. 2 Suicide bombing step-by-step (Adapted from *Outsmarting the Terrorists*, Table 5.3)

1. *Preparation*: Find safe house(s) for operations → Obtain or manufacture bomb paraphernalia → Select target or targets → Recruit bomber candidates → Specify exact location for detonation → Specify route to target → Establish group commitment → Train bombers → Prepare propaganda → reinforce individual commitment →
2. *Operation*: Intelligence gives “all clear” to target → Dispatch bomber → Bomber journeys to target →
3. *Follow up*: Claim responsibility → Broadcast propaganda → Review operation → Plan new attack → Establish supply chain for weaponry → To step 1

The journey to crime has been studied for a wide range of crimes including burglary, car theft, sex offenses, and robbery. Recent work emphasizing the step-by-step approach has adopted the concept of “scripting” from cognitive psychology which examines in detail the decision-making process and techniques used by offenders to exploit opportunities and accomplish their task. This approach has been used in analyzing credit card and check fraud, suicide bombing, Internet child pornography, identity theft, and trafficking in endangered species. The second important aspect of intervention is to conduct a careful assessment of the opportunity structure to identify weak points of security that may be exploited by criminals or terrorists. For example, particular targets may be more vulnerable, exposed, or accessible than others. Once these have been identified, the appropriate security systems may be installed. This is an important feature of situational crime prevention because it implies that there is much that can be done to protect targets even without knowing who the terrorists are or what their motivation may be.

Guided by the three principles of opportunity, rational choice, and intervention, Professors Clarke and Newman have developed a comprehensive approach to explaining terrorism which also provides a structure for organizing responses to terrorism. This helps to connect the often disjointed study of terrorism that focuses either entirely on the terrorist’s characteristics or entirely on law enforcement responses to those characteristics. Instead, Clarke and Newman advocate that we “think terrorist” in order to identify a modus operandi

according to the opportunities available. The terrorist is placed within the opportunity structure of terrorism which is primarily composed of what they call the four pillars of terrorist opportunity: targets, weapons, tools, and facilitating conditions.

The Four Pillars of Terrorist Opportunity

1. Targets

In theory, terrorists could attack any target, as many politicians and law enforcement individuals have worried ever since 9/11. In practice, they must choose their targets carefully. And this choice will be strongly conditioned by the inherent attractiveness of those targets to terrorists, summarized as EVIL DONE:

Exposed: The Twin Towers were sitting ducks.

Vital: Electricity grids, transportation systems, and communications are vital to all communities.

Iconic: Of symbolic value to the enemy, for example, Statue of Liberty, the Pentagon, and the Twin Towers.

Legitimate: Terrorists’ sympathizers cheered when the Twin Towers collapsed.

Destructible: The Murrah building in Oklahoma City and Twin Towers on 9/11.

Occupied: Kill as many people as possible.

Near: Within reach of terrorist group and close to home base.

Easy: The Murrah building was an easy target to a car bomb placed within 8 feet of its perimeter.

Of the above characteristics, the proximity of the target to the base of operations (near) is probably the most important. Research has noted that terrorists will base their operations in regions and

countries where the local population is sympathetic to their cause. As noted earlier, the importance of proximity to target and base of operations has been known for some time in situational crime prevention research on the offender’s journey to the crime. The reason is that the closer to base of operations, the more detailed information can be collected concerning the accessibility of the target and the route to the intended attack. Street and traffic conditions, for example, may be significant in carrying out a suicide bombing. If a target is distant, it will be much more difficult to obtain the necessary operational information. This is why all Al-Qaeda attacks against the USA, except the attacks on the World Trade Center, have been conducted against US targets overseas (embassies and military bases). These targets were generally much closer to the bases of operations of Al-Qaeda. In the case of the Al-Qaeda attacks on the US World Trade Center in 1993 and 2001, it was necessary to set up satellite bases close to the targets. Subsequent successful and failed attempts in the UK and USA have revealed that attacks inspired by Al-Qaeda have been carried out by individuals who were born of immigrant parents within the home country, though in some cases received training from terrorist camps overseas. Thus, the traditional distinction between domestic and foreign terrorism no longer holds, but the necessity of proximity of the terrorist base of operations to the target remains supreme.

2. Weapons

As with targets, terrorists favor certain weapons over others. The majority of terrorist attacks are conducted using small arms and munitions. Attacks with weapons of so-called mass destruction (biological or nuclear) have been few, and where conducted, of limited effect. The characteristics of weapons that make them appealing to terrorists are MURDEROUS:

Multipurpose	A high-powered rifle has a specific use, while explosives have a much wider application
Undetectable	Semtex is small, lightweight, and largely undetectable, ideal for penetrating layers of security

(continued)

Removable	The weapons of terrorism must be portable, so they must be relatively light and reasonably small, unless transportation is available (see tools below)
Destructive	Explosive devices have a greater kill rate than guns targeted at specific individuals
Enjoyable	Terrorists are clearly attracted to their weapons. In fact, it is not just terrorists who enjoy weapons. Many ordinary people do also
Reliable	If terrorists have used a weapon, or one like it, many times before, they are likely to favor that weapon over another
Obtainable	How easy is it to get the weapon? Can it be bought or stolen easily? Or can it even be manufactured in-house?
Uncomplicated	All weapons require practice and training. A weapon that demands considerable skill, such as a free-flight armor-piercing missile launcher, will rarely be used or if used miss the target
Safe	The use of bombs as weapons is inherently more dangerous than the use of other weapons, especially when made in-house. But they remain the terrorist weapon of choice because they are so destructive

There will almost always be a trade-off between the target of choice and the weapon. If a target is particularly desirable, or difficult to reach, there may be a search for an appropriate weapon. For example, the first attack on the US World Trade Center in 1993 used a truck bomb parked under one of the towers. Though it did much damage, it failed to destroy the target which at the time was considered indestructible. The second attack used a new weapon, a passenger jet as missile, which was successful. The planning required to use this new weapon was very extensive and took years to accomplish. It also required Al-Qaeda to locate a satellite base of operations in the USA close to the targets (the relevant airports and target itself). Finally, it ignored the one attractive characteristic of weapons listed above: it was not “safe.” The users, as with suicide bombers, were willing to die in order to reach the target. However, we should note that in the case of “regular suicide bombers,” the explosives used are well

developed and “safe” for those managing the training of the suicide bombers. The explosive vest must be well constructed so that it is detonated only when the bomber reaches the target.

Weapons of mass destruction, including biological weapons, have been rarely used by terrorists. The reason is that such weapons generally do not fit the MURDEROUS requirements of weaponry to terrorists. However, using this approach, Hassan Naqvi reviewed the main biological and chemical weapons theoretically available to terrorists and concluded that ricin toxin would be the most attractive weapon because of its exceptionally MURDEROUS qualities:

<i>Multipurpose</i>	Ricin can poison individuals by means of ingestion, inhalation, and subcutaneous exposure
<i>Undetectable</i>	Ricin is tasteless and water soluble, making it extremely concealable. Laboratory setups designed to extract ricin toxin from castor beans range from very simple to as complex as a drug lab. Any of these labs can be set up in a garage
<i>Removable</i>	Ricin is extremely toxic to humans in low doses. Very small amounts of poorly purified ricin are enough to harm multiple individuals. A single terrorist will be able to carry vast quantities of ricin without help
<i>Destructive</i>	Ricin is a thousand times more deadly than cyanide. In addition, the dissemination of ricin by means of food, water, or ventilation vectors would cause considerable panic and economic turmoil
<i>Enjoyable</i>	The extraction of ricin from castor beans is a fairly simple science experiment, and its clandestine use may add to the enjoyment of its production. Admittedly, it may not be quite as spectacular in its effects as are violent explosions and guns
<i>Reliable</i>	Ricin is an extremely hardy protein. Because it is resistant to changes in temperature, acidity, and even ultraviolet light, it retains much of its poisonous properties in environments as extreme as the stomach
<i>Obtainable</i>	Castor beans are available for purchase on the Internet, and laboratory equipment can easily be stolen

(continued)

<i>Uncomplicated</i>	The extraction and dissemination of ricin is the most basic of all bioweapons. The instructions for its weaponization are easily found on the Internet
----------------------	--

<i>Safe</i>	Ricin, although extremely toxic, will not harm terrorists if fairly basic precautions are taken, such as the use of masks
-------------	---

3. Tools

Terrorist attacks cannot be accomplished without the tools of everyday life. Without such tools, it is much harder for terrorists to reach a target or use their weapons. For many of the commonest attacks, such as car or truck bombings, drive-by shootings, and targeted assassinations, terrorists are likely to need most of the following:

- Cell phones or other means of communication
- Cars or trucks to transport themselves and weapons
- Cash or (false) credit cards, bank accounts, or other means of transferring money
- Documents (false or stolen) – for example, drivers’ licenses, passports or visas, and vehicle registration documents
- Maps (and increasingly GPS), building plans, and addresses so that the target location can be pinpointed
- Television sets and monitors
- Video and still cameras for surveillance
- Internet access to collect information on street closures, traffic patterns, weather conditions, local news, and disseminate propaganda worldwide

Depending on the availability of such tools, the weaponry employed and targets reached may vary.

It may seem obvious that terrorists would need these everyday tools, but it is not obvious how they can use them without divulging at some point who they are and what they are up to. These tools are certainly widely available, but their visibility is also considerable. Using cash instead of a credit card to rent a car, for example, draws unwanted attention to the terrorist. Using a credit card exposes one’s identity – unless it is stolen. For these reasons, terrorists steal many of the everyday tools that are widely available,

which of course opens them up to risk of getting caught. This is why there is a considerable, and increasing, overlap between terrorism and “traditional crime” such as money laundering, drug trafficking, and even bank robbery – a favorite way of the IRA (Irish Republican Army) to raise money. Some terrorist groups have even created entire departments that specialized in forging documents. Others have made their own money.

Despite the generality of tools needed by terrorists, some attacks do, of course, require specific tools – for example, a belt or vest to carry the suicide bomber’s explosives – and it will always be important to understand what these tools might be and the supply chain that produces them.

4. Facilitating Conditions

Targets, weapons, and tools exist within physical, economic, and social environments. These environments serve to enhance their use; otherwise, they would be useless. At particular points in time or in particular regions or places, conditions may arise that facilitate terrorist ability to exploit these opportunities. These conditions make it ESEER for terrorists:

Easy: When local officials are susceptible to corruption

Safe: When ID requirements for monetary or retail transactions are inadequate

Excusable: When family members have been killed by local antiterrorist action

Enticing: When local culture/religion endorses heroic acts of violence

Rewarding: When financial support for new immigrants is available from local or foreign charities

One can readily see that the characteristics of ESEER are more or less those of failed or failing states, which is why they are the preferred abode of terrorist bases. ESEER is also exacerbated by the conditions of globalization that include the worldwide marketing of small arms whether legal or illegal, porous borders between countries making movement of terrorist operatives easier, the proliferation of nuclear technology and other toxic materials creating opportunities for terrorists to obtain or manufacture WMDs, and lax international banking practices that facilitate money laundering. The global reporting of

savage or violent terrorist attacks on mass media and the Internet has also probably facilitated terrorist recruitment into various terrorist organizations. It is to the organizational aspect of terrorism that we now turn.

The Situational Approach to Terrorist Groups

With some exceptions, terrorism is a group exercise, which brings with it both advantages and disadvantages to doing terrorism. The advantages include being able to (a) mount a more complex one-off mission (the Northern bank robbery by the IRA or the 9/11 attack) and (b) sustain a continued series of attacks over a long period of time (e.g., Hezbollah in the Palestinian territories). However, the majority of terrorist groups dissolve within 1 year or less. Whether they completely desist or morph into different terrorist groups is still an open question. But this important fact suggests that there are serious problems that affect the structure and dynamics of terrorist groups.

Terrorist Group Structure

There are a number of different types of organizational structure identified among terrorist groups. Much is made of their formal or informal structures, how they arose, and their similarity to other types of organizations such as business models, specialist groups for hostage taking, or cell formation to insulate members from the overall structure of the organization. The latter protects against penetration by external agents and minimizes internal group dissention and distrust.

It is also likely that developments in communication technology may mitigate the difficulties faced in organizational structure of criminal and terrorist groups. Mobile phones in particular, especially smart phones that provide cheap access to the Internet in third world countries, have made it much easier for groups to coalesce and to overcome some of the external and internal constraints of terrorist group operations. Indeed, as has occurred in the Arab spring uprisings of 2011, many of these have been fuelled by social

networking technologies. Demonstrations, sometimes turning into violent uprisings, have occurred without any formal organization at all. Social networking technologies may overcome the need to maintain a lasting formal group structure. Groups can come together – almost spontaneously for a specific event – then melt away until the next occasion. However, if routine, that is, repeated, attacks are to be sustained, obviously an enduring organization is essential, made clear as the uprisings in Libya, Egypt, and Syria unfold.

Terrorist and Criminal Groups

The situational approach views terrorist attacks that are carried out by groups or at least more than one person in much the same way as any other crime. That is, it follows the commission of the crime step-by-step; it does not presume in advance that there is a particular organized hierarchical structure. For example, the 2010 book *Situational Prevention of Organized Crime* shows that drug trafficking, contraband cigarettes, sex trafficking, mortgage fraud, and timber theft are commonly carried out by small networks of friends or family groups who live on different sides of a national border and take advantage of marketing and trafficking opportunities that are thus available.

Groups are viewed as part of the situational environment in which decisions must be made. The extent to which decision making is facilitated or constrained when it is done by groups or in group situations rather than by individuals is a difficult question to answer, though classic research on group decision making has suggested that in groups that are not well established, there is a tendency for polarization in decision making and a shifting of decision making to converge with the views of an authoritarian leader. So the group setting is yet another factor that limits or conditions terrorists' rational choice.

It is also important to recognize that in many, if not all instances, groups organized to commit traditional crimes ("organized crime") such as drug trafficking and extortion, corruption of government officials already existed before terrorist groups came along. Therefore, the existence of such groups is simply another opportunity for

terrorist groups to exploit, and it is logical that they would do so. There is mounting research that demonstrates the ways in which terrorist groups develop their organizational capacities as a result of the existence of organized crime and eventually collaborate, overlap with organized crime groups, or even morph in some instances into criminal groups that do organized crime. From the situational crime prevention point of view, however, it does not matter much whether what these groups accomplish is organized crime or terrorism. All we need to know is how they make their choices, what interventions will constrain them, and how working within a group organized in a particular way facilitates or constrains the choices made along the way to accomplishing the mission.

The Situational Dynamics of Terrorist Groups

There are many sources of conflict within terrorist groups. Newman and Clarke in *Outsmarting the Terrorists* have identified the internal dynamics of terrorist groups that may constrain them as:

- (a) The conflict between charismatic leadership and strict discipline. Charismatic leadership is probably the most significant attribute of terrorist groups and is necessary in order to transform terrorist groups into terrorist movements. It is needed to maintain faith in violence as a means to an end. But at the same time, strict discipline is needed to maintain operational security and efficiency. If charisma wanes (which we know always does), faith and thus discipline are undermined.
- (b) The conflict between morality and hypocrisy. The moral contradiction between killing people (by terrorists) as a protest against killing people (by whatever enemy) is inherently a shaky morality, needing to be constantly fed by self-serving propaganda.
- (c) Rules of entrance and exit. Some groups have no other rules than being a friend or family member. Others have strict rules for entry (applicants must prove themselves) and of exit – those who leave are obviously a security risk.
- (d) Operational security. Complex operations require well-trained individuals who know

each other well. But the more they know of each other, the more security will be jeopardized if one of them is captured.

The Situational Approach to Counterterrorism

The situational approach to counterterrorism is of course focused on prevention of terrorist attacks. As in crime prevention, this is its prime objective. Over the years, situational crime prevention has accumulated a range of techniques for preventing many different types of crime. These techniques, known simply as the “Twenty Five Techniques” of crime prevention, are grounded on five psychological principles that are directed at further limiting the rational choice of the offender. Applied to preventing terrorism these are:

1. Increasing the effort of mounting an attack or mission, such as protecting targets
2. Increasing the risk of carrying out an attack, such as surveillance and checkpoints
3. Reducing the rewards of attacks such as swift mitigation by first responders to reduce the injury and damage of attacks
4. Reducing provocations of terrorists such as avoiding controversial weapons
5. Removing excuses such as adopting clear rules of interrogation

If we superimpose these principles on the four pillars of terrorist opportunity, a broad array of techniques that suggest points of intervention is revealed (Fig. 3). Each of these five principles must be tailored to a specific type of terrorist attack after a step-by-step analysis has identified points of weakness.

Policy Implications

The four pillars of terrorist opportunity and the five modes of intervention that form the bases of the 25 techniques provide a framework for a systematic policy that would assess the risks of attack against various targets in various locations. After 9/11, “risk assessment” was promoted by the US Department of Homeland Security as a way of protecting against terrorist attacks. Unfortunately, because it lacked

a systematic structure of how to assess comparative risks, or even what might be at risk, the major portion of resources allocated to US states by the federal government was distributed according to political largesse rather than according to a systematic comparative assessment of what was at risk. This error was in part also produced by the widespread panic by both politicians and security professionals such as Richard Clarke the then White House counterterrorism czar and reflected in the cry, “we can’t protect everything.” But the situational approach demonstrates that not all targets are equally attractive to terrorists. We do not have to protect everything, at least not to the same degree. By developing a systematic procedure for identifying risk, adequate planning and countermeasures can be set up both to prevent attacks and to mitigate the fallout of attacks should they occur. The manual, *Policing Terrorism: An Executive’s Guide*, by professors Newman and Clarke and published by the Center for Problem-Oriented Policing, shows how this can be done by local police departments.

However, the complaint that “we can’t protect everything” reflects not only panic but also a basic misunderstanding of the situational approach to terrorism, which is that if we protect one target, the terrorist will simply shift to another target that is more easily accessible. This is called in the research literature “displacement” or the “substitution effect”. And it is the major criticism of the situational approach to crime and to terrorism.

Displacement

Echoed in *The Economist* on March 8, 2008, and in other trade media, the displacement criticism argues that terrorists are so committed to their violence; if they can’t attack one target, they will keep on trying until they reach another. However, we have already seen that terrorist groups, although they may be composed of committed individuals, are very likely to dissolve within a year, so the idea of commitment at least by a group to carry out an attack against anything anywhere remains highly doubtful. Furthermore, recent research in situational crime prevention by Professors Guerette and Bowers

	Targets	Tools	Weapons	Facilitating Conditions
Increase the Effort	<ul style="list-style-type: none"> Identify vulnerable targets Prioritize targets for protection Close streets, build walls and barriers Security training for VIPs Control dissemination of weapons technology 	<ul style="list-style-type: none"> Reduce supply of cash Design electronic products to prevent use as detonators, timers, High tech passports, visas, driving licenses etc. National ID cards 	<ul style="list-style-type: none"> Restrict weapons sales Hold contractors liable for stolen explosives Reduce explosives shelf-life Bomb recognition publicity "User unfriendly" weapons Restrict information on weapons use 	<ul style="list-style-type: none"> Tighten identity and credit authentication procedures Tighten border controls Destroy safe houses/training camps Disrupt recruitment (e.g. Madrasas)
Increase the Risks	<ul style="list-style-type: none"> Strengthen formal and informal surveillance through CCTV, citizen vigilance, hot lines 	<ul style="list-style-type: none"> Technology to identify and locate cars, trucks, cell phones Internet surveillance RFIDs for parts on vehicles and electronic products GIS chips in terrorist tools 	<ul style="list-style-type: none"> RFIDs/GIS chips to track weapons Screen incoming cargo for weapons Outlaw technology to circumvent screening 	<ul style="list-style-type: none"> "Know your customer" bank policy. Track all financial transactions Monitor foreign student activity Promote ties between local police and immigrant communities Conduct sting operations to foster climate of risk
Reduce the Rewards	<ul style="list-style-type: none"> Conceal or remove targets Bomb proof buildings/Kevlar curtains Design guidelines to minimize injury from explosions Swift cleanup of attack site 	<ul style="list-style-type: none"> Anticipate terrorist innovation in use of tools 	<ul style="list-style-type: none"> Avoid overreaction to attacks 	<ul style="list-style-type: none"> Use publicity to isolate terrorist group from community Use publicity to portray hypocrisy, cruelty of terrorist acts Anti-money laundering regulations
Reduce Provocations	<ul style="list-style-type: none"> Unobtrusive public buildings at home and abroad Unobtrusive dress, living accommodations for personnel abroad 	<ul style="list-style-type: none"> Use unobtrusive vehicles in foreign locations 	<ul style="list-style-type: none"> Clear and consistent rules of engagement Minimal weapons use at protest demonstrations 	<ul style="list-style-type: none"> Work closely with immigrant communities and host community abroad Clear rules for public demonstrations Avoid provocative announcements
Remove Excuses	<ul style="list-style-type: none"> Train personnel abroad in local customs, culture 		<ul style="list-style-type: none"> Avoid use of controversial weapons (eg phosphorous bombs) Don't overestimate utility of bioweapons 	<ul style="list-style-type: none"> Avoid maltreatment of prisoners Clear rules for interrogation Counter terrorist propaganda Win hearts and minds

Situational Approaches to Terrorism, Fig. 3 Situational prevention techniques applied to terrorism (Adapted from *Outsmarting the Terrorists*, pp. 193–194)

has shown clearly that displacement does *not* occur in some 70 % of cases studied where interventions were introduced for a wide range of crimes, and the same has been found for terrorism in respect to the effectiveness of airport security. Finally, should displacement be shown to occur, this is actually a demonstration that the intervention *had an effect* on the terrorists' behavior. It means that the terrorist in displacing to a less preferred target is taking additional risks, thereby increasing the chances of being caught or thwarted in the attempt. And having to attack a target of second choice, a successful attack will produce fewer rewards to the terrorist: it may be less vital and produce less injury and damage. This leads to the final important policy implication of the situational approach to terrorism.

Attack Mitigation

If it is acknowledged that displacement *may* occur (not *will* occur), then it is possible to develop a plan of security for protection of targets that rates them according to how attractive they are to terrorists (EVIL DONE) but also the expected loss should they be attacked. This may easily be incorporated into the standard disaster response planning that occurs in most communities. Then, if a target is attacked, the fallout of that attack can be mitigated by an effective disaster response, and thus the rewards of the attack to the terrorists reduced. The importance of this mitigation should not be underrated. Certainly Al-Qaeda terrorists have understood this from the beginning. It is why they commonly include in their attacks bombs designed to disrupt the mitigating actions of first responders. It is also why they try to commit a number of attacks simultaneously. This methodology has been copied by many other terrorist groups, especially the Taliban.

The situational approach to terrorism fits nicely into the established disaster response infrastructure already established. This not only involves training of first responders, but an extensive organizational infrastructure that ensures that confusion does not occur in dealing with an attack (which occurred in the 9/11 attack) but also

ensures that local communities plan systematically to protect themselves according to a rational assessment of risks. Much of this planning for protection involves regular situational crime prevention activities, well known in the field of problem-oriented policing especially partnering with businesses and local community organizations to solve crime and disorder problems.

Conclusion

In sum, the situational approach to terrorism bridges the well-known gap between the academic study of the causes of terrorism and the practical challenges of preventing and responding to terrorist attacks that are faced by policymakers, police, and security professionals. While waiting for the rest of social and political science to eradicate the *root causes* of terrorism, the situational approach provides a conceptual, practical, and evidence-based guide for eliminating the *immediate causes* of terrorist attacks.

Related Entries

- ▶ [Crime Prevention Through Environmental Design](#)
- ▶ [Environmental Criminology](#)
- ▶ [Identity Theft](#)
- ▶ [Money Laundering](#)
- ▶ [Organized Crime](#)
- ▶ [Problem-Oriented Policing](#)
- ▶ [Situational Crime Prevention](#)

Recommended Reading and References

- Bullock K, Clarke RV, Tilley N (eds) (2010) Situational prevention of organized crimes. Willan/Taylor and Francis, Abingdon
- Clarke RV, Mayhew PM (1988) Crime as opportunity: a note on domestic gas suicide in Britain and the Netherlands. *Br J Criminol* 29:35–46, 1989
- Clarke RV, Newman GR (2006) Outsmarting the terrorists. Praeger Security International, Westport/London
- Cornish DV, Clarke RV (eds) (1986) *The reasoning criminal*. Springer, New York

- Freilich JD, Newman GR (eds) (2009) Reducing terrorism through situational crime prevention, vol 25, Crime prevention studies. Criminal Justice Press, New York
- Guerette RT, Bowers KJ (2010) Assessing the extent of crime displacement and diffusion of benefits: a review of situational crime prevention evaluations. *Criminology* 48(1):1331–1368
- Hsu H (2011) Unstoppable? A closer look at terrorism displacement. Doctoral dissertation, University at Albany
- Jones SG, Libicki MC (2008) How terrorist groups end: lessons for countering Al Qa'ida. RAND, Santa Monica
- Makarenko T (2004) The crime-terror continuum: tracing the interplay between transnational organised crime and terrorism. *Global Crime* 6(1):129–145
- Naqvi H (2012) Situational crime prevention applied to ricin and bioterrorism. In: Freilich JD, Shoham SG (eds) Special issue on policing terrorism around the globe. *Israel studies in criminology*, vol 11 (in press)
- Newman GR, Clarke RV (2008) Policing terrorism: an executive's guide. USDOJ, Washington, DC
- Rapoport D (1988) Inside terrorist organizations. Columbia University Press, New York
- Rollins J, Wyler L (2010) International terrorism and transnational crime. Congressional Research Service, Washington, DC, pp 1–56
- Stohl M (2008) Networks, terrorists and criminals: the implications for community policing. *Crime Law Soc Change* 50(1–2):59

Situational Crime Prevention

Nick Tilley¹ and Aiden Sidebottom²

¹Jill Dando Institute of Security and Crime Science, University College London, London, UK

²Department of Security and Crime Science, University College London, London, UK

Overview

Situational crime prevention is concerned with reducing opportunities for crime. It comprises “opportunity-reducing measures that (1) are directed at highly specific forms of crime, (2) involve the management, design and manipulation of the immediate environment in as systematic and permanent a way as possible, (3) make crime more difficult and risky, or less rewarding and excusable as judged by a wide range of offenders (Clarke 1997: 4).”

Since its inception, situational crime prevention has undergone several refinements, incorporating developments in crime prevention research and practice. Opportunity reduction measures have taken numerous forms and have been found effective in reducing a wide range of crime types. It has also encountered staunch resistance, both on theoretical grounds and on the implications deemed to arise from the implementation of situational interventions.

This entry charts the development of situational crime prevention. It begins by describing its origins, method, and key principles. Next, it discusses evidence of its effectiveness. This is followed by a review of the major criticisms of situational crime prevention and associated rebuttals. It finishes by outlining future directions for the field.

The Origins of Situational Crime Prevention

Just as there was evolution before evolutionary theory, there was situational crime prevention before situational crime prevention theory. The core tenet of situational crime prevention is that unwanted actions by others can be averted by changing the immediate circumstances of their prospective behavior. It is not necessary (and indeed it may not be possible) to alter the dispositions inclining them to act in undesirable ways.

Hedgehogs practice situational prevention. In the interests of not being attacked, they have evolved sharp spikes that increase the risk and reduce the reward to those animals that might otherwise attack and kill them. Similarly, skunks have developed smells, butterflies the appearance of large eyes, and stick insects – well – the appearance of sticks, to make them less attractive to or less easily identified by potential predators. Situational prevention can even be seen in plants: think, for example, of stinging nettles, cacti, and bramble bushes.

Humans too invented situational measures long before late twentieth-century criminologists began to articulate the relevant theory. Most of us, for example, carry around with us milled

coins. These became commonplace in the seventeenth century as a way of preventing coin clipping, which was then widespread: the face value of the coins was supposed merely to reflect the value of the precious metals of which they were minted, and profit could be made by clipping off and keeping some of the metal but preserving the purchasing power of the coin. Unclipped coins thus went out of circulation as clipped coins had the same exchange value. In terms of Sir Thomas Gresham's famous aphorism: bad money drove out good.

The same fundamental attribution error may underlie the surprisingly long time it took for both evolutionary theory and situational crime prevention theory to surface. The "fundamental attribution error" refers to the human tendency to explain good and bad by the actions of agents with corresponding good or bad intentions (Ross and Nisbett 1991). Crime is bad so its causes must lie in bad intentions: in particular we should blame bad agents and the source of their wickedness. Indeed, echoes of this are still found in much criminology, which looks for the causes of crime mainly in criminal dispositions borne of bad genes, bad upbringing, or bad social conditions. In response to this, situational crime prevention theory presents revolutionary challenges to much taken-for-granted thinking.

The stimulus to the first formulations of situational crime prevention came from a crisis of confidence in conventional wisdom in the 1970s. Improvements in welfare during the post-war period and the simultaneous increases in crime cast doubt on any notion that there might be a simple causal relationship between deprivation and crime, which would be prevented by improvements in social conditions. Moreover, evaluations of efforts at offender rehabilitation, reflecting the dominant criminological opinion at the time, were not encouraging. The slogan "Nothing Works" came to be believed. Situational crime prevention emerged in the British Home Office and promised an alternative approach that focused on preventing crime events rather than addressing offenders' dispositions to commit crime. The approach was a pragmatic one. Situational crime prevention seemed to

work, and policies might usefully be directed at working out how and why situations could be modified to reduce opportunities for crime.

Theoretical and Methodological Underpinnings of Situational Crime Prevention

Crime as Opportunity was published by the British Home Office in 1976 (Mayhew et al. 1976). It was followed in 1980 by a paper in the *British Journal of Criminology* with the title "Situational Crime Prevention" (Clarke 1980). These two papers articulated a fresh agenda for understanding and preventing crime.

Mayhew et al. mainly focus on physical means of reducing crime opportunities. Their starting point, however, is "the extent to which deviance may be a temporal response to the provocations, attractions, and opportunities of the immediate situation" (1976: 1). They refer to Hartshorne and May's classic study, finding that honesty in children is not a stable attribute but instead that situational changes in opportunity can lead them sometimes to cheat, steal, and lie (Hartshorne and May 1928). Mayhew et al. refer also to research showing the importance of the immediate situation on the behavior of those placed in institutions, such as borstals. The language is of "stimulus conditions," including "opportunities for action," providing "inducements for criminality," which are "modified by the perceived risks involved in committing a criminal act and – in a complex, interrelated way – the individual's past experience of the stimulus conditions and the rewards and costs involved" (Mayhew et al. 1976: 2–3). Although opportunity had been referred to in passing in previous studies, Mayhew et al.'s argument was that it warranted closer, center-stage attention.

One of the most important grounds for arguing that situational factors are worthy of consideration in their own right was the discovery that changes in the toxicity of the domestic gas supply in the UK had a substantial effect on suicide rates. Mayhew et al. cite a study which found that the 45 % drop in Birmingham's suicide rate from 122

to 67 per million between 1963 and 1969 was almost entirely attributable to reduced toxicity of the gas supply in that city (there were 87 gas-related suicides in Birmingham in 1962 but only 12 in 1970). The drop in Birmingham's suicide rate was twice the national fall. A later study by Clarke and Mayhew (1988) corroborated these findings on a national scale. It might be expected that suicides are not committed lightly, but are provoked by deep feelings of despair and, therefore, can only be prevented by looking at the underlying psychological or social source of these sentiments. Finding that the gradual switch from toxic coal gas to nontoxic natural gas largely accounted for the overall drop in suicides challenged that assumption. Having an easily available method of committing suicide, in particular one that was painless and non-disfiguring, seemed to be associated with suicidal behavior. Removing that method had a substantial effect on suicide rates. Even in relation to this deeply caused behavior, there did not appear to be displacement to alternative methods (an issue we will return to shortly). There are many ways of taking one's own life. Changing the gas supply can have no influence on the cause of suicidal dispositions. It does not make suicide impossible. It nevertheless produced a substantial effect on total numbers of suicides. If deeply motivated behavior could be prevented by a simple change in the situation, how much more readily might criminal behavior, lacking such profound emotional sources, also be prevented?

Mayhew et al. see opportunities deriving from various sources: potential offenders' age, sex, and general lifestyle; victims' status; patterns of daily activity; properties of objects (such as their abundance); physical security; and levels of supervision and surveillance (Mayhew et al. 1976: 6–7). They also spell out variables that would be important in testing opportunity theories, for example, “the greater freedom of men to be away from their homes at night, the density of residential properties in rural as against urban areas, the security of these properties, or the extent to which routes of access to them are visible to neighbours and passers-by” (Mayhew et al. 1976: 7). The main emphasis, however, was

on the modifiable features of opportunities, notably “the physical security and surveillance aspects of opportunity” (Mayhew et al. 1976).

The Home Office paper goes on to discuss some examples of opportunity reduction as a means of preventing crime: the use of steering wheel locks to prevent car theft and the effects of design (front or back entry) and staffing (presence or absence of a conductor) on bus vandalism. Both studies find that situational variables were important in generating the crime patterns of interest. The introduction of steering wheel locks to all new cars in the UK, from 1971, reduced their theft but with some displacement to older cars. Areas of buses that were not easily supervised experienced far higher levels of vandalism than those open to surveillance.

Clarke (1980) adopts a more critical stance towards the then prevailing criminology than that found in the Home Office paper. He refers to the widespread “dispositional bias” which sees crime as mainly the preserve of a small minority inclined to criminality and which also continued to inform crime prevention policy and practice. He highlights the practical difficulties in changing criminal disposition or in altering the social conditions deemed to foster criminality, as well as uncertainties over the causal processes at work. He stresses, instead, the potential benefits of looking at crimes as results of consciously made, situated choices, even though those involved may not fully appreciate why they are making their decisions.

Clarke refers to three distinctive concerns of situational crime prevention theory: crime events (rather than criminality), separate categories of crime (rather than all crime), and current conditions for criminal acts (rather than dispositions). Two major types of preventive mechanism are discussed: (a) reduction of physical opportunities and (b) increase in chances of being caught, although (c) housing policies that improve chances of parental supervision of children are also mentioned. Clarke distinguishes between opportunistic offenders, whose crimes are less likely to be displaced by situational crime prevention measures and those who make their living from crime, for whom displacement is more

likely. A middle group, who supplement their income from crime, produces the main puzzle, where there are plentiful alternative targets. Clarke speculates that classes of target may be protected by situational measures, for example, when much stronger steel coin boxes replaced vulnerable aluminum ones in British telephone kiosks.

While Clarke's suggestions marked a wide departure from much criminological thinking at the time, noticeable similarities are observed between the situational approach and developments in several areas of psychology. For example, Clarke's concerns with criminology's "dispositional bias" spoke directly to the ongoing person-situation debate between personality psychologists and social psychologists. The then prevailing narrative was that human behavior is best explained by more or less stable traits that make up an individual's personality. This was challenged by the pioneering work of Mischel (1968), who claimed that personality traits were poor predictors of behavior and that greater attention be paid to situational influences. Later "situationalists" could point to a series of classic experiments, notably Zimbardo's Stanford Prison Experiment and Milgram's electric shock study that demonstrated the ease with which situational forces could be manipulated in ways that trumped dispositions, often to dramatic effect. There are also hints of behaviorism in Clarke's remark that "decisions are much influenced by past experience (as) people acquire a repertoire of different responses to meet particular situations and if the circumstances are right they are likely to repeat those responses that have previously been rewarding" (Clarke 1980: 138). Likewise, Clarke acknowledges explicitly that situational crime prevention "is, perhaps, closest to a social learning theory of behaviour" (p. 139). Clarke also notes, however, that situational prevention may owe "something to the sociological model of crime proposed by the 'new criminologists'" (Clarke 1980). Moreover, towards the end of his paper, he avers that opportunity reduction "is entirely compatible with a view of criminal behaviour as predominantly rational and autonomous" (Clarke 1980: 145).

The Home Office report of 1976 and Clarke's manifesto for situational crime prevention of 1980 laid the foundations for subsequent developments. These can be considered under the headings of methodology, models of the offender, and classification of preventive techniques.

Methodology

Action research has come to be the main method used in situational crime prevention. In accordance with this, Ekblom (1988) describes the "preventive process" as one involving (a) obtaining data on a crime problem, (b) analysis and interpretation of the data, (c) devising preventive strategies, (d) implementation, and (e) evaluation, with (f) feedback via continuous monitoring of crime.

Because situational measures are applied on the basis of the analysis of specific crime problems, evaluation methods that try to determine whether a specific type of measure, say lighting upgrades or CCTV, is effective are not normally deemed appropriate. Instead action-research-based case studies attuned to the problem context have become the preferred method.

Several tools and analytic techniques to inform situational crime prevention have been developed. In particular:

- Crime pattern analysis has advanced rapidly better to specify the problems to be addressed, as well as their spatial and temporal components.
- Repeat victimization patterns have been very widely found and victimization (by place, person, or target) identified as the best predictor for future risk and hence most deserving of preventive attention.
- The phenomenon of "near repeats" has also been identified: the elevated risks faced by those near to the victimized.
- The CRAVED mnemonic (concealable, removable, available, valuable, enjoyable, and disposable) has been devised to capture the attributes of products that are most likely to be stolen.
- "Crime scripts" are used to tease out the steps involved in committing more or less complex

crimes, which can then be drawn on to work out where the most promising points of preventive intervention can be found.

- The “problem analysis triangle” organizes the critical elements that give rise to crime events.

Models of the Offender: The Rational Choice Perspective

A rational choice model has been proposed for offenders (Cornish and Clarke 1986). The simplest way of understanding how situational measures work is to assume some level of rationality on the part of offenders. A change in the situation leads prospective offenders to reassess the expected costs and benefits from engaging in a particular criminal act. If the change in the situation suggests increased rewards or decreased costs, then more of those at the margins of participation in crime will be drawn in. Likewise, if the change suggests decreased rewards or increased costs, fewer will. Little of the formal apparatus of rational choice theory, as used by economists and game theorists, has been adopted. Rather, the assumption is simply that crime choices are open to influence by changing the expected balance of advantages and disadvantages. Indeed, even if there is a very strong disposition to commit a particular crime, the presence of a police officer at the scene, for example, may be sufficient to lead the individual to decide in the event not to proceed.

The rational choice model does not suggest that each individual carefully considers all the options that are available to determine which course of action will maximize utility. Lack of information, lack of ability, and the costs involved in arduous and prolonged decision-making make that untenable. Rather, potential offenders, like everyone else, operate with “bounded rationality.” This conceives of them as decision-makers who are responsive to changes in situations, but it does not embrace the unrealistic assumption that they are capable, conscious, careful, conscientious, and comprehensive calculators of costs and benefits to estimate expected utility in advance of each choice made.

Classification of Preventive Techniques

Clarke’s early references to reductions in physical opportunities and increases in the chances of being caught as two categories of situational measures have since been refined and expanded. These have been generalized, respectively, to “increasing effort” and “increasing risk” for the offender. “Reducing reward” was then added, followed by “reminding of rules” (now called “removing excuses”) and “reducing provocation” (Wortley 2001; Clarke and Homel 1997; Cornish and Clarke 2003). Taken together, these form a classification system of situational techniques. Each of the five major headings has a variety of subtypes, as shown in Fig. 1.

The classification of situational techniques serves four purposes. First, it provides a catalogue of the ways in which proximal conditions may affect criminal behavior. Second, it provides a repertoire of possible situational measures that can be countenanced when trying to determine what might be done in relation to a specific crime problem. Third, it provides a user-friendly teaching aide when introducing researchers and practitioners to situational crime prevention. And fourth, it highlights evidence gaps where case studies are required.

In some versions of the typology of techniques, the importance of perception is stressed in relation to risk, reward, and effort. That is, it is recognized that it is not only, or perhaps always most importantly, a question of *real* increases in effort or risk or *real* decreases in reward. For the person who might contemplate crime, the issue is that of the risk, effort, and reward that they *perceive* in relation to the criminal acts of interest. Similarly, provocation and rule recognition are about the ways in which situations are understood.

The Effectiveness of Situational Crime Prevention

A very large number of studies demonstrating the effectiveness of situational crime prevention have now been undertaken (e.g., see Clarke 1997 for a collection of successful case studies

Increase the effort	Increase the risks	Reduce the rewards	Reduce provocations	Remove the excuses
<p>1. <i>Target Harden</i></p> <ul style="list-style-type: none"> immobilizers in cars anti-robbery screens 	<p>6. <i>Extend guardianship</i></p> <ul style="list-style-type: none"> Routine precautions: go out in groups at night 'Cocoon' neighborhood watch 	<p>11. <i>Conceal targets</i></p> <ul style="list-style-type: none"> gender-neutral phone directories off-street parking 	<p>16. <i>Reduce frustrations and stress</i></p> <ul style="list-style-type: none"> efficient queuing soothing lighting/music 	<p>21. <i>Set rules</i></p> <ul style="list-style-type: none"> rental agreements hotel registration
<p>2. <i>Control access</i></p> <ul style="list-style-type: none"> alley-gating entry phones 	<p>7. <i>Assist natural surveillance</i></p> <ul style="list-style-type: none"> improved street lighting support whistleblowers 	<p>12. <i>Remove targets</i></p> <ul style="list-style-type: none"> removable car radios pre-paid public phone cards 	<p>17. <i>Avoid disputes</i></p> <ul style="list-style-type: none"> fixed cab fares reduce crowding in pubs 	<p>22. <i>Post instructions</i></p> <ul style="list-style-type: none"> 'No parking' 'Private property'
<p>3. <i>Screen exits</i></p> <ul style="list-style-type: none"> tickets needed for exit electronic merchandise tags 	<p>8. <i>Reduce anonymity</i></p> <ul style="list-style-type: none"> taxi driver ID's school uniforms 	<p>13. <i>Identify property</i></p> <ul style="list-style-type: none"> property marking vehicle licensing 	<p>18. <i>Reduce emotional arousal</i></p> <ul style="list-style-type: none"> controls on violent porn prohibit paedophiles working with children 	<p>23. <i>Alert conscience</i></p> <ul style="list-style-type: none"> roadside speed display signs 'Shoplifting is stealing'
<p>4. <i>Deflect offenders</i></p> <ul style="list-style-type: none"> street closures separate bathrooms for women 	<p>9. <i>Utilise place managers</i></p> <ul style="list-style-type: none"> CCTV for double-decker buses Two clerks in liquor stores 	<p>14. <i>Disrupt markets</i></p> <ul style="list-style-type: none"> monitor pawn brokers licensed street vendors 	<p>19. <i>Neutralise peer pressure</i></p> <ul style="list-style-type: none"> 'idiots drink and drive' 'it's OK to say NO' 	<p>24. <i>Assist compliance</i></p> <ul style="list-style-type: none"> litterbins public lavatories
<p>5. <i>Control tools/weapons</i></p> <ul style="list-style-type: none"> toughened beer glasses disabling stolen cell phones 	<p>10. <i>Strengthen formal surveillance</i></p> <ul style="list-style-type: none"> security guards CCTV in town centres 	<p>15. <i>Deny benefits</i></p> <ul style="list-style-type: none"> ink merchandise tags graffiti cleaning 	<p>20. <i>Discourage imitation</i></p> <ul style="list-style-type: none"> rapid vandalism repair V-chips in TV's 	<p>25. <i>Control drugs and alcohol</i></p> <ul style="list-style-type: none"> breathalysers in pubs alcohol-free events

Situational Crime Prevention, Fig. 1 Typology of situational crime prevention techniques. **Source:** Tilley, N. (2009). *Crime Prevention*. Willan, Cullompton, Devon

and for a fuller bibliography, with over 200 examples, go to <http://www.popcenter.org/library/scp/pdf/bibliography.pdf>. These cover a range of crime types and different settings. For example, in one study on residential burglary

in England, lockable gates were installed to alleyways that ran behind terraced properties. The gates restricted access only to those residents in the affected properties, thereby reducing the opportunities for burglars to enter such properties

from the rear, as was evidently the norm. The scheme was found to reduce residential burglary by over a third relative to the control area. Moreover, the scheme was cost-effective, recouping £1.86 for every one pound spent (Bowers et al. 2004). In another study, La Vigne (1994) describes how the implementation of a computerized telephone system in the New York City jail system led to considerable reductions both in illicit phone use by inmates and phone-related violent disputes. In other examples, the effect was immediate, as with the near disappearance of bus robberies in the USA in the 1970s following the implementation of cash drop safes and the shift to an exact-fair payment system.

Such studies do not find some specific measure or type of measure that “works” invariably – specific crime prevention panaceas are not expected. Instead, what they show is that strategies making alterations in the balance of rewards, efforts, and risks or removing provocations or adding rule reminders at the point of crime commission can be devised, which speak to specific crime problems.

The true extent of the effectiveness of situational crime prevention is likely to be higher than that gleaned from the published literature. Retailers and private businesses routinely employ situational measures to protect their assets and prevent loss or theft. Fitting merchandise with electronic tags or placing items in protective displays is commonplace. Such efforts are, however, rarely formally evaluated and hence are typically absent from the scientific literature.

Criticisms of Situational Crime Prevention

Both the theory and practice of situational crime prevention have been subject to an array of criticisms. The most common are itemized below, followed by the associated rebuttals.

Theoretical Criticisms

Situational Crime Prevention Only Displaces Crime

Arguably, the most enduring criticism of situational crime prevention concerns displacement.

It is argued that situational crime prevention simply shifts crime around because it fails to address the underlying “root causes” of crime. Displacement can take several forms: by place, time, method, crime type, or offender. There are enormous measurement problems that mean that it is, in effect, impossible to rule out all forms of displacement. The best that is possible is to assess whether there is displacement where, when, and in relation to which offenses it would be most expected. On that basis, it can be concluded with some confidence that displacement rarely matches the crimes prevented by the application of situational crime prevention (Bowers and Guerette 2009). This is because (a) offenders like to offend in areas with which they are familiar and (b) crime opportunities are not uniformly available. Contrariwise, there is growing evidence that “diffusions of benefits” often occur whereby preventive effects extend beyond the operational range of the measures applied. The possible types of diffusion of benefits match those of displacement. They have been found quite widely, for example, by time, where “anticipatory benefits” include reductions in crime before the preventive measures become operative (Smith et al. 2002), and place, where the geographical range of crime reductions extend beyond the area covered by the measures (Clarke and Weisburd 1994). Proposed explanations for these patterns are that offenders tend to be unaware of the exact boundaries (temporal and spatial) of prevention measures.

Situational Crime Prevention Is Superficial and Simplistic

Crime is complex with many causes. Its explanation calls for attention to a wide range of variables relating to genetics, childhood experience, social structure, and the workings of the criminal justice system. Even if situational variables have any role to play, it is a small one that needs to be considered in the context of a wide range of other causal influences. In placing less emphasis on them, the situational approach is criticized for speaking only to the descriptive components of crime – *where* and *when* crime happens – and not to *why* crime occurs. Situational crime prevention

is therefore alleged to be overly simple: lacking the theoretical depth of mainstream criminological theories and merely extending common sense.

The criticism is a weak one. Situational crime prevention draws heavily on three well-established crime event theories: routine activity, rational choice, and crime pattern. It is also underpinned by the widely accepted principle that situations yield causative influences on human behavior, in combination with predisposition, nor can it accurately be considered simple. While the superficial logic is often consistent with day-to-day observations, why the interaction between certain individuals and certain situations under contrasting conditions does or does not generate crime is in no way less complex than the question of why certain individuals do or do not enter a life of crime.

Situational Crime Prevention Is Applicable Only to Instrumental Offenses

Although situational crime prevention may be relevant to instrumental crime where some calculation of returns in relation to effort and risk might reasonably be expected and where, therefore, situational changes that alter expected utility might change an offender's decisions, for other types of crime, it is much less relevant. Many crimes, especially those involving violence, are more expressive than instrumental or are undertaken under the influence of drugs or alcohol. In these cases, those involved are unlikely to notice or adjust their behavior in response to more or less subtle situational factors.

There are clearly more examples of situational crime prevention being applied to instrumental crimes than there are expressive crimes. But generalizing this pattern to reduce situational crime prevention to an instrumental-only form of crime control is misleading. That many early situational measures relate to acquisitive crimes reflects demand not discordance. Situational crime prevention was developed by researchers based in the British Home Office, which, at the time, was primarily concerned with the high levels of residential burglary and car crime experienced in England and Wales. More generally, there is a situational component to all behaviors. The

association of changes in the supply of gas in the UK with dramatic changes in suicide patterns was described earlier. There is also strong evidence to show that situational changes can lead to dramatic reductions in the probability of and harm caused by violence. Bars are a prime example. The shift from glass pint pots to shatterproof receptacles and changes to the internal configuration of bars and nightclubs have all been found to produce positive reductions in alcohol-related assaults in such settings (see Graham and Homel 2008).

Ethical and Practical Criticisms

Situational Crime Prevention Blames Victims

Offenders commit crime and so should be held fully responsible for their criminal behavior. Situational crime prevention suggests that crime problems derive from the opportunities created by those who have not committed any criminal act. Moreover, it suggests that they should accept some responsibility (and in certain cases bear some of the costs) of reducing opportunities. For many critics this is an affront, leading to accusations that situational prevention "blames the victim." Offenders choose to commit their crimes. Fault lies entirely with them. Moreover, it is the duty of the state in general, and the criminal justice system in particular, to control their behavior.

One suspects that the understandable emotions associated with victim blaming, including most obviously the unwarranted notion that some rape victims may be blamed for wearing provocative outfits, have placed greater prominence on this criticism than is deserved. Situational crime prevention recognizes that targets (broadly defined) can play a causative role in crime events. This is true of people, places, and products. Evidence suggests that certain measures are associated with reduced vulnerability to victimization (as well as its opposite). Most individuals welcome such advice and use it to take recommended precautions against crime. Most individuals would also accept that this is a reasonable expectation and that government agencies are largely impotent to protect them from crime in every situation all of the time. Moreover, there are examples where victims are open to blame and where they

may be held culpable for the criminogenic effect of their actions (or inactions). For example, a shop that repeatedly experiences theft because of the way it displays goods, but passes the costs of the crimes to the state, may rightly be blamed for failing to take adequate precautions. Likewise, bar managers who blatantly supply alcohol to already intoxicated patrons may be held responsible for the violent crimes they produce.

Situational Crime Prevention "Criminalizes" Policy, Planning, and Design

The notion that the physical and social architecture of everyday life shapes crime opportunities and that it should therefore be adjusted to reduce crime, risks "criminalizing" areas of policy and practice (shaping them by a situational crime prevention agenda) to the detriment of other ends they may have. Pursuing a situational crime prevention agenda, it is argued, undermines social welfare policy, and diverts resources away from welfare provision and social crime prevention. This also applies to schools, hospitals, retailers, product designers, planners, and architects. In particular it may suggest acceptance, in the name of crime prevention, of an ugly "fortress society."

Knepper (2009) argues that situational crime prevention contributes more to social policy than the literature suggests. Many situational crime prevention case studies take place in settings that support welfare, namely, housing, hospitals, public transport, and schools. This is most clearly seen in the work of Poyner and Webb on how the design and configuration of social housing influences rates of crime (Poyner and Webb 1991). Effective situational crime prevention can thus increase the likelihood of meeting social welfare objectives.

With respect to the criminalization of products and places, it is trite but worth emphasizing that situational crime prevention amounts to more than simply target hardening. Situational interventions exhibit huge diversity, from high-tech baggage detectors to low-tech bicycle parking stands designed to facilitate improvements in locking practice. Nor are such measures ugly. To situational advocates, particularly those in

the tradition of design against crime, manipulating the immediate environment in ways that reduce opportunities for crime without compromising, say, aesthetics, costs, or function holds sovereign (see Ekblom 2008). Finally, while some might consider gated communities as emblematic of a "fortress society," there are many examples of situational interventions altering the physical environment in ways that promote community togetherness by facilitating neighborly interactions. Lighting upgrades and subsequent increases in pedestrian street usage after dark is a case in point (Painter 1996).

Future Directions for Situational Crime Prevention

There is scope for some theoretical development in situational crime prevention. It is not clear that the rational choice model of the offender is necessary for it. The rational choice model emerged only after the core ideas about situational crime prevention had been formulated, at which point, as indicated earlier, various possible models of the offender were mooted. Moreover, the incorporation of provocation reduction and rule reminders suggests that agents can be responsive to situational changes that operate at the level of feelings and morals as well as at the level of utility calculation.

Displacement is a further area deserving of research. Evidence on the presence of displacement is lopsided. The assessment of spatial displacement and to a lesser extent temporal displacement predominates, while other types of displacement have been little considered, in part because of the challenging measurement problems. The empirical research that is available is also limited in what it can capture, even when technically highly sophisticated. It has long been recognized that displacement is likely to be contingent on factors such as alternative available targets and types of offender motivation. There is scope for further theoretical development to devise propositions that predict what kinds of displacement can be expected in what conditions, again delivering on the suggestions of early

papers on situational prevention. The same is true for diffusion of benefits. That theoretical work then needs to be followed by empirical tests.

There is also scope to advance the application of situational crime prevention. While the range of crime types that have been targeted by research and practice has expanded enormously over the past 30 years, there remain important gaps. Studies into “unconventional” crimes, for example, organized crimes of various types (Bullock et al. 2010), terrorism (Clarke and Newman 2006), wildlife crime (Pires and Clarke 2012), and cyber crime (Newman and Clarke 2003), are available but at present are largely programmatic. Case studies evaluating measures designed to activate provocation-reduction or excuse-removal mechanisms are limited (Guerette 2009). Moreover, situational interventions have rarely been formally evaluated when applied in industrial, office, and agricultural settings (Guerette 2009). More broadly, situational crime prevention has yet to be systematically trialed in developing resource-limited countries, although interest appears to be growing (Pires and Clarke 2012). A key question relates to the challenges of carrying out the preventive process in settings where data are often absent.

The final point concerns exposure and dissemination. Despite being widely advocated and implemented in both North America and much of Europe, it is fair to say that situational crime prevention is yet to capture the imagination of the public and wield widespread influence on policymakers. There is a banality to situational crime prevention that belies its sophistication. This often means it is not credited for providing valid explanations for shifting crime patterns where traditional criminology fails. Nor is it always recognized for its role in producing major crime reductions. Take the much-storied crime drop experienced in several industrialized nations from the early to mid-1990s. Numerous hypotheses have been put forward for such declines, many receiving widespread media coverage. Yet arguably the most convincing explanation of the observed trends relates to the prevalence and quality of in-car (situational) security measures (Farrell et al. 2011). Failure

to bring the successes of situational crime prevention to the attention of the public may stymie its progression and confine it to the margins of crime prevention policy and practice.

Related Entries

- ▶ [Crime Mapping](#)
- ▶ [Crime Prevention Through Environmental Design](#)
- ▶ [Crime Science](#)
- ▶ [Effectiveness of Situational Crime Prevention](#)
- ▶ [Multiple Victims and Super Targets](#)
- ▶ [Problem-Oriented Policing](#)
- ▶ [Rational Choice Theory](#)
- ▶ [Routine Activities Approach](#)
- ▶ [Theories for Situational and Environmental Crime Prevention](#)

Recommended Reading and References

- Bowers KJ, Johnson SD, Hirschfield A (2004) Closing off opportunities for crime: an evaluation of alley-gating. *Eur J Crim Policy Res* 10(4):285–308
- Bullock K, Clarke R, Tilley N (eds) (2010) *Situational prevention of organised crimes*. Willan, Cullompton
- Clarke RV (1980) Situational crime prevention: theory and practice. *Br J Criminol* 20:136–147
- Clarke RV (ed) (1997) *Situational crime prevention: successful case studies*, 2nd edn. Harrow and Heston, Guilderland
- Clarke RV, Homel R (1997) A revised classification of situational crime prevention techniques. In: Lab SP (ed) *Crime prevention at a crossroads*. Academy of Criminal Justice Sciences and Anderson, Highland Heights/Cincinnati
- Clarke RV, Mayhew P (1988) The British gas suicide story and its criminological implications. In: Tonry M, Morris N (eds) *Crime and justice*, vol 10. University of Chicago Press, Chicago, pp 79–116
- Clarke RV, Newman GR (2006) *Outsmarting the terrorists*. Praeger Security International, Westport
- Clarke RV, Weisburd D (1994) Diffusion of crime control benefits: observations on the reverse of displacement. In: Clarke RV (ed) *Crime prevention studies*, vol 2. Criminal Justice Press, Monsey, pp 165–182
- Cornish D, Clarke RV (eds) (1986) *The reasoning criminal*. Springer, New York
- Cornish DB, Clarke RV (1987) Understanding crime displacement: an application of rational choice theory. *Criminology* 25:933–947
- Cornish D, Clarke RV (2003) Opportunities, precipitators and criminal decisions: a reply to Wortley’s critique of

- situational crime prevention. In: Smith MJ, Cornish DB (eds) *Crime prevention studies: theory for practice in situational crime prevention*, vol 16. Criminal Justice Press, Monsey
- Eklblom P (1988) *Getting the best out of crime analysis*. Home Office, London. Crime prevention unit series, Paper No. 10
- Eklblom P (2008) *Designing products against crime*. In: Wortley R, Mazerolle L (eds) *Environmental criminology and crime analysis*. Willan, Cullompton, pp 195–220
- Farrell G, Tseloni A, Mailley J, Tilley N (2011) The crime drop and the security hypothesis. *J Res Crim Delinq* 48:147–175
- Graham K, Homel R (2008) *Raising the bar: preventing aggression in and around bars, pubs and clubs*. Willan, Cullompton
- Guerette RT (2009) The pull, push and expansion of situational crime prevention evaluation: an appraisal of thirty-seven years of research. In: Tilley N, Knutsson J (eds) *Evaluating crime reduction initiatives*, vol 24, *Crime prevention studies*. Criminal Justice Press, Monsey, pp 29–58
- Guerette RT, Bowers K (2009) Assessing the extent of crime displacement and diffusion of benefit: a systematic review of situational crime prevention evaluations. *Criminology* 47(4):1331–1368
- Hartshorne H, May MA (1928) *Studies in deceit*. Macmillan, New York
- Knepper P (2009) How situational crime prevention contributes to social welfare. *Liverp Law Rev* 30:57–75
- La Vigne NG (1994) Rational choice and inmate disputes over phone use on Rikers Island. In: Clarke RV (ed) *Crime prevention studies*, vol 3. Criminal Justice Press, Monsey, pp 109–125
- Mayhew PM, Clarke RV, Sturman A, Hough JM (1976) *Crime as opportunity*. Her Majesty's Stationary Office, London. Home Office Research Study, No. 34
- Mischel W (1968) *Personality and assessment*. Wiley, New York
- Newman GR, Clarke RV (2003) *Superhighway robbery: preventing e-commerce crime*. Willan, London
- Painter K (1996) The influence of street lighting improvements on crime, fear and pedestrian street use, after dark. *Landsc Urban Plan* 35(2–3):193–201
- Pires S, Clarke RV (2012) Are parrots CRAVED? An analysis of parrot poaching in Mexico. *J Res Crime Delinq* 49:122–146
- Poyner B, Webb B (1991) *Crime free housing*. Butterworth Architect, Oxford, UK
- Ross L, Nisbett RE (1991) *The person and the situation: perspectives of social psychology*. McGraw-Hill, New York
- Smith MJ, Clarke RV, Pease K (2002) Anticipatory benefit in crime prevention. In: Tilley N (ed) *Analysis for crime prevention*, vol 13, *Crime prevention studies*. Criminal Justice Press, Monsey
- Wortley R (2001) A classification of techniques for controlling situational precipitators of crime. *Secur J* 14(4):63–82

Situational Crime Prevention and the Wild West

Shawna Cleary¹, Rashi K. Shukla¹ and Marcus Felson²

¹School of Criminal Justice, University of Central Oklahoma, Edmond, OK, USA

²School of Criminal Justice, Texas State University-San Marcos, San Marcos, TX, USA

Synonyms

[Crime prevention](#); [Natural crime prevention](#)

Overview

Criminology now includes an increasing body of research and theory that subsumes crime prevention within a larger topic of protection and survival in nature. That body of work not only considers crime prevention as representative of a larger survival process but also takes into account human-animal interactions. Such interactions include human abuses of animals and animal populations and nature in general – including abuses that are against the law. But human-animal interactions also include humans protecting against animal incursions into their herds and home territories.

Background and Review of Literature

The situational prevention literature has recognized historical examples of human self-protection as a background topic but has increasingly brought the interactions with nature to the foreground. After reviewing a bit of that literature, the current entry shifts to a pictorial review of situational crime prevention in America's "Wild West." That term often applies to a historical period that has disappeared, but even today many Western states include sections with ranches, horses, cattle, and exposure to predatory incursions. Political struggles are found

between metropolitan environmentalists who wish to reintroduce wolves to rural areas and ranchers who fear the loss of cattle and sheep to those same wolves. Migrants from city to suburban fringe and rural areas seek greater security than the areas from which they came; but they introduce new risks and fears to the areas that they enter. Thus, one person's protection becomes another person's threat. Survival from outside threats is a perennial problem in human history as in natural history.

The situational prevention of attacks also was found in the Old West. Consider this problem: ingots of silver bullion were carried by stagecoaches that were often victims of robbery by gangs of armed men on horseback. The problem was completely solved by instead producing 300 lb ingots that could not be carried on a single horse (Lingenfelter 1986). That was the end of stagecoach robbery. This example has been noted by several criminologists (Tilley and Laycock 2002; Clarke 1999; Farrell and Pease 2006).

Another example from the Old West came from pre-cowboy times. The ancient cliff-dwelling Indians in what is now Arizona and New Mexico entered their homes by climbing up ladders. At night they pulled up their ladders to prevent intrusions. This illustrates that situational prevention is far from new in behavioral terms. Moreover, Oscar Newman (1972) called our attention to territoriality and defensible space. Although situational crime prevention now includes many more ideas than that, it is still recognized that Oscar Newman placed defense of oneself, one's offspring, and one's dwelling area within a larger natural history perspective. That perspective was elaborated greatly in Felson's (2006) *Crime and Nature*, which included symbiotic and competitive ideas about offenders and the larger society.

Yet the breadth of "environmental criminology" is widened still further by new criminological papers on human-animal interactions. Recent research on parrot poaching strongly supports the idea that such crimes are highly structured in time and space and therefore subject to situational crime prevention (Pires 2012; Pires and Clarke 2011, 2012).

Additional important research considers human attacks on elephants (Lemieux and Clarke 2009), and a more general review of wildlife crime is found in Pires and Moreto (2011), explaining why such offenses are subject to situational prevention. It is also clear from Eliason (2012) that wildlife crime is theoretically understood using the routine activity approach, a thread running through each of the articles reviewed in this paragraph.

Clarke (1999) offers us a framework for evaluating security by offering 25 techniques for situational crime prevention, along with many examples. His framework is widely employed to analyze crime situations and recommend solutions. These techniques are organized in five columns based on how they contribute to crime reduction. Included are (1) increasing the effort to carry out the crime; (2) increasing the risks to the offender; (3) reducing the rewards that the crime gives to the offender; (4) reducing provocations, especially for dispute-related crimes; and (5) removing excuses for carrying out the crime.

The current review considers how ranchers and others in today's Wild West use situational prevention techniques to protect against intrusions and attacks. The more important goal is to add examples and hence to widen the repertoire of ideas for studying the prevention-protection process and for understanding routine activities and situational prevention of crime. This entry shall demonstrate how the techniques of situational crime prevention have been and are currently being used in the American West, especially for nonhuman predators. Despite the urban image of rural life as idyllic, the reality is far from that. Ranchers, farmers, and other property owners often have a great deal of land and animals that must be controlled and protected from people, animals, and insects. A protective mentality is very important for their economic survival and is taken very seriously.

A Note About the Pictures

The pictures used within this entry were taken by the primary author. Many of the pictures were taken in rural Oklahoma; a few were taken at the

Situational Crime Prevention and the Wild West, Fig. 1 Old style bank teller window with bars



National Cowboy and Western Heritage Museum in Oklahoma City and the Woolaroc Museum in Bartlesville, Oklahoma.

Crime Prevention in the West

We begin (Fig. 1) with a bank teller window, more heavily barred than most modern banks in the United States. Although the design of the bars is attractive, their main purpose is to keep offenders out of the money area – an example of target hardening along with a kind of perimeter control. That did not prevent very aggressive robbery by multiple armed men, but it did narrow the span of vulnerability. Figure 2 shows a safe from the Cattlemen’s State Bank, designed to be too heavy for easy removal and too hard to break, given its heavy lock. Although we take these safe designs for granted today, they originated in an earlier era in response to the mobility offered by horses and a desire to respond.

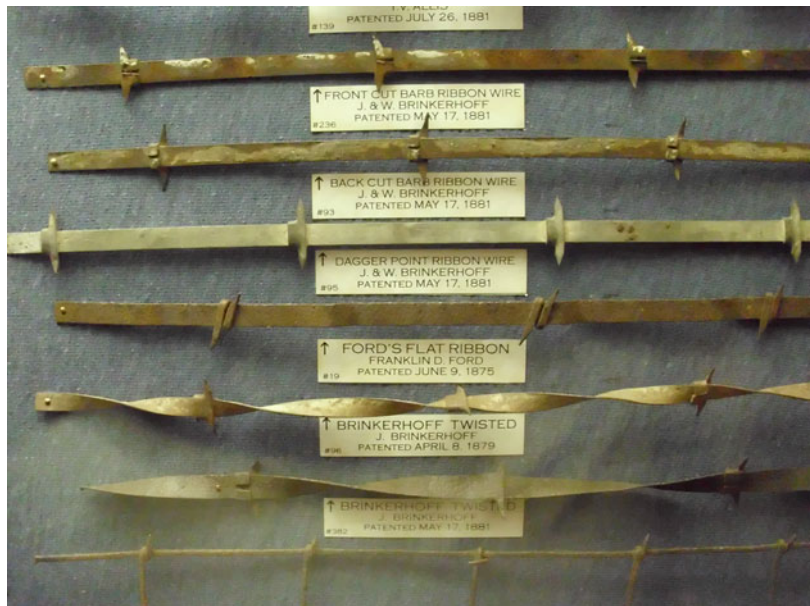
One of the most important inventions in the American West was barbed wire. Without it, one can argue that the West would never have developed, since mass grazing of animals would have been impractical if they could easily wander off. The Biblical image of a shepherd watching

a small herd of sheep or goats contrasts with the American process of maintaining very large herds with minimal supervision. Barbed wire fencing was developed in the late 1800s as a method of frugally enclosing animals and land while also preventing human theft. Barbed wire fence design styles range from legal and patented to illegal versions of patented designs to custom designs, resulting in thousands of versions (Liu 2009). In fact, the Cowboy and Western Heritage Museum in Oklahoma City displays 1,300 different types of barbed wire fencing. Figure 3 contains a few examples of different barbed wire styles from the late 1800s, each with its own patent. None of these styles is currently in use today, but the basic idea continues to the present in other forms. Figure 4 shows a rudimentary rural fence, while Fig. 5 shows a more professional version – each of which serves to contain cattle and signal other people; modern fencing in the West (not depicted) consists of wire woven in a small v-pattern that encloses cattle and horses without catching their hooves, is very difficult to climb, and keeps out large predators unless they can jump. Each of these fencing types applied defensible space long before Newman articulated it.

Situational Crime Prevention and the Wild West, Fig. 2 Cattlemen's State Bank safety vault



Situational Crime Prevention and the Wild West, Fig. 3 Examples of different styles of barbed wire

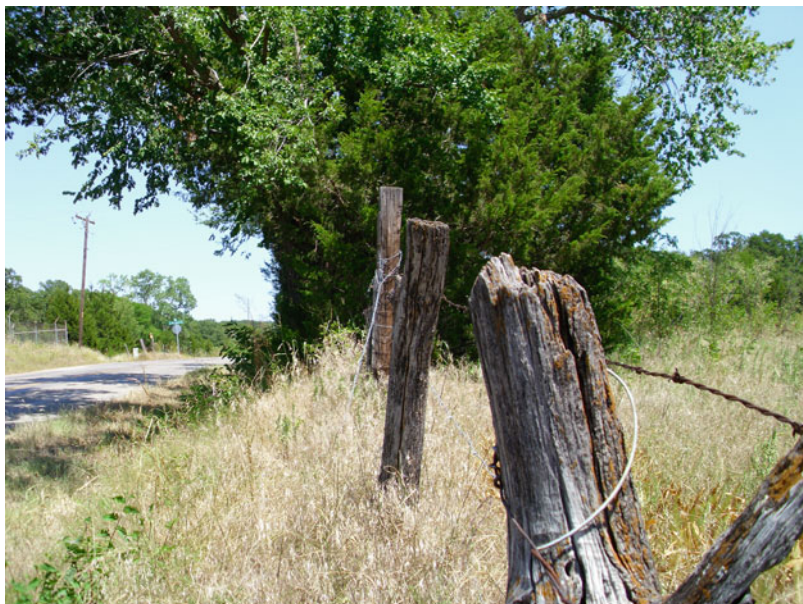


Not so well known outside the West is the old technique of hanging dead coyotes on a fence as a signal to live coyotes that their lives are endangered if they cross (Fig. 6).

It is important to consider in reading the current entry that predatory animals, from insects to

coyotes, do more than just nibble. Attacks sometimes decimate crops or herds and other times remove the profit margin. Early ranchers worked on a much smaller scale, but today's West includes not only large-scale operations but also many small-scale ranchers who work part-time in

Situational Crime Prevention and the Wild West, Fig. 4 Rudimentary barbed wire fence in rural Oklahoma



Situational Crime Prevention and the Wild West, Fig. 5 Well-constructed barb wire fence in rural Oklahoma



the city and are subject to considerable loss of income and way of life. Thus, protection of livelihood is a serious issue for them.

The use of predator against predator is an important part of their defense strategy. Owls are natural predators for farm pests, including

small birds and rodents. Many rural people in the West place statues of owls in their barns to act as artificial place managers (Fig. 7). Plastic snakes and hawk statues are sometimes used for the same purpose. Scarecrows are designed to look like a farmer and are sometimes used as

Situational Crime Prevention and the Wild West, Fig. 6 Dead coyotes hanging on fence as warning



Situational Crime Prevention and the Wild West, Fig. 7 Plastic owl serving as place manager in barn

S

artificial place managers to scare off birds and protect crops; many of the domestic pets in metropolitan areas become very important economically in the American West. The role of dogs was already mentioned above. Not so well known to urban people is the role of barn cats – commonly used as place managers at farms, ranches, and

stables. These cats live in and around barns for the primary purpose of hunting and killing mice, rats, birds, moles, and small snakes.

Animal containment is an important rural issue for many reasons, especially for large animals. Ranchers and farmers do not want to lose their cattle or find them injured and have to

Situational Crime Prevention and the Wild West, Fig. 8 Cattle guard



Situational Crime Prevention and the Wild West, Fig. 9 Guard donkey with longhorn cattle in rural Oklahoma



ethanize them, nor do drivers want to run into stray cattle. [Figure 8](#) illustrates a cattle guard, designed to contain large animals in pastures while allowing gates between them to remain open. Cattle guards are formed by placing steel bars close together horizontally across a shallow ditch under a gate. The bars discourage hoofed animals from exiting through the gate

as their hoofs slip on the bars and get caught between them. Cows and horses will generally avoid cattle guards, although they are not foolproof. Some horses may jump over them, and both cattle and horses have been known to catch their hoofs between the bars, breaking one or more legs, leading to certain euthanization.

Situational Crime Prevention and the Wild West, Fig. 10 Longhorn cow wearing branded saddle



Situational Crime Prevention and the Wild West, Fig. 11 Stall divider



One of the most interesting prevention techniques is to use guard donkeys near cattle to watch fences and prevent intrusions (Fig. 9). Donkeys are good place managers because they have a dangerous kick and tend to attack animal predators. Similarly, llamas or large dogs are sometimes used for the same purpose. This

gives us a new way to look at “formal surveillance,” with animals functioning as security guards, requiring relatively inexpensive maintenance, and demanding no pensions or benefits.

Figure 10 illustrates branding – burning unique marks into the flesh of cattle, horses, or saddles to identify them with a particular ranch.

This fits very closely with the situational prevention method of identifying property to discourage burglary (Laycock 1991). Sometimes a brand would be altered by the thief in order to allow sale. A more modern technique is less easily thwarted: Freeze branding leaves a white brand mark on an animal's hide and is more effective than burning. Thus, we see that Ekblom's (2001) "arms race" between offenders and crime prevention techniques applies as much in the West as in metropolitan areas around the world. Modern animals and property are also identified with brass nameplates (not illustrated) that offer some protection if offenders do not have time to remove them.

Because large-scale horse operations have been commonplace in the West, managing animal interaction is important. Because horses are herd animals, it is important for them to see and communicate with each other, but physical interaction can be problematical if they begin to fight. Horses within a single farm or ranch are usually separated from one another with a variety of walkways and stall dividers (see Fig. 11) allowing them to see, hear, and smell one another, without biting or kicking.

Conclusions

Numerous types of situational crime prevention apply to the American Wild West and today's remaining Western frontier. Target hardening includes bank teller windows and heavy bank safes. Barbed wire fencing and cattle guards serve to control access. Guard donkeys and barn cats serve as place managers. Branding cattle identifies property. Stall dividers and horse barn designs serve to avoid equine disputes. Dead coyotes on fences and owl statues in barns give instructions to intruders.

We can readily see that situational prevention depends on routine security built into the design and management of daily life. This applies in the metropolis and Wild West alike, but not always exactly alike. Much more needs to be known about human-animal interaction and predation and

human protection of their own animals from others. Crime is part of a much larger ecological process which we are only beginning to understand.

Related Entries

- ▶ [Crime Prevention Through Environmental Design](#)
- ▶ [Designing Products Against Crime](#)
- ▶ [Informal Guardianship](#)
- ▶ [Routine Activities Approach](#)

Recommended Reading and References

- Arnold C (2001) *The ancient cliff dwellers of Mesa Verde*. Clarion Press, New York
- Clarke RV (1999) Hot products: understanding, anticipating and reducing demand for stolen goods. Police Research Series Paper 112. Home Office, London, UK
- Ekblom P (2001) Future imperfect: preparing for the crimes to come. *Crim Just Matt* 46(1):38–40
- Eliason SL (2012) Trophy poaching, a routine activity perspective. *Dev Behav* 33(1):72–87
- Farrell G, Pease K (2006) Criminology and security. In: Gill M (ed) *The handbook of security*. Palgrave Macmillan, New York, pp 179–199
- Felson M (2006) *Crime and nature*. Sage, Thousand Oaks
- Felson M, Rachel Boba Santos (2009) *Crime in everyday life*, 4th edn. Sage, Thousand Oaks
- Jeffrey CR (1977) *Crime prevention through environmental design*. Sage, Thousand Oaks
- Laycock G (1991) Operation identification, or the power of publicity? *Secur J* 2(2):67–71
- Lemieux AM, Clarke RV (2009) The international ban on ivory sales and its effects on elephant poaching in Africa. *Brit J Criminol* 49(4):451–471
- Lingenfelter RE (1986) *Death valley and the Amargosa: a land of illusion*. University of California Press, Berkeley
- Liu JS (2009) *Barbed wire: The Fence that changed the west*. Mountain Press, Missoula
- Newman O (1972) *Defensible space, crime prevention through urban design*. Macmillan, New York
- Pires SF (2012) The illegal parrot trade: a literature review. *Global Crime* 13(3):176–190
- Pires SF, Clarke RV (2011) Sequential foraging, itinerant fences and parrot poaching in Bolivia. *Brit J Criminol* 51:314–335
- Pires SF, Clarke RV (2012) Are parrots craved? An analysis of parrot poaching in Mexico. *J Res Crime Delinq* 49(1):122–146
- Pires SF, Moreto WD (2011) Preventing wildlife crimes: solutions that can overcome the 'tragedy of the commons'. *Eur J Crimin Pol Res* 17(2):101–123

Tilley N, Laycock G (2002) Working out what to do: evidence-based crime reduction. Crime Reduction Research Series Paper 11, Home Office, London
 Wortley R, Mazerolle L (eds) (2008) Environmental criminology and crime analysis. Routledge, New York

SNAP[®]

▶ [Stop Now and Plan \(SNAP[®]\) Model](#)

SNAP[®] Boys

▶ [Stop Now and Plan \(SNAP[®]\) Model](#)

SNAP[®] GC

▶ [Stop Now and Plan \(SNAP[®]\) Model](#)

SNAP[®] Girls

▶ [Stop Now and Plan \(SNAP[®]\) Model](#)

SNAP[®] ORP

▶ [Stop Now and Plan \(SNAP[®]\) Model](#)

So What Criminology?

Roger Matthews¹ and Jennifer Gibbs²

¹University of Kent, Canterbury, Kent, UK

²West Chester University at Pennsylvania, West Chester, PA, USA

Synonyms

[Administrative criminology](#); [Critical criminology](#); [Public criminology](#); [Realism](#)

Overview

Despite the continued expansion of academic criminology and the proliferation of publications, the policy relevance of criminology has noticeably decreased in recent years. Some commentators have suggested that academic criminology is becoming irrelevant socially and politically and that it has become increasingly marginalized from the process of policy making (Currie 2007). This has led to the identification of a growing body of what has become referred to as “So What?” criminology.

“So What?” criminology has a number of characteristics according to Elliott Currie (2007) including the use of impenetrable language; the adoption of highly technical research techniques which are often dauntingly quantitative, as well as a tendency to focus on trivial issues; and an orientation towards what C. Wright Mills (1959) called “abstracted empiricism.” This makes research results difficult to read by anyone other than academic experts in the discipline, while the bulk of the literature remains inaccessible to both policy makers and the general public.

This body of work is produced in an academic context where there is an increased emphasis on publishing in peer-reviewed journals combined with a growing reluctance to engage in the challenging and time-consuming business of policy formation and development. Criminology is also becoming increasingly fragmented and organized into different theoretical and methodological silos, which has to some extent undermined the possibility of constructive debate in the subject producing a body of largely incompatible and partial knowledge. In contrast to these developments, realist criminology argues for the integration of theory, method, and practice and aims to move criminology in the direction of developing a more constructive and socially engaged approach to crime control. Thus, it is suggested that adopting a realist approach provides the possibility of developing an alternative to “So What?” criminology, which is both critical and useful.

Introduction

There is a growing paradox in criminology. Despite the increasing scale of the criminological enterprise involving an increasing number of students, courses, and publications, the policy relevance of criminological research has been decreasing. To make matters worse, this growing body of criminological endeavor suffers from a paucity of theory combined with the adoption of inadequate or inappropriate methodologies. This has provided the basis for the development of what has become known as “So What?” criminology.

This entry will outline some of the main features of “So What?” criminology, which include weak conceptualization, inadequate methodologies, and limited policy relevance. It will also examine some of the recent developments within criminology that have contributed to the growth of “So What?” criminology. Finally, we will offer some thoughts on the contribution that realist criminology can make to overcoming the limitations of “So What?” criminology.

The Main Features of “So What?” Criminology

The term “So What?” criminology refers to those forms of criminology that involve a low level of theorization and thin, inconsistent, or vague concepts and categories; embody a dubious or inappropriate methodology; or have little or no policy relevance (Matthews 2010). Currie (2007) has suggested that in order to overcome the deficiencies of “So What?” criminology, researchers should always aim to present their findings in a way that is easily accessible to policy makers, the general public, and practitioners. Thus:

...some areas of criminological research have become so specialized, so technical and so driven by what has gone before that even many people within the discipline cannot really follow it without help. This means that the need for people who can do the work of synthesis and contextualization of research has become ever greater. (Currie 2007: 183)

Research findings and “criminological wisdoms,” he suggests, should be disseminated wherever possible through public debate in order to foster a “public criminology.” However, the problem with this approach is that there is little consensus in the field and that most “criminological wisdoms” are highly contested, and it is difficult to find much agreement among criminologists on any issues of significance.

Others have advocated a sort of division of labor within the subject dividing researchers into professional, policy, and critical criminologists, each developing their own forms of expertise and ideally contributing to each other’s endeavors (Burawoy 2004). However, fostering such divisions within academic criminology is likely to increase the existing fragmentation of the subject and is more likely to exacerbate than resolve the growing irrelevancy of the ever-expanding criminological enterprise.

One of the most noticeable limitations of much contemporary criminology is poor conceptualization and lack of theoretical rigor. This is most evident in the way in which the central concepts like “crime” and “race” are used. Within the criminological community, there is little clarity or agreement how the concept of crime is to be conceptualized, and most criminologists are divided between those who think that it is an act and those that see it is a product of social reaction. Some use essentially legalistic definitions of crime, while others want to dispense with the concept of crime altogether and replace it with notions of “harm” or human rights (Muncie 1996; Hillyard et al. 2004). Realists have argued that “crime” is a process of action and reaction and have developed the sensitizing notion of the “square of crime” pointing out that the construction of “crime” is a *process* involving four basic elements – the offender, the victim, the state and public opinion, and the social norms (Young and Matthews 1992).

Similarly in relation to “race,” rather than deconstructing this concept, all too often “race” is equated with skin color, and diverse ethnic groups are lumped together as if they formed a homogeneous group. Strangely in country like

America that has one of the most diverse ethnic populations in the world, much criminological investigation operates as if the American (criminal) population were divided into “blacks” and “whites,” resulting in a form of “monochromatic criminology.” In this bifurcated world, even Hispanics, which are soon to be the second largest ethnic group in America, are paid relatively little attention.

The theoretical and conceptual limitations of contemporary criminology, however, go far beyond these two examples. Administrative criminology, which emanates from governmental departments and which plays an increasingly major role in funding and conducting criminological research, is notoriously weak conceptually. Its theoretical weakness, often combined with dubious methodological approaches, tends to result in weak or equivocal conclusions and policy proposals. Often, using common sense or taken-for-granted categories of “crime,” administrative and related criminologies veer towards a form of pragmatism which paradoxically is neither very practical nor makes much contribution to developing a cumulative knowledge about crime.

Administrative criminology also tends to gravitate towards empiricism or positivism and assumes that scientific investigation is based on what can be observed. Believing that the lack of theory can be compensated for by the use of “sophisticated” statistical methods, these criminologies tend to reduce the richness and complexities of crime and punishment to a “numbers game” (Young 2004). However, no amount of statistical manipulation can compensate for weak conceptualization.

The problem of method, however, goes far beyond debates about sample size, statistical techniques, and the problems of representation and generalization. Instead, criminological researchers must begin by asking a more fundamental question about how the social world can be best appropriated and understood. Criminologists operate in open systems, which are complex and messy. Social reality does not present itself in an obvious and unproblematic way. To paraphrase Marx: if the social world were intelligible

at the level of appearances, there would be no need for social scientific investigation at all. Thus, the aim of investigation is to go beyond appearances and that which is directly observable and to uncover the complexity of a socially stratified and diverse reality. Thus, we need to move beyond simple description of the world and instead try to understand the causal processes involved. At the same time, we need to consider the relation between agency and structure and how the actions of particular individuals or groups are conditioned or shaped by the social structures in which they operate. The ultimate aim of criminological investigation is therefore not to produce descriptions or “facts,” as if these speak for themselves, but *explanations*.

Empiricism comes in many forms. In criminology, there is a tendency to what we might call “functional empiricism” and “inverted empiricism.” Functional empiricism refers to those approaches which see two observable phenomena that seem to change together and conclude that one must be influencing the other. Thus, Loic Wacquant (2009), for example, claims that there is a “functional equivalence” between the decline of black ghettos in America and the increased incarceration of African Americans and that this is a consequence of neoliberalism. However, why neoliberalism with its commitment to minimal statism would want to expand state provision in this way rather than leaving ethnic minorities to fend for themselves in the ghetto is not clear. It is also the case that in periods of neoliberalism, the number of people sent to prison has decreased as well as increased, suggesting that there is no necessary relation between neoliberalism and penal expansion (Matthews 2009b).

Inverted empiricism refers to those forms of investigation that claim, for example, that there is no causal relationship between crime rates and incarceration on the basis that these two rates do not always covary (see, e.g., Tonry 2001). However, comparing overall crime rates and imprisonment rates is problematic because measurement of these two variables changes over time and between different locations. This line of thinking by focusing on immediate observables tends to neglect a more detailed critical

investigation into the attrition rates of different offenses particularly those offenses that are more likely to result in a prison sentence. Thus, naïve empiricists like Michael Tonry (2001) who claims that crime rates and imprisonment rates are unconnected because they do not correlate avoid examining the causal processes in play, and how changing patterns of crime over time as well as changes in the way in which different crimes are processed can influence the scale of imprisonment. An investigation of this type would reveal that the increase in incarceration rates in America is not so much a function of the police and judiciary getting tough on drug offenders, particularly those from ethnic minority groups, as Michael Tonry (1995) claims, but a significant increase in the number of violent offenders being processed (Bureau of Justice 2008).

The limited policy relevance of much criminology is associated with the theoretical and methodological weaknesses outlined above. In addition, many criminologists see themselves as professional academics in the way that Burawoy (2004) describes and largely divorced from policy and practice. Often, under increasing pressure to produce well-cited publications and manage increased teaching and administrative loads, a growing body of criminologists does not have the time or inclination to engage in public or political debate. Among those who are interested in engaging with policy and practice, it is often the case that because they present “facts” rather than credible explanations, their work often does not resonate with politicians or the public. In addition, findings are not presented in an accessible form.

All too often, criminologists dwell on the failures or limitations of different programs and fail to address the question of “what is to be done.” Thus, in relation to major issues such as imprisonment and policing, we see endless publications listing the failures and the problems of these institutions and organizations, rather than engaging in a constructive discussion on how they might be improved or usefully reformed.

The Construction of “So What?” Criminology

There are a number of developments that have taken place within criminology over the last 20 or 30 years that have contributed to the development of “So What?” criminology in different ways. First, as mentioned above, there has been a considerable expansion in administrative criminology as a result of the increased dominance of government funding of criminological research. Reese Walters (2003) has questioned how “market-led criminology,” which focuses on risk management, privatization, and cost-effectiveness, has influenced the production of criminological knowledge. Government-funded forms of administrative criminology, he suggests, are not interested in generating critical and reflexive research and have in recent years come to undermine and sideline critical criminological research. There can be little doubt that the changing nature of the academy coupled with the increasingly narrow focus of much government-sponsored research has produced a growing body of largely atheoretical research with a limited policy edge. Consequently, much “administrative criminology” often involves policy-driven evidence rather than evidence-driven policy.

Second, postmodernism has had a major influence on the social sciences in general and criminology in particular. Postmodernists have been critical of what they refer to as “grand narratives” and of the modernist belief that social progress can be achieved through the application of scientific knowledge to social problems. Postmodernists consequently present a form of relativism and generally express an anti-objectivism and antirealism. Although postmodernists often end up in producing their own meta-narratives and engage in debates which imply at least some degrees of objectivism and an independent reality about which we can debate, the advent of postmodernism has fuelled a relativism that was already present in criminology. This relativism suggests on one side that everybody’s view is equally valid or that certain “standpoints” have privileged position in relation to truth.

Thus, Carol Smart (1992) in her well-referenced critique of “malestream criminology” in which she berates criminology for what she sees as its adoption of grand narratives argues instead for situated knowledge. Suspicious of what she sees as an essentially male-centered criminology, she argues that the truth claims of criminology have to be deconstructed and challenged from a feminist standpoint. Although rightly critiquing partial views of the world that claim to be universal, Smart wants to replace these one-sided views with another set of one-sided views. As John Lea (1998) has argued, replacing the views of the dominant group with those of the marginalized, victimized, and excluded women is to replace one form of fundamentalism with another. In the end, Smart muddies the waters rather than constructing a basis for the development of a feminist criminology. By advocating a form of standpoint feminism together with an antimodernism and antirealism, Smart gravitates towards relativism and foundationalism. From this perspective, it is difficult to see how she might develop a policy response that would reduce rape, domestic violence, and other crimes that victimize women. Thus, postmodernism tends towards defeatism and does little to move the field forward in developing meaningful and effective criminal justice policies.

Third, in conjunction with postmodern defeatism, there is a significant trend in academic criminology towards liberal pessimism. The implicit message of this stance as Stanley Cohen (1985) pointed out is to try to do less harm rather than more good. This brand of pessimism was popularized in the 1970s with the publication of Martinson’s (1974) report on corrections – a report that was widely interpreted as claiming that “nothing works.” Although Martinson in a subsequent article admitted that certain programs did work for some people under certain conditions, the pessimists embraced the “nothing works” mantra and applied it to different areas of criminology. Significantly, the claim that rehabilitation strategies were ineffective in reducing recidivism allowed a shift away from a commitment to try to

rehabilitate prisoners to a policy of incapacitation and “just deserts.” This in turn led to the warehousing of prisoners and the creation of what were described in the UK as “penal dustbins.”

Liberals, in fact, have found it difficult to develop policies and interventions that are credible and which might resonate with the policy makers and the general public. This is partly because many are antistate or at least minimal statist, while most are anti-punishment. For the most part, their arguments are directed towards the reduction of state intervention and of limiting the use of punishment. Thus, there is a general policy of reductionism among liberal criminologists who argue for a decrease in the prison population and in many cases the use of community-based sanctions instead. However, because liberals often fail to address the issue of who should go to prison and for what purpose and for how long, the question of how far the prison population needs to be reduced can never be answered. The issue of imprisonment is, however, not just a question of numbers. Moreover, as some criminologists have pointed out, the greater use of community-based alternatives to incarceration can lead to “net widening” and the simultaneous expansion of both inclusive and exclusive strategies of control (Cohen 1985).

From the vantage point of liberal pessimists, there is a general wariness about proposing “solutions” or “alternatives.” For some, there is little interest in improving prison conditions as this is only seen to relegitimize the use of incarceration. On the other hand, there is a growing skepticism about the use of community-based sanctions. At the same time, there is some suspicion concerning the use of the “welfare sanction” which is seen to deepen the level of state intervention into individual lives while extending the range of intervention to include the wider community (Cohen 1985; Garland 1981). Thus, liberal pessimists find it difficult to commit to any of these forms of regulation.

More recently, liberal pessimists have claimed that there has been a rise in “populist punitiveness” and that both politicians and the public are becoming increasingly punitive (Garland 2001;

Simon 2001; Wacquant 2009). Although these claims are based on thin theoretical and empirical foundations, there is a substantial consensus among liberal pessimists that we are witnessing a punitive turn (Matthews 2005). The main reference points for many of these liberal pessimists are the growing prison population on both sides of the Atlantic. The increase in the number of people in prison coupled with decreasing crime rates is seen as uncontroversial evidence of an increase in punitiveness. However, in relation to the American situation, the prison population leveled off at the end of the 1990s, while in a number of states, the prison population has actually decreased over the past decade. In the UK, we have seen a bifurcated response in recent years with the judiciary giving longer sentences for certain offenses like violence and murder, while giving reduced sentences for other offenses like theft and burglary (Matthews 2009a). This does not look much like the blanket surge in punitiveness that pessimistic liberals suggest. Also, in other countries such as Italy, Canada, and Germany, the incarceration rate has remained stable or declined in recent years (Meyer and O'Malley 2005). In short, these liberal pessimists add little that is useful or constructive to policy formation and in fact in relation to the debate around punitiveness have led us into a conceptual cul-de-sac. Unfortunately, doing nothing or trying to find reasons why proposed interventions are destined to fail creates a form of disengagement that allows policy makers and politicians a relatively free hand to orchestrate criminal justice policy. Criminal justice, however, is too important to allow politicians to make uninformed decisions.

Fourth, the demise of critical criminology in recent decades has contributed to the resurgence of conventional criminology with its natural scientific orientation. At its height of influence in the 1970s, critical criminology provided useful theories on crime and deviance including labeling and subcultural theories. Importantly, critical criminology encouraged critical discussions about the role of power and politics in relation to crime and justice, as well as raising questions regarding values and meanings.

However, critical criminology was unable to offer a credible alternative to conventional criminology. In developing a skeptical view of convention attitudes towards crime, it tended towards a romantic and idealistic view of the criminal while downplaying the impact of crime on victims.

Eventually critical criminology was largely absorbed within the framework of mainstream criminology and became at best the "bad conscience" of conventional criminology. The demise of critical criminology was unfortunate, as it left the door open for positivism to gain ascendancy within academic criminology. In recent years, however, we have witnessed the growth of cultural criminology that embodies an important critical impulse (Ferrell et al. 2008). Cultural criminology focuses on the experiences of marginalized groups, emphasizing questions of agency, motivation, emotions, and the generation of meaning. Cultural criminology, though, largely ignores the role of the state and wider issues of power. It has been criticized for acting as "zookeepers of deviance," while there is some confusion over what is meant by "culture" (O'Brien 2005). The problems are seen to limit its critical capacity, while from a realist perspective, cultural criminology like its critical criminological predecessors has a limited engagement with policy and social reform. Thus, while cultural criminology has no doubt reinvigorated critical interest among academic criminologists and in many respects breathed new life into the subject, its ability to provide a coherent alternative to mainstream criminology in its current form remains limited.

Countering "So What?" Criminology

By focusing on identifying causal connections between phenomena, realists seek to understand how and why change occurs. In contrast to post-modern defeatism and liberal pessimism, realist criminology involves a modernist perspective, promoting an evidence-based approach to social reform. Following its critical criminological predecessors, realist criminology aims to

deconstruct key concepts like “crime” and “race” and to provide explanations rather than present a series of facts or decontextualized descriptions. Also drawing on the work of cultural criminologists, it aims to incorporate an appreciation of the cultural dimensions of crime control recognizing that policy formation will always be subject to cultural variations. In this way, realist criminology aims to develop a joined-up approach that incorporates sophisticated conceptualization and a robust methodology while making a contribution to policy formation.

Realists on Theory. First and foremost, realists recognize that “crime” is not simply a given nor is it an “act” or simply a product of “social reaction.” Rather it is a process which involves a complex and changeable relation between offenders, victims, the state, and the public. It aims to avoid the relativism of postmodernists and the idealism of some forms of critical criminology. The aim is neither to romanticize nor demonize the offender but rather to understand the motivational and structural dynamics that create crime and victimization on one hand and which shape formal and informal responses on the other. Understanding the “square of crime” is important in order to remind ourselves that “crime” is neither a “top-down” construction imposed by the criminal justice system nor a “bottom-up” process involving certain “acts” or “behavior” or changing levels of tolerance but a complex relation between these different determinants. For realists, the development of theory is important but is not seen as an end in itself. Theory must be useful and usable.

Realists on Method. Although realists are critical of empiricism and positivism, they are also critical of those qualitative approaches that are purely descriptive. Consequently, from a realist perspective, purely descriptive ethnographies are of no greater value than the most dense and impenetrable forms of inferential statistics. Realists also reject those forms of “cookbook criminology” that claim that there is one methodological approach that is intrinsically superior to all others and should be followed on all occasions. In many respects, realists are methodological pluralists and argue that the selection

of methods will be dependent on the nature of the object under study and the type of research questions that are developed (Sayer 2010).

Rather than rely on statistical correlations, the aim of realist investigation is to identify causal connections. The objective is to find out not only what works but how and why it works. It is also recognized that “what works” will vary according to context and also according to the subjects that interventions are directed towards. As Pawson and Tilley (1997) point out, it is not so much that certain programs “work” but that they are effective to the extent that they connect with the capacities of the subjects to whom they are directed.

Realists on Practice. The move away from “nothing works” (Martinson 1974) to “what works” (Sherman et al. 1997) has been a welcome break from the pessimism and impossibilism that permeate contemporary criminology. While the field today focuses on using evidence to form policy, the evidence for “what works” has been disappointing, largely due to a low level of theory and poor conceptualization combined with inadequate and inappropriate methodologies. A great deal of administrative criminology adopts an instrumentalist view of “what works” and pays scant regard for differences in contexts and in the capacities and propensities of the subjects at whom interventions are aimed. Given the complex and ever-changing contexts of social life, the difficulty is to find situations in which successful interventions can be replicated. Indeed, nothing works for everyone all the time. Thus, instead of claiming that a program or policy “works,” perhaps it is more appropriate to claim that some programs work for some people under certain circumstances.

Critical realists also have a distinctive view of the nature and meaning of interventions. Interventions are not just practices but theories or hypotheses that postulate the possibility of bringing about an improved outcome. Consequently, interventions are potentially fallible particularly because they deal with complex social realities, as well as dealing with different groups of subjects and will invariably be implemented differently in different contexts.

Intervention provides tests for theories and hypotheses and potentially offers a basis for developing cumulative knowledge about what works in different contexts and for different populations.

Conclusion

Realist criminology, it has been suggested, offers an alternative to “So What?” criminology and aims to provide a coherent and constructive alternative that is able to provide an evidence-based approach to social reform. The task facing criminology is to develop a coherent and integrated approach that emphasizes the role of a theory and which is critical but grounded. Realist criminology argues for the deployment of methods of investigation that are responsive to the object under study and are designed to identify how and why measures work and ultimately to fashion interventions that are aimed at making tangible improvements in the world. In short, realist criminology is theoretical but not theoreticist, critical but not negative or impossibilist, Utopian but grounded in lived experience, methodologically flexible but rigorous, practical but not pragmatic, and policy relevant rather than policy driven.

Related Entries

- ▶ [Cultural Criminology](#)
- ▶ [Left Realism](#)
- ▶ [Postmodern Criminology](#)
- ▶ [Randomized Experiments in Criminology and Criminal Justice](#)

Recommended Reading and References

Burawoy M (2004) Public sociologies: contradictions, dilemmas and possibilities. *Soc Forces* 82(4):1603–1618

Bureau of Justice (2008) Adult correctional population 1980–2007. www.ojp.gov/bjs/glance/covr2.htm.

Cohen S (1985) Visions of social control. Polity, Cambridge

Currie E (2007) Against marginality: arguments for a public criminology. *Theor Criminol* 11(2):175–190

Ferrell J, Hayward K, Young J (2008) *Cultural criminology*. Sage, London

Garland D (1981) The birth of the welfare sanction. *Brit J Law Soc* 8:29–45

Garland D (2001) *Culture of control*. Oxford University Press, Oxford

Hillyard P, Pantazis C, Tombs S, Gordon D (eds) (2004) *Beyond criminology: taking harm seriously*. Pluto Press, London

Lea J (1998) Criminology and postmodernity. In: Walton P, Young J (eds) *The new criminology revisited*. Macmillan, London

Martinson R (1974) What works? Questions and answers about prison reform. *Public Interest* 35:22–54

Matthews R (2005) The myth of punitiveness. *Theor Criminol* 9(2):175–202

Matthews R (2009a) Beyond ‘so what?’ criminology: rediscovering realism. *Theor Criminol* 13(3):341–362

Matthews R (2009b) *Doing time: an introduction to the sociology of imprisonment*, 2nd edn. Palgrave/Macmillan, London

Matthews R (2010) The construction of ‘So What?’ criminology: a realist analysis. *Crime Law Social Ch* 54:125–140

Meyer J, O’Malley P (2005) Missing the punitive turn? Canadian criminal justice, balance and penal modernism. In: Pratt J, Brown D, Brown M, Hallsworth S, Morrison W (eds) *The new punitiveness*. Willan, Collumpton

Mills CW (1959) *The sociological imagination*. Oxford University Press, New York

Muncie J (1996) The construction and deconstruction of crime. In: Muncie J, McLaughlin E (eds) *The problem of crime*. Sage, London

O’Brien M (2005) What is cultural about cultural criminology? *Brit J Criminol* 45:599–612

Pawson R, Tilley N (1997) *Realist evaluation*. Sage, London

Popora D (1998) Four concepts of social structure. In: Archer M, Bhaskar R, Collier A, Lawson T, Norrie A (eds) *Critical realism: essential readings*. Routledge, London

Sayer A (2010) *Method in social science: a realist approach*, 2nd edn. Routledge, London

Sherman LW, Gottfredson D, MacKenzie D, Eck J, Reuter P, Bushway S (1997) *What works, what doesn’t, what’s promising: a report to the United States congress*. University of Maryland, College Park, Maryland

Simon J (2001) Entitlement to cruelty: neoliberalism and the punitive mentality in the United States. In: Stenson K, Sullivan R (eds) *Crime, risk and justice*. Willan, Collumpton

Smart C (1992) Feminist approaches to criminology or postmodern woman meets atavistic man. In: Morris A, Gelsthorpe L (eds) *Feminist perspectives in criminology*. Open University Press, Buckingham

Tonry M (1995) *Malign neglect: race, crime and punishment in America*. Oxford University Press, New York

- Tonry M (2001) *Penal reform in overcrowded times*. Oxford University Press, New York
- Wacquant L (2009) *Punishing the poor: the neoliberal government of social insecurity*. Duke University Press, Durham
- Walters R (2003) New modes of governance and the commodification of criminological knowledge. *Soc Legal Stud* 12(1):5–26
- Young J (2004) Voodoo criminology and the numbers game. In: Ferrell J, Hayward K, Morrison W, Presdee M (eds) *Cultural criminology unleashed*. Glasshouse Press, London
- Young J, Matthews R (1992) *Rethinking criminology: the realist debate*. Sage, London

research has established that collective efficacy is an important predictor of differences in the level of crime between neighborhoods and may be significant to our understanding of the disproportionate prevalence of crime in neighborhoods characterized by economic disadvantage and residential instability (Sampson et al. 1997; Morenoff et al. 2001). However, important questions remain about the way residents perceive collective efficacy, the different ways they can choose to intervene when they become aware of a crime problem, and the connections between collective efficacy and other neighborhood phenomena including cultural forces and more formal crime control efforts.

Social Bond

► Social Control

Social Capital and Collective Efficacy

Kevin M. Drakulich
School of Criminology and Criminal Justice,
Northeastern University, Boston, MA, USA

Overview

Criminologists have long theorized that some neighborhoods have a greater capacity than others to recognize common problems and act collectively to solve them. Though this basic notion is derived directly from social disorganization theory, collective efficacy – a recent reformulation – has brought substantial new attention to the idea. Collective efficacy borrows from the idea of social capital to rethink the ways that this capacity is rooted in the local connections among neighbors. Specifically, *social capital* refers to a resource potential facilitated by the structure of local networks, while *collective efficacy* is the ability of a group to draw on this resource to recognize common interests and achieve specific tasks related to local social control. This reformulation has also garnered attention for innovations in the operationalization and measurement of the concept. Recent

Introduction

In 1942, a now-classic volume was published which made a radical suggestion: the high crime rates found in immigrant and African-American neighborhoods were not the product of the inherent criminality of individual members of these groups – a popular explanation at the time. Instead, Shaw and McKay revealed that delinquency rates were related to the social-structural organization of the city and were highest in the “transitional zone,” an area characterized by low home ownership, socioeconomic disadvantage, and other social-structural disadvantages. The enduring power of their work is derived from the carefully collected and presented evidence of this link between neighborhood structure and crime rates.

Shaw and McKay (1942) also pose an explanation for this link, suggesting that, in simple terms, there is something about some neighborhoods that allows them to organize in ways that prevent crime problems or respond to them as they occur. Notably, however, this is somewhat of an amorphous concept, and both Shaw and McKay (1942) as well as a number of subsequent researchers have struggled with two important issues. The first is how to *theoretically* conceptualize exactly what is important about the social organization of a neighborhood. The second is how to *measure* this theoretical concept systematically such that its role as a mechanism can be tested.

After more than 50 years of incremental and mixed progress on these issues, a 1997 study purported to address both issues. *Collective efficacy* attempts to capture “the differential ability of neighborhoods to realize the common values of residents and maintain effective social controls” (Sampson et al. 1997, p. 918). While earlier work relied on either administratively collected data or aggregations of individual reports of personal behaviors, collective efficacy employed individuals as reporters of neighborhood conditions. Any one of these respondents, on their own, may be biased or simply in error in their assessments. Collectively, however, and when controlling for potential individual-level sources of bias, such respondents may produce a reliable and valid measure of a collective phenomenon – or, as Warner and Rountree characterize it, a Durkheimian *social fact*: “a thing distinct from its individual manifestations” (1997, p. 520).

By situating this collective capacity for control in the resource potential of specific forms of personal and organizational networks (Sampson et al. 1999, p. 635), collective efficacy potentially represents significant progress in our understanding of the group social processes relevant to the spatial distribution of crime. This entry begins by providing background on collective efficacy as a theoretical mechanism, tracing its development through several previous propositions. Collective efficacy, however, departs from previous propositions through several theoretical and measurement innovations. The following section details these innovations and surveys research on its importance to a variety of neighborhood phenomenon. The final section raises a number of potential controversies and open questions facing future work on collective efficacy.

Background

Simply stated, social disorganization theory suggests that the structural conditions of a neighborhood, especially the level of poverty and residential instability, disrupt local social organization and thereby increase crime. The

direct association between the structural conditions and the crime rate is relatively easy to observe using administrative data, such as the US Census and the Uniform Crime Reports. More difficult to observe is the social mechanism by which this structure affects crime rates. Over time, research on social disorganization has frequently revised the theoretical operationalization of what it is about the organization of the neighborhood that matters, while simultaneously experimenting with a variety of different more or less successful means of measuring these theoretical mechanisms.

Shaw and McKay (1942), in their initial proposal of the theory, suggest several theoretical mechanisms. Some components of their explanation – including value heterogeneity and the cultural transmission of delinquent traditions between generations of boys within neighborhoods – have been largely ignored or dropped in subsequent revisions to social disorganization, including in the notion of collective efficacy. One component of their explanation, however, is the direct intellectual antecedent of collective efficacy. Drawing on a classic account of Polish communities in Chicago by W.I. Thomas, Shaw and McKay (1942) suggest that an “organized” community is characterized by “the presence of social opinion with regard to problems of common interest, identical or at least consistent attitudes with reference to these problems, the ability to reach approximate unanimity on the question of how a problem should be dealt with, and the ability to carry this solution into action through harmonious co-operation” (p. 184).

In subsequent years, Kornhauser’s (1978) critique of the theory led to an increasing focus on the mechanism of social control to the exclusion of the subcultural and learning mechanisms – those explanations interested in value heterogeneity or the transmission of delinquent skills and values. Simultaneously, an emphasis on social ties among neighbors rose to prominence following Kasarda and Janowitz (1974) argument for a systemic model of social organization in which the local community is “a complex system of friendship networks and formal and informal associational ties” which are in part a product of

local structural conditions (p. 329). Out of these trends came a new operationalization of the mechanism at the heart of social disorganization: a disruption of the social networks among neighbors which damages the neighborhood's capacity to self-regulate (e.g., see Bursik and Grasmick 1993).

While research inspired by these ideas represented both theoretical and methodological progress, it was still hindered by basic measurement issues. In perhaps the best example of this work, Sampson and Groves (1989) operationalize the mechanism of social disorganization as a combination of sparse local friendship networks, the presence of unsupervised teenage peer groups creating a nuisance, and low levels of organizational participation. In the measure of nuisance-causing unsupervised teenage groups, Sampson and Groves are attempting to directly measure a neighborhood's capacity for informal social control. Unfortunately, by operationalizing social disorganization via an outcome (the presence of rabble-rousing teenagers) rather than the social mechanism itself (the capacity for informal social control), Sampson and Groves open themselves up to the classic criticism of social disorganization research as using delinquency "both as an example of disorganization and something caused by disorganization" (Pfohl 1985, p. 167).

Sampson and Groves (1989) find a strong relationship between their measure of the presence of trouble-causing unsupervised youth groups and a variety of measures of local crime, but their measure of local friendship networks has somewhat less success. Warner and Rountree (1997) find similarly mixed effects for the role of local social ties in relation to local crime rates: they find that measures of the prevalence of interactions between neighbors (borrowing tools, having lunch or dinner, helping neighbors with problems) are negatively related to the crime rate only in largely white neighborhoods and appear unrelated to crime in neighborhoods with larger numbers of racial or ethnic minorities. Critics of the idea of social disorganization pointed out that some high-crime communities in fact had strong *criminal* organization (recent example in

Venkatesh 1997; see discussion in Sampson 2002). Critics of the systemic model of social disorganization pointed out that many poor communities had dense social networks (Gans 1962) and that in some communities, ties between neighborhoods prevented informal social control efforts when they bridged the gap between residents who were involved in crime and those who may otherwise have organized against it (Pattillo 1998).

As of the 1990s, then, research on social disorganization theory was faced with several key issues. First, the exact role of social networks in social organization remained ambiguous. Second, there appeared to be a need to distinguish organizational activities that suppress crime versus those that encourage crime. Third, research in the area continued struggling to identify direct measures of any of these theoretical concepts to test the core hypotheses.

Social Capital and Collective Efficacy

The concept "collective efficacy" attempts to resolve some of the issues in social disorganization theory by grounding it in new ideas about the dynamics of interpersonal connections within groups and the capacity these connections create for collective action. By asking residents to report on their perceptions of their neighbors' reactions to hypothetical local problems, collective efficacy attempts to capture a task-specific *potential* for collective action rooted in the social networks of a neighborhood. Collective efficacy, then, is the collective realization among neighbors that they have a common goal in the safety of their neighborhood and a collective willingness to act to achieve this goal (Sampson et al. 1997).

Collective efficacy has found success by making three key theoretical revisions to older versions of social disorganization theory and by using these revisions to better operationalize and measure the presence or absence of this phenomenon within neighborhoods. Theoretically, Sampson et al. (1997); (see also Sampson 2002) define collective efficacy as rooted in social capital: "the resource potential of personal and

organizational networks” (p. 635), as task specific, and in particular as oriented toward the universally shared goal of living in a safe neighborhood. They find empirical evidence for a role for collective efficacy by operationalizing it as potential rather than actualized organization and as an emergent property of neighborhoods best captured by using local residents as imperfect reporters of this group-level phenomenon. The following sections discuss each of these theoretical and measurement dimensions that distinguish collective efficacy from earlier versions of social disorganization.

Social Capital

Perhaps the most important theoretical development for the reformulation of social disorganization theory is Coleman’s (1988) version of social capital. Coleman (1988) combines economic rational choice theory with sociological normative theory to create the notion of “social capital,” with the premise that “rational or purposive action, in conjunction with particular social contexts, can account not only for the actions of individuals in particular contexts but also for the development of social organization” (p. S96).

In a simple sense, social capital suggests that certain kinds or qualities of connections among people can facilitate particular actions that may not otherwise have been possible. It can take a variety of forms which may be better or worse at facilitating different kinds of actions. One such form is a marketplace of obligations and expectations in which people do favors or help one another with the expectation of indirect reciprocation (Coleman 1988). This form of social capital provides a framework for understanding why a resident would undertake an action – for example, intervening when they see teens vandalizing public property or a neighbor’s house – that has minimal direct benefits for themselves. The key is that when that resident lives in a community with an active marketplace of obligations and expectations, they can expect that their favor will be reciprocated by others at some other point when the resident currently performing the favor has a need of their own arise. Other forms of social capital include the

transmission of potentially useful information through networks maintained for other purposes or the presence of norms and effective sanctions. Coleman (1988) describes, for example, a norm in which all local adults are responsible for the safety of all local children in public spaces.

These forms of social capital are facilitated by social structure. The trust required for a marketplace of obligations and expectations, for instance, is more likely to emerge in a community with relative residential stability – the presence of long-term residents who may be expected to remain in the neighborhood long enough to repay debts into the collective pool. Such a marketplace may also be more likely when residents have enough resources to be able to do favors for neighbors but not so many resources that they will never need to ask neighbors for favors (Coleman 1988). Coleman (1988) also identifies particular kinds of relational structures that produce social capital, one of which is relevant here. *Closure* in social networks refers to the presence of sufficient ties between network actors. One specific form of closure relevant to collective efficacy is intergenerational closure, in which parents know the parents of their children’s friends, through which they can more effectively monitor and sanction their own children. In simpler terms, it is harder to do something without your parents’ knowledge when they know all your friends as well as the parents of your friends. Thus, network closure is an aspect of social structure that facilitates certain kinds of actions within the structure.

Social capital, then, suggests that elements of local social organization provide neighborhoods with a social resource that can be drawn upon to address local problems. For instance, in highly residentially stable neighborhoods, there is more likely to be a high degree of interconnectedness among residents. Due to this interconnectedness, residents may feel a sense of trust that if they help a neighbor, another neighbor may be willing to help them when they need it. This, Coleman argues, provides a social resource potential that can be drawn upon to address local problems. Similarly, interconnections between local parents and local children may both increase the

surveillance of and provide greater adult resources for local children. Thus, social capital describes a process by which structural characteristics of a neighborhood like residential stability assist in the development of particular forms of social organization, which in turn can facilitate specific kinds of local actions.

Collective Efficacy as Task-Specific Action

One key to social capital is that – like physical or human capital – it often exists as a value-neutral tool that could be appropriated to different kinds of actions. A person could use a biology degree to cure cancer or create a weapon; spray paint can be used for art or vandalism. Similarly, a neighborhood could use comparable kinds of social capital alternatively to address a rash of burglaries, to prevent African-Americans from moving into their neighborhood, or even to encourage residents to refrain from reporting crimes to or cooperating with the police.

One type of action that can result from this social resource potential is attempts to control local delinquency or crime. *Collective efficacy* is the collective realization among neighbors that they have a common goal in the safety of their neighborhood and a collective willingness and ability to act to achieve this goal (Sampson et al. 1997). Notably this conception of social organization has changed little since Shaw and McKay's (1942) description of "harmonic cooperation." The theoretical advancements since Shaw and McKay have largely come in two areas: conceptualizing this organization as task specific (organization specifically against crime will be more relevant to crime as an outcome than more general kinds of organization) and developing a better model of how such a capacity might develop, especially as a product of the structural characteristics of a community. In practice, the task-specific potential to address crime is so highly related to the presence of specific local forms of social capital that the most frequent measure of collective efficacy combines a direct measure of the capacity for informal social control with a measure of social cohesion and trust (Sampson et al. 1997; Morenoff et al. 2001).

This conceptualization helps resolve some of the questions about the role of social ties and networks from prior work. Specifically, Sampson et al. (1999) suggest that collective efficacy is a task-specific potential for action that is rooted in social capital: "the resource potential of personal and organizational networks" (p. 635). In this way, collective efficacy departs from some of the more recent versions of social disorganization by de-emphasizing the importance of close friendship networks and instead suggesting that a resource potential exists in particular forms of local social organization, including more utilitarian and less affective connections among persons. Thus, a place where neighbors trust one another enough to develop a marketplace of obligations and expectations may possess collective efficacy even when neighbors are not actively friendly with one another, while a different neighborhood with dense friendship ties may still lack it. In fact, Granovetter (1973) has described the ways in which the preponderance of dense and largely redundant ties will be less useful in exposing people to new ideas or information than ties that bridge the gaps between different groups of people.

Measurement as Potential for Action

The key to measuring this new concept is to capture the *potential* for collective action rather than measuring the presence or absence of potential outcomes of such actions, as Sampson and Groves (1989) did with their measure of perceived juvenile delinquency. The potential, in this case, is reflected in a norm: *shared expectations* for informal social control. The measurement of this kind of norm – something that exists between neighbors rather than within any one neighbor – requires the perspectives of those residents. Sampson et al. (1997) accomplish this by using questions which ask residents to speculate about how their neighbors would respond to a series of hypothetical situations involving delinquent acts by youths. Respondents are asked about the likelihood that their neighbors would take action if they encountered one of a number of instances of youth delinquency. These include children skipping school and hanging out

on a street corner, children spray-painting graffiti on a local building, children disrespecting an adult, and circumstances in which a fight breaks out in front of their house. Additionally, some versions of collective efficacy have included an additional noncrime outcome: the likelihood of neighbors acting to save a local fire station threatened by budget cuts.

By asking respondents to report their perceptions of the level of collective efficacy in their neighborhood, they attempt to capture a *potential* for action which exists in the connections among persons in the neighborhood rather than within any one resident or in any simple aggregation of residents. In this sense, collective efficacy represents an emergent property of neighborhoods – shared expectations about the control of local delinquency – that theoretically would not be captured by the simple aggregation of reports of individuals' willingness to engage in such control.

Reflecting the theoretical relationship between social capital and collective efficacy, most measures of collective efficacy also include items capturing the degree of social cohesion and trust in the neighborhood. Though the two are distinct concepts theoretically and it appears plausible that collective efficacy is to some degree a product of local social capital elements such as cohesion and trust, in practice the two appear highly correlated and thus have frequently been combined to create a single scale.

In doing so, collective efficacy appears to be more successful than earlier measures of the social disorganization mechanism in predicting local levels of violent crime and even mediating some of the association between structural conditions and violence (Sampson et al. 1997; Morenoff et al. 2001). Research has also found collective efficacy to be predictive of fear of crime (Gibson et al. 2002), partner violence (Browning 2002; Wright and Benson 2011), as well as outcomes less directly related to crime including health and mental health and sexual behavior outcomes such as pro-social competency and problem behavior, birth weight, premature mortality, short-term sexual partnering, and age of sexual initiation. Collective efficacy is also

suggested to be the source of a spurious disorder-crime relationship suggested by broken windows theory (Sampson and Raudenbush 1999).

Controversies, Unresolved Issues, and Open Questions

By situating the collective capacity for control in the resource potential of specific forms of personal and organizational networks, collective efficacy represents significant progress in our understanding of the group social processes relevant to the spatial distribution of crime. Operationalizing collective efficacy as an emergent task-specific potential for collective action against delinquency and crime is a noteworthy effort to resolve the difficult problem of measuring phenomena which are a property of groups, or a property of the connections among persons, rather than a property of the persons themselves. This concept represents substantial progress over earlier attempts to theorize and measure a neighborhood's capacity for social control. However, several unresolved issues remain which pose questions for future research.

Lack of Mediation and Measurement Issues

Despite evidence for an apparent role for collective efficacy in the distribution of crime over neighborhoods, the structure of a neighborhood continues to exert strong direct effects on the crime rate. Although Sampson et al. (1997) claim strong mediation effects, Morenoff et al. (2001) suggest this claim may have been overstated. Though collective efficacy appears to have a direct effect on crime, it may not be the long-sought mechanism explaining the link between social structure and crime – or it may only be one component of such a mechanism. Alternatively, this lack of mediation may be a product of measurement error related to ways in which collective efficacy has been operationalized – specifically the difficulty involved in speculating about the potential reactions of other people to hypothetical situations. This raises a key issue for both the measurement of the concept of collective efficacy

and the theory itself: we still know too little about how residents make assessments of the local capacity for collective efficacy.

How Do People Intervene?

Another area for future research is to develop a better understanding of how residents choose to intervene when they become aware of crime problems. Residents may directly intervene by confronting troublemakers or they may indirectly intervene by mobilizing external resources including the police department. Residents may be fearful that intervening in local problems could result in retaliation or simply put them at greater risk for victimization. Recent work has suggested just this: that a fear of reprisals acts as an impediment to participation in traditional forms of informal social control (Carr 2003; St. Jean 2007). Instead, such residents may choose to work through the police department to exert control over local crime, actions which do not appear to require the cohesion necessary for informal efforts (Carr 2003; Warner 2007). On the other hand, neighborhoods which lack faith in the police may be wary of choosing to intervene through the police, and in some of these neighborhoods, residents may choose to attempt to address some specific crime problems through local criminal organizations like gangs (Venkatesh 1997; Pattillo 1998). Thus, while collective efficacy research has largely been interested in *whether* neighborhoods are willing to intervene, we may also want to consider *how* neighbors choose to intervene and the differential consequences for neighborhoods resulting from these choices.

Culture and Norms About Crime: Universality or Heterogeneity?

Collective efficacy assumes that communities share common values with respect to local public safety (Sampson 2002) and – following Kornhauser (1978) – has largely eschewed a role for culture in encouraging crime. Kubrin and Weitzer (2003) suggest, however, that culture may play several different roles that are relevant to the collective organization against crime; future work would do well to consider the potential links.

Early criminological work on culture emphasized the prevalence of pro-crime values in lower-class or minority communities. Often portrayed as a reaction to strain over status, respect, or access to legitimate or illegitimate economic opportunities (Cohen 1955; Cloward and Ohlin 1960), such communities were characterized as having pro-crime or crime-tolerant cultural norms. These norms emerge as a negativistic or oppositional reaction to societal norms (Cohen 1955; Anderson 1999), or as an appropriate or predictable reaction to economic circumstances (Miller 1958; Cloward and Ohlin 1960). Shaw and McKay's (1942) early work in social disorganization characterized disorganized neighborhoods as suffering from a kind of normative heterogeneity in which children will be exposed to examples of both criminal and conventional behavior. Sutherland (1947) suggests that both organization against crime and organization in favor of crime will be relevant to understanding the local crime context.

Alternatively, instead of pro-crime values mattering, it may be that communities simply differ in the strength with which they claim conventional norms (Kornhauser 1978; Warner 2003). It may be that these communities are simply more resigned or accustomed to the presence of crime – that crime has infiltrated their cognitive landscape. Finally, culture may matter in ways that have little to do with values about crime. Small (2002), for instance, suggests that cultural framing mechanisms have the ability to either incite or sustain collective organization, including organization toward the goal of addressing local crime problems. For instance, a framing of the neighborhood as in crisis or under threat may *incite* organization (Small 2002; Carr 2003), while a cultural frame defining a neighborhood as one's best opportunity for success is necessary to *sustain* continued organization (Small 2002).

Connections Between Formal and Informal Social Control

Collective efficacy is, in essence, the capacity for *informal* social control. Of course, the police also act as agents of more *formal* social control

efforts, and prior work has suggested that neighborhoods differ in both the behavior of police and perceptions of the police. The capacity for informal social control, then, may also be related to local formal efforts. This idea has only just begun to receive serious attention (see Kubrin and Weitzer 2003). Recent work has found evidence for a link between perceptions of the police and informal social control efforts.

Conclusion

Collective efficacy draws on the notion of social capital to present a new theoretical and methodological conceptualization of differences between neighborhoods in their ability to recognize common problems and act collectively to solve them. Social capital is grounded in social networks – it “exists in the relations among persons” (Coleman 1988, p. S100) – and is facilitated by particular kinds of local structural conditions such as residential stability or the presence of formal organizations. In turn, collective efficacy is the ability of a community to draw on this resource to recognize a shared interest in the safety of the local area and to achieve specific tasks related to local informal social control. Recent research has provided evidence of a role for collective efficacy in explaining inter-neighborhood differences in crime rates. Important questions remain, however, about how respondents estimate their neighbor’s capacity for informal social control, the different ways people choose to intervene, connections to cultural dimensions of a neighborhood, and the links between formal and informal social control efforts.

Related Entries

- ▶ [Differential Social Organization](#)
- ▶ [Disadvantage, Disorganization and Crime](#)
- ▶ [Early Chicago School Theory](#)
- ▶ [Informal Social Control](#)
- ▶ [Neighborhood Effects and Social Networks](#)
- ▶ [Race and Ethnicity in Social Disorganization Theory](#)

Recommended Reading and References

- Anderson E (1999) *Code of the street: decency, violence and the moral life of the inner city*. W. Norton, New York
- Browning CR (2002) The span of collective efficacy: extending social disorganization theory to partner violence. *J Marriage Fam* 64:833–850
- Bursik RJ Jr, Grasmick HG (1993) *Neighborhoods and crime: the dimensions of effective community control*. Lexington, New York
- Carr PJ (2003) *Clean streets: controlling crime, maintaining order, and building community activism*. New York University Press, New York
- Cloward R, Ohlin L (1960) *Delinquency and opportunity: a theory of delinquent gangs*. Routledge, London
- Cohen AK (1955) *Delinquent boys: the culture of the gang*. Free Press, Glencoe
- Coleman JC (1988) Social capital in the creation of human capital. *Am J Sociol* 94:S95–S120
- Gans HJ (1962) *The urban villagers; group and class in the life of Italian-Americans*. Free Press, Glencoe
- Gibson CL, Zhao J, Lovrich NP, Gaffney MJ (2002) Social integration, individual perceptions of collective efficacy, and fear of crime in three cities. *Justice Q* 19:537–564
- Granovetter MS (1973) The strength of weak ties. *Am J Sociol* 78:1360–1380
- Kasarda JD, Janowitz M (1974) Community attachment in mass society. *Am Sociol Rev* 39:328–339
- Kornhauser RR (1978) *Social sources of delinquency*. University of Chicago Press, Chicago
- Kubrin CE, Weitzer R (2003) New directions in social disorganization theory. *J Res Crime Del* 40:374–402
- Miller WB (1958) Lower class culture as a generating milieu of gang delinquency. *J Soc Issue* 14:5–19
- Morenoff JD, Sampson RJ, Raudenbush SW (2001) Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence. *Criminology* 39:517–560
- Pattillo ME (1998) Sweet mothers and gangbangers: managing crime in a black middle-class neighborhood. *Soc Forces* 76:747–774
- Pfohl SJ (1985) *Images of deviance and social control*. McGraw-Hill, New York
- Sampson RJ (2002) Organized for what? Recasting theories of social (dis)organization. *Crime Soc Disorgan* 10:774–802
- Sampson RJ, Groves WB (1989) Community structure and crime: testing social-disorganization theory. *Am J Sociol* 94:774–802
- Sampson RJ, Raudenbush SW (1999) Systematic social observation of public spaces: a new look at disorder in urban neighborhoods. *Am J Sociol* 105:603–651
- Sampson RJ, Raudenbush SW, Earls F (1997) Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science* 277:918–924

- Sampson RJ, Morenoff JD, Earls F (1999) Beyond social capital: spatial dynamics of collective efficacy for children. *Am Sociol Rev* 64:633–660
- Shaw CR, McKay HD (1942) Juvenile delinquency in urban areas. University of Chicago Press, Chicago
- Small ML (2002) Culture, cohorts, and social organization theory: understanding local participation in a Latino housing project. *Am J Sociol* 108:1–54
- St. Jean PKB (2007) Pockets of crime: broken windows, collective efficacy, and the criminal point of view. University of Chicago Press, Chicago
- Sutherland EH (1947) Principles of criminology, 4th edn. Lippincott, Philadelphia
- Venkatesh SA (1997) The social organization of street gang activity in an urban ghetto. *Am J Sociol* 82–111
- Warner BD (2003) The role of attenuated culture in social disorganization theory. *Criminology* 41:73–98
- Warner BD (2007) Directly intervene or call the authorities? A study of forms of neighborhood social control within a social disorganization framework. *Criminology* 45:99–129
- Warner BD, Rountree PW (1997) Local social ties in a community and crime model: questioning the systemic nature of informal social control. *Soc Probl* 44:520–536
- Wright EM, Benson ML (2011) Clarifying the effects of neighborhood context on violence “behind closed doors”. *Justice Q* 28:775–798

behaviors of its members. These mechanisms take many shapes and sizes and are often classified by type and/or level. For example, social control is often divided into two types: informal and formal. Informal sanctions encompass such things as not engaging in an action so as not to disappoint one’s parents or because one believes the act to be wrong. Formal social controls involve more direct action against a person such as suspension from school or arrest and incarceration in the criminal justice system. These mechanisms also traverse various levels or domains of control ranging from the private to the public realm. The private level refers to those modes of control that are the most intimate or close to the person such as the control exhibited by family and friends and is the level most likely to contain informal measures. The next level, often referred to as the parochial or institutional level, involves those controls invoked by social institutions such as churches and schools. The public domain represents manners of control that are the most removed from the individual. This includes broader social connections between individuals, institutions, and the greater society and is often realized through formal legal sanctions. The levels of control are often complementary, interactive, and interdependent as well.

The following essay provides the reader with an introduction to the study of social control, formal theories of social control, and modern extensions and analytical techniques that have evolved to account for its multifaceted and multilevel nature. This research is grounded in two complementary sociological perspectives: social disorganization and social control. Originating in studies of the break down, or “disorganization” of communities, research in the first tradition looked to structural characteristics of a neighborhood that were associated with high crime rates. Later theorists then drew clear connections between the concepts of “disorganization” and “social control” clearly noting that the two concepts are but distinct ends of a continuum between high and low levels of social solidarity (Kornhauser 1978). When high, the community is able to exert both levels of informal and formal social control upon its members to achieve

Social Constructivism

► Transitional Justice

Social Control

Megan C. Kurlychek

School of Criminal Justice, University at Albany,
State University of New York, Albany, NY, USA

Synonyms

[Social bond](#)

Overview

Social control refers to the mechanisms through which a society is able to regulate and direct the

desired social outcomes. When low, on the other hand, individuals are left to their own devices.

From a social control perspective, this is not a good thing. That is, social control theorists do not begin with the question of why individuals commit crime, but rather why they do not (see, Hirschi 1969). The assumption being that left to our own devices, humans will always pursue pleasure and will do so in the easiest way possible. Since crime often offers high excitement and pleasure with less effort than traditional mechanisms (e.g., hard work), social control is then necessary to inhibit individuals from taking this path.

The following section reviews early studies of the organization, or rather the breakdown of the organization, of communities and their elaboration into formal theories of crime and delinquency. Connections between social disorganization and social control theories are described. Later sections then detail modern extensions of these theories and advances in scientific research that have informed the current state of knowledge and set forth pathways for future research.

Early Studies of Community Structure and Social Control

The 1996 publication of Hilary Clinton's *It Takes a Village: And Other Lessons Children Teach Us* brought national attention to the notion that the socialization of our nation's children is not just a family affair but involves all members of the local community. While this heightened attention to the crucial function of the community in the control and socialization of our nation's youth may seem somewhat novel, it is in fact nothing new to those who study societies and the mechanisms by which societies control their citizenry. Indeed, philosophers as far back as Aristotle noted the importance of community governance in maintaining social order and raising children. Beyond mere philosophy, a rich tradition of research and theory exists that details the important and complex relationship between community structure, interpersonal social relationships,

and the exertion of social control upon a community's residents, particularly its youth.

The formal study of social control in America can be traced back to studies of the growth of American cities such as Chicago around the turn of the twentieth century. This school of thought, often referred to as "The Chicago School," approached the question of social control from the opposite direction – that is, this research examined what happens when a community is not able to exert social control upon its members. W. I. Thomas and Florian Znaniecki provide some of the earliest empirical work in this area writing on the life experiences of immigrant families trying to "make it" in the new world (1927). Of the many struggles these families faced, one of the greatest was raising their children in a community setting that was unfamiliar to them. The traditions in which they were raised no longer existed, and new traditions had yet to be established in their place. Thus these immigrant parents were ill prepared to socialize their children to this new environment. As immigrant populations tended to settle together (e.g., Polish families with Polish families, Irish families with Irish families, etc.), these problems rose above the individual or family level to become a larger group and even community phenomenon.

With the breakdown in traditional parental socialization roles, this research also found peer networks to take on a heightened influence in the lives of children. This is further demonstrated in an influential writing by Thrasher, *The Gang* (1927), that presented a chronicle of the evolution of neighborhood childhood playgroups into more formal "gangs" – some of which were criminal, some of which were not. According to this study, one primary feature of delinquent gang formation was that it filled a space or crack in social solidarity. These cracks, referred to as interstitial areas, were seen as pockets in society in which the community could not exhibit social control. These were the areas in which gangs were then free to form. As described by Thrasher, the gang provided the essential human needs of belonging and camaraderie that the family and community had failed to provide.

In addition to these qualitative accounts of life in a growing city, researchers at the University of Chicago began the development and formal testing of theories of community growth and governance. The most prominent of these theories was social disorganization theory developed by Clifford Shaw and Henry McKay. Applying a model of community growth developed by colleagues Robert Park and Ernest Burgess, Shaw and McKay were able to document patterns of arrest in the city over a 30-year time period. During this time, Chicago was experiencing rapid population growth primarily due to large influxes of immigrants from Europe as well as southern states. The city, in turn, started to experience high levels of crime. While some tended to place blame on the latest group of immigrants – be they from Poland, Ireland, Italy, etc. – Shaw and McKay conducted a scientific study of the arrest of delinquent youths and found that the criminal activity was concentrated in a small area of the city encircling the central business district which became known as the “Transition Zone.” This area was in flux as neighborhoods that were once homes were swallowed up by business as the central business district (referred to in Chicago as the Loop) expanded. Housing in this area became relatively cheap as landowners stopped investing in properties they knew would soon be demolished. Because the rents in this area were cheap (and the housing in poor condition) it was often that the newest group of immigrants would first live in this area until they found jobs and gained solid footing in the “New World.”

By mapping out crime data for over 30 years, Shaw and McKay were able to show that the crime remained in this area of the city regardless of the population inhabiting the area. In addition, they found that when crime did “move” to a new area it was because that area was experiencing conditions similar to those neighborhoods where crime had previously been concentrated. Thus, this study found that “social disorganization” and the high crime rates that attend it are primarily characteristics of a place and not the people who inhabit that space.

The characteristics of the place most often associated with social disorganization were high

poverty, high residential mobility (people moving in and out quickly), and high racial heterogeneity. The first, poverty, was influential because there was little money to invest in improving the area. Homes would become dilapidated and community institutions such as schools, churches, and social organizations would break down. Second, because the neighborhood was “deteriorating” rather than being maintained, this is not an area where people desired to put down roots. Rather it was a stepping stone, an area where they would live briefly only until they could afford something better. Thus, once an individual or family acquired some wealth, they did not invest it in this area but rather they would move away leading to high residential mobility. Thirdly, as people came and went, especially during this time of high immigration, the area became populated by people with very different cultures, values, and even languages. This racial and ethnic heterogeneity lead to conflict and mistrust between groups as well as the inability of inhabitants to form basic social networks that help to hold a community together. Social Disorganization theory therefore posited that it was a combination of these three elements – poverty, residential mobility, and racial/ethnic heterogeneity – that lead to high levels of crime and delinquency in a given community.

Social Disorganization and Social Control: Making the Connection

While Shaw and McKay offered these basic renditions explaining why these structural factors might be of importance, it wasn’t until much later that this theory was formally recognized as a version of Social Control theories and further elaborated from a pure macro-level theory to an examination of the microlevel processes that occur within a community. In a classic 1978 writing, Ruth Kornhauser examined many existing theories of crime and delinquency detailing their primary assumptions and classifying them based upon these underlying assumptions. Through this process she was able to clearly demonstrate the inherent ties between social

disorganization and social control theories. According to Kornhauser, the structural factors related to disorganization of a community were indeed important predecessors of crime and delinquency *because* they lead to the disruption of the various components of both informal and formal social control.

In addition to highlighting the connections between social disorganization and social control, Kornhauser's work brought renewed attention to the social disorganization perspective by pointing out that the primary overlooked components of Shaw and McKay's work were the processes through which social disorganization at the neighborhood level is linked to individual-level delinquency. Not only are community structural characteristics important for understanding neighborhood crime levels. These ecological and social causes set the stage for a variety of undesirable consequences including weak social bonds developed between the individual and society, the inability of institutions to enforce common standards and goals, the defective socialization of children, the lack of continuous social networks, and, even the lack of the ability to take action to fix these problems. During the 1970s, 1980s, and 1990s, the search for a better understanding of the social processes by which social disorganization leads to reductions in social control became a fruitful branch of research for criminologists. The following section outlines some of the key knowledge gained during this time.

Extensions of a Theory: Systems of Social Control

John Kasarda and Morris Janowitz (1974) introduced what they coined the "systemic model" of a community control which envisioned the local community as a complex system of friendship and kinship networks. These networks were rooted primarily in family life and were the basis for the ongoing processes of socializing youth in the family and community. This research then tied these friendship and kinship networks to the social structural characteristic of residential mobility proposed by Shaw and McKay.

Specifically, the authors proposed that residential mobility is a temporal variable as the development of social bonds and networks takes time. The longer one lives in a community, the more people one knows and the more opportunity one has to become involved in community activities and institutions. When people are moving in and out of an area fairly rapidly, these basic mechanisms naturally break down. This then leads to a reduction in the social networks a person is able to develop. Fewer social networks in turn diminish the ability of the family to properly socialize children and to enforce social controls, informal and formal, upon youth.

Subsequent research has further documented the importance of social ties. For example, Coleman (1988) argued that it is the resources transmitted through social ties, not the ties per se that are key to facilitating neighborhood social control. Similar to the familiar notion of financial capital (e.g., one's financial resources), these social ties and networks became known as a person's "social capital." Research in this area has examined such things as the number of family members living close by, the number of neighbors one knows by sight and by name, family composition (e.g., single-parent families), and the number of moves a family makes as potential indicators of social capital.

Although research has evolved to measure and explain social capital in various ways, it is important to remember that its primary nature is that it is something not housed within an individual but rather within the structure of the community in which the individual lives. Research by Robert Sampson and colleagues has been highly influential in the development of this concept and has resulted in the development of several terms and measures that are now generally accepted as a standard in the social control literature. The first of these is Intergenerational Closure. Intergenerational Closure reaches beyond the immediate family to measure the extent to which children are linked to adults within the community. While this includes immediate family, it also includes the parents of friends, as well as other adult figures such as teachers, coaches, church officials, or even just other concerned resident.

A second concept, reciprocal exchange, moves beyond the mere existence of such relationships and places emphasis on the ability of community members to rely on each other to exchange items (e.g., borrow a tool), information (e.g., tryouts for a sport team or new community organization), and even advice such as turning to one another for advice on child rearing. This concept encompasses more than merely knowing other community members to incorporating levels of trust between community members.

A third concept that has been highly influential in research in this area is that of “collective efficacy.” According to research social capital may be a necessary, but not sufficient, aspect of neighborhood relations in the exhibition of social control. In addition to forming ties and resources in one’s community, collective efficacy then refers to the willingness and ability of community members to come together and take action on a specific task. One such important task is the socialization of youth. Collective efficacy could also be visualized as a community coming together to earn funds to build a park, start a youth soccer league or improve a library. A classic example often given is the creation of a neighborhood or block watch group in which community members take turns patrolling the neighborhood and reporting suspicious people or activity in a direct effort to reduce crime. Not surprisingly, empirical evidence shows strong relationships between high levels of collective efficacy in a community and low levels of crime. In addition, some research has suggested that factors such as friendship networks and participation in community associations matter *only* through their ability to promote collective efficacy among residents (Morenoff et al. 2001).

Another aspect of social relationships brought to light in this research was social cohesion. While this particular term has several various definitions in general it refers to the sharing of common goals and values. For example, it would be hard for a community to come together to achieve a common goal (e.g., collective efficacy), if members did not agree on the goal. Just as residential mobility was shown to influence crime through its

impact on social ties and networks, another social structural characteristic – racial and ethnic heterogeneity – appears to exert its influence mainly through its impact on another social interactive process. In this case, the effect is upon levels of social cohesion and the sharing of common beliefs and goals. Logically, the more different people are in a community the less likely they may be to share common goals and/or to be willing to come together and take action to achieve such goals.

Scholars note, however, that social cohesion can indeed go both ways when it comes to influencing crime and delinquency. That is, a family with strong cohesion with criminal members might then promote criminality in other members. As such, a community that accepts drug dealing, prostitution, or other types of deviance as normal can exhibit social “cohesion” in a fashion that promotes high levels of crime.

As the body of research on systemic social control continues to accumulate knowledge is gained about additional mediating processes and the reasons why they are important for the imposition of social control. For example, in addition to the factors noted above, variables such as family disruption, percent of single-parent families, sparse local friendship networks, mixed land use, population density, and the presence of unsupervised teenage peer groups all have been found to be indicators of decreased social control and, in turn, to lead to increased levels of crime and delinquency. Thus, the systemic model of social control informs us that the social structural conditions identified by Shaw and McKay as indicators of social disorganization do not directly cause crime but rather are related to high levels of crime and delinquency *because* they disrupt the more microlevel social process of human interaction that are necessary for a community to exert social control upon its residents.

Levels of Social Control

At the same time as the systemic model of social control was being developed and elaborated, complementary research was being conducted that elaborated upon the various levels of social

control. Research by Albert Hunter (1985), and later Robert Bursik and Harold Grasmick (1993), detailed three distinct levels of control: private, parochial, and public. While these levels have been hinted at by the previous research, this section will more specifically identify the levels and their relationships.

The first level, private social control, is primarily informal in nature and is evoked through personal relationships. This level thus relates closely to the measures of social ties, networks, and cohesion noted above while adding in aspects of emotional attachments as well. For example, according to Travis Hirschi's original Social Bond Theory (1969), what stops an individual from committing delinquent acts are his/her bonds to society. These bonds take the form of personal attachments to people, commitment to conventional goals, involvement in conventional activities and the adoption of pro-social beliefs. Thus, simply knowing that a behavior will bring a negative response by a friend or family member or may negatively impact the ability of one to reach a goal exhibits enough "control" to stop the individual from engaging in the behavior.

The parochial level then incorporates the role of local social institutions in a community such as schools, churches, and community groups. This level of social control relates closely to the institutional controls noted in previous research and the importance of these institutions for the socialization of children. As children age such institutions take on increased importance in their lives and therefore become primary sources of socialization. The parochial level of social control may reinforce that exhibited at the private level, or in some cases, may protect the youth from further problems if the private level is not sufficiently developed. For example, a youth who is not highly attached to a parent in the home may form a strong bond with a coach or a teacher. That youth would then be less likely to break the rules than one who had no social control at either the private or parochial level. Thus, when the informal controls traditionally exhibited at the private level break down, the parochial or institutional level of social control may then

serve as a second line of defense to prevent anti-social behavior.

The third level, public social control, was described by Albert Hunter (1985) as the ability of the community to access public goods and services to maintain order. This included but was not limited to police services. At this level the focus is either on the development of programs to create order and prevent crime or to catch those who are breaking the law and invoke enough formal control upon them to discourage them from continuing to break the law. At this level community financial resources and ties to greater society and social resources are of central importance.

It is easy to see the interdependent nature of these levels. If private social controls are strong, there is less need for parochial and public controls. On the other hand when private and parochial levels of control are weak, a greater use of public social control would be warranted (Black 1976). For example, an increased use of police force in an area is actually seen as a breakdown in private and parochial levels of social control. If all three levels are lacking, however, an area will be ridden with high rates of crime and violence. Research also suggests the potential for these levels to interact with one another. That is, for the effect of low social control at one level to have varying effects upon individuals based on other levels of risk or social control.

Moreover, it is important to note that this multilevel model of social control is not distinct from the systemic model but rather a further elaboration of the dynamics of social processes used to invoke social control. For example, levels of private social control in a neighborhood are determined by the ties between individuals and the strength of families. At the parochial level the strength of institutions and their ability to socialize youth is dependent on community member support both in terms of finances and participation. Research has further suggested that the social processes leading to the determination of private and parochial social controls are inherently tied to the larger structural characteristics originally identified by Shaw and McKay.

Advances in Theory, Testing, and Techniques

As highlighted by the previous discussion, social control is exhibited not only at distinct levels but through intricate interpersonal processes that are linked to the larger social institutions and structural characteristics of a community. Thus, the study of social control has become more complex in nature as have the theories that explain it.

One particular direction of theory development has been the integration of theories to address the various causes of crime and levels of social control. One example is Terence Thornberry's Interactional Theory that connects elements of social control and social learning theories as well as incorporating elements of the social structure as the interactional setting in which delinquent behavior is learned and even reinforced (Thornberry 1987). This theory also holds that different factors command primary influence at different stages of life. For example, while the family is key in early childhood through early adolescence, by mid-adolescence the school and other institutions outside of the family become more prominent with peer groups and the broad community taking dominance by late adolescence. Thus, according to this theory, the level of social control (private, parochial, public) bearing the strongest relationships to delinquency would change over the lifecourse.

John Laub and Robert Sampson (1993) also look at social control over different points in the life course and offer some of the first theories to extend this examination into adulthood noting what they call "adult social bonds." The institutions and social bonds that they incorporate as mechanisms of social control include marriage and employment as these offer new mechanisms of informal control upon the adult individual. For example, in adulthood the introduction of a spouse adds a new type of private social control while employment would expand connections to both parochial and public levels of social control.

Another fruitful theoretical avenue has been the risk-protective factor approach which seeks to identify factors at the individual, institutional, and neighborhood level that increase or decrease the

possibility of deviant outcomes. Research in this area has tended to emphasize those factors that increase deviance such as previously noted correlates of delinquency (e.g., broken families, low attachment to schools). Analyses at the neighborhood level have also expanded to include such new things as number of gangs, availability of weapons, and ease of access to drugs.

Research utilizing this approach has been highly informative in gaining a better understanding of the interdependent and interactive nature of the multiple levels of social control. For example, one study asks "Do disadvantaged neighborhoods cause well-adjusted children to become adolescent delinquents?" (see Wikstrom and Loeber 2000). The answer here generally appears to be "No." However, some neighborhood enticements were found to lead some well-adjusted youth into delinquency particularly in later adolescence when influences outside of the home gain dominance. Other studies approach the question from the opposite direction asking whether "good" communities can help "not so well-adjusted" youth avoid negative outcomes (see Sampson et al. 2002; Kurlychek et al. 2011). In sum the basic consensus of these studies appears to be yes, that is, at-risk children fair better when raised in communities with high levels of parochial social control (e.g., strong social institutions and connections/cohesions among residents).

As theories of social control and crime become more complex, so do the data sets and analytical techniques needed to test them. For example, if social control spans three levels, then so must the data collected to explore it. The compilation of such data sets is time consuming and expensive thereby prohibiting a great deal of research in this area. However, several recent data collection efforts have evolved that are specifically designed to advance our understanding of the multilevel nature of social control. One such effort, the Project on Human Development in Chicago Neighborhoods (PHDCN), was specifically designed to examine how families, schools, and neighborhoods combine to affect child and adolescent outcomes. As an example of just how complicated such a data collection

process can be it should be noted that this project has been ongoing for since 1994 and has received funds from a variety of entities including the National Institute of Justice, the John D. and Catherine T. MacArthur Foundation, the US Department of Health, the US Department of Education, the National Institute of Mental Health the Harris Foundation, and the Turner Foundation.

In addition to the collection of data, the analyses of the data has been plagued by complexities such as the identification of compositional versus contextual effects and the spatial dependence of data as individuals are nested within schools and communities. Fortunately, recent developments in statistical techniques and packages have allowed researchers to effectively tackle many of these issues.

Perhaps the most significant advance in this area has been the use of hierarchical linear modeling (HLM) techniques to enable researchers to further disentangle the impact of individual-, institutional-, and neighborhood-level effects as well as exploring how factors at the various levels interact with one another. In addition this technique allows the researcher to identify what are known as "contextual" effects, which are effects above and beyond the mere aggregation of individual factors. For example, if being from a single-parent family increases the risk of delinquency by 5 % and one community has 10 % single-parent homes and another 50 %, the neighborhood with 50 % single-parent homes would naturally be expected to have more delinquency because more youth exhibit the 5 % greater risk of delinquency. However, a contextual effect would emerge if being in a community with more single-parent homes actually increased the risk of each youth, for example, instead of a 5 % greater risk of delinquency, it might be a 10 % greater risk of delinquency. Existing literature suggests that a pattern such as this might emerge as the percent of single-parent homes impacts other things such as the presence of adult role models, the ability of adults to supervise their children, and even the ability of parents to participate in community institutions/associations.

Examples of findings using this technique are abundant in the criminological literature. For example, research has found that the concentration of social disadvantage in neighborhoods has an independent contextual effect on adolescents' serious offending behavior even after controlling for multiple measures of individual-level risk. Studies have also found that adolescents residing in more economically disadvantaged neighborhoods were more likely to be involved in a range of serious violent behaviors after controlling for levels of individual-level risk. On the other hand, research has begun to show that youth with high levels of personal risk actually have lower chances of becoming delinquent if they reside in neighborhoods with strong social institutions and high levels of collective efficacy.

Future Directions

Although much knowledge has been gained about the role of social control, its types and levels in the causation of crime and delinquency, much is left to learn about the interactive and interdependent nature of social control. In particular, as noted above, research has just begun to show that living in an area with high levels of parochial social control can actually reduce the risk of a delinquent outcome for even high-risk youth. Research further exploring these connections holds the potential to inform crime prevention and intervention policy. For example, we might know that a youth is at high personal risk for delinquency because of family turmoil or other issues, but there may be ways to still reduce the overall probability of a criminal outcome for this youth by increasing other social controls such as attachments to school, involvement in community groups, and improving overall community levels of cohesion and collective efficacy. Or, working in the opposite direction, for youths living in high-risk communities, this research informs what might be done at the private/informal level to improve outcomes.

To answer questions such as these, there is also a need for more systematic social

observation which better captures both the notion of a neighborhood and the real “feel” of living in that community. According to experts there are simply too many features of a neighborhood that are not measured readily by available public data sets (e.g., census) or cannot be systematically captured in surveys. To gain this type of information more observational and qualitative research is needed. Key researchers in this area stress the importance of collecting this information at the smallest unit possible (e.g., a block face – one side of a street within a block) (Sampson et al. 2002). This allows the researcher to then study these small geographic units, as well as providing the capacity to aggregate these units into larger geographical constructs (e.g., block groups, neighborhoods).

In conducting future research it is also important to remember that no community is an island. That is, each geographic area may be impacted by those areas surrounding it as well. As one example, when studying an adolescent it may be important to consider not only his/her home neighborhood but also the neighborhoods through which he/she travels to attend school or to which he/she travels to spend time with friends. It may also be important to understand concentrated disadvantage by not only looking at one community, but to also incorporate the resources of those communities that border it. One of the few studies in this area suggests that the consequences of being poor are worse for African-Americans in our society because not only do these individuals live in poor neighborhoods but they are much more likely than poor white individuals to live in communities embedded in areas of sustained poverty and disadvantage (Pattillo-McCoy 1999).

In summary, while we know much about social control and its importance in preventing crime, there is still much more to learn about the complex and dynamic processes through which it impacts individuals and the communities in which they live. Future research in this area holds the promise to help build stronger families, safer communities, and provide better life outcomes for even the most at-risk youth.

Recommended Reading and References

- Black D (1976) *The behavior of law*. Emerald Group Publishing, Bingley
- Bursik RJ, Grasmick HG (1993) *Neighborhoods and crime: the dimensions of effective community control*. Lexington Books, New York
- Coleman JS (1988) Social capital in the creation of human capital. *Am J Sociol* 94(Suppl):S95–S120
- Hirschi T (1969) *Causes of delinquency*. University of California Press, Berkeley
- Hunter AJ (1985) Private, parochial and public school orders: the problem of crime and incivility in urban communities. In: Suttles GD, Zald MN (eds) *The challenge of social control: citizenship and institution building in modern society*. Albex, Norwood
- Kasarda JD, Janowitz M (1974) Community attachment in mass society. *Am Sociol Rev* 39:328–339
- Kornhauser R (1978) *Social sources of delinquency*. University of Chicago Press, Chicago
- Kurlychek M, Krohn MD, Beidi D, Penly-Hall G, Lizotte AT (2011) Protection from risk: an exploration of when and how neighborhood-level factors can reduce violent youth outcomes. *Youth Violence Juv Justice* 10(1):83–106
- Morenoff JD, Sampson RJ, Raudenbush SW (2001) Neighborhood inequality, collective efficacy and the spatial dynamics of urban violence. *Criminology* 39(3):517–559
- Park RE, Burgess EW, McKenzie RD (1925) *The growth of the city*. University of Chicago Press, Chicago
- Pattillo-McCoy M (1999) *Black picket fences: privilege and peril among the black middle class*. University of Chicago Press, Chicago
- Rankin BH, Quane JM (2002) Social contexts and urban adolescent outcomes: the interrelated effects of neighborhoods, families and peers on African-American youth. *Soc Probl* 49(1):79–100
- Sampson RJ, Byron Groves W (1989) Community structure and crime: testing social-disorganization theory. *Am J Sociol* 94(4):774–802
- Sampson RJ, Laub JH (1993) *Crime in the making: pathways and turning points through life*. Harvard University Press, Cambridge
- Sampson RJ, Morenoff JD, Earls F (1999) Beyond social capital: spatial dynamics of collective efficacy for children. *Am Sociol Rev* 64:633–660
- Sampson RJ, Morenoff JD, Gannon-Rowley T (2002) Assessing “neighborhood effects”: social processes and new directions in research. *Annu Rev Sociol* 28:443–478
- Shaw CR, McKay HD (1942) *Juvenile delinquency and urban areas*. University of Chicago Press, Chicago
- Taylor RB (1997) Social order and disorder of street blocks and neighborhoods: ecology, microecology, and the systemic model of social disorganization. *J Res Crime Delinq* 34(1):113–155
- Thomas WI, Znaniecki F (1920) *The Polish peasant in Europe and America*. Gorham, Boston

Thornberry TP (1987) Toward an interactional theory of delinquency. *Criminology* 25(4):863–891

Thrasher F (1927) *The gang*. University of Chicago Press, Chicago

Wikstrom P-OH, Rolf L (2000) Do disadvantaged neighborhoods cause well-adjusted children to become adolescent delinquent? A study of male juvenile serious offending, individual risk and protective factors and neighborhood context. *Criminology* 38(4):1109–1142

Social Control and Self-control Through the Life Course

Michael L. Benson

School of Criminal Justice, University of Cincinnati, Cincinnati, OH, USA

Overview

The concept of social control has a complex and controversial history in the field of sociology. Originally, it was defined as the ability of a group to regulate itself, but the term was subsequently redefined to refer either to socialization or social repression (Janowitz 1975). Within criminology, social control typically is used in the more classical sense to refer to the mechanisms through which groups attempt to orchestrate behavior and control deviance. Self-control, on the other hand, is a relatively new concept in criminology that refers to the differential ability (or inability) of individuals to refrain from taking advantage of opportunities to satisfy their immediate desires by engaging in criminal or deviant behavior (Gottfredson and Hirschi 1990). Both social control and self-control have been proposed as important factors that influence individual involvement in crime throughout the life course (Gottfredson and Hirschi 1990; Sampson and Laub 1993). Empirical support for their importance in the etiology, longevity, and patterning of criminal behavior is substantial, but a number of issues remain to be resolved, including the origins and stability of self-control, the effects of social control on criminal trajectories at different stages of the life course, and the factors

that influence a person's exposure to informal social control. This entry addresses these issues by focusing on how social control and self-control may change over the life course.

Crime and Informal Social Control

Within criminology, informal social controls are seen as emerging from “the role reciprocities and structure of interpersonal bonds linking members of a society to one another and to wider social institutions such as work, family, and school” (Sampson and Laub 1993, p. 18). For example, husbands and wives are bonded to one another through socially recognized obligations and expectations regarding proper normative behavior. These interpersonal bonds act as constraints on the participants in a role relationship, in the sense that there is pressure on participants to meet their role expectations if they wish for the relationship to continue and if they wish to be viewed favorably by others. Children may be bonded to their parents by emotional connections of love and caring, and children understand that their parents have expectations regarding how they should behave. To the extent that children care about their parents, these expectations can act as a constraining force on their behavior. Similarly, being employed requires that one meet the expectations of one's employer to show up at work, do one's job, and not bring disrepute on the work organization. Thus, informal social controls that arise out of social roles guide behavior by placing responsibilities and constraints on the individual and by directing individual action in some directions rather than others. Prosocial behavior is promoted by strong social controls, while deviancy and delinquency are promoted by the attenuation of social controls over individual conduct (Thornberry 1987).

At different stages in the life course, individuals are potentially subject to different forms of informal social control that arise out of different role relationships and participation in different social institutions (Sampson and Laub 1993; Thornberry 1987). For children, informal family and school bonds are important. Children who are strongly bonded to their parents and who care about school are less likely to be involved in

delinquency than children who have difficult relations with their parents or who do not like school. As children move through the life course, the major sources of informal social control change. Parents and school are not as important for young adults as they are for children and teenagers. For young adults, employment and marriage are potential sources of informal control. Some scholars argue that variation in the strength of informal controls influences the likelihood and degree of involvement in crime and deviance at all stages of the life course (Sampson and Laub 1993; Thornberry 1987).

For young children, families and specifically parents are the most important source of informal social control. The fundamental processes that seem to be present in successful families and missing in unsuccessful ones are attachment and control. Attachment refers to an emotional connection between parent and child, in which the child has feelings of love, respect, and admiration toward the parent and the parent feels similarly toward the child. It is a reciprocal process, involving both parent and child, but it starts with the parent (Hirschi 1969). By giving love and expressive support early on, parents can foster a sense of attachment in their children. Children who develop strong attachments to their parents care about their parents' feelings and opinions. They are aware of and sensitive to the impact that their behavior can have on their parents. They understand that if they are caught doing something wrong, it will embarrass and disappoint their parents, and they do not want that to happen. Rather, they want their parents to be proud of them. This emotional connection functions as a sort of internal monitor of the child's behavior when parents are not present. Parents who work at developing a strong sense of attachment in their children when they are young are exercising informal social control and are likely to be rewarded for their efforts when their children enter adolescence in reduced levels of delinquency.

Showering children with love and affection when they are young in the hopes of developing a strong sense of attachment, however, is not enough in and of itself to prevent involvement

in delinquency or deviance. Expressive support is crucial, but parents who want to keep their children out of trouble must also exercise more direct types of informal social control. Direct control involves monitoring the child's behavior, recognizing deviance when it occurs, and correcting misbehavior when it happens (Gottfredson and Hirschi 1990). Parents must pay attention to their children and be aware of what they are doing and recognize when they are doing something that is wrong or inappropriate. By monitoring their children, recognizing deviance, and correcting misbehavior, parents can foster the development of conformity in their children and help start them off on trajectories aimed away from serious delinquency (Gottfredson and Hirschi 1990; Patterson et al. 1992).

As children grow older and move into adolescence, the sources of informal social control change from families and parents to schools and peers. Children who are strongly attached to school are significantly less involved in deviance than those with weaker attachments. School attachment is a multidimensional concept that has both objective and subjective elements. Objective attachment can be seen in school performance (that is, how well one does in the classroom) and involvement in school-related activities, such as committees, sports teams, or school clubs. Subjective attachment refers to one's aspirations and expectations in regard to educational achievement and one's sense of satisfaction with and affection toward the people and activities that constitute the school system (Sampson and Laub 1993).

That people often behave differently when they are in a group as opposed to being by themselves is so well known that it is something of a sociological and psychological truism. This truism appears to apply with particular force to juveniles, as most delinquency is committed in groups (Warr 2002). Thus, peers have long been recognized as an important influence on juvenile delinquency and as a source of informal social control in adolescence in the form of peer pressure. Juveniles who have delinquent friends tend to be delinquent themselves, while their nondelinquent counterparts tend to have

nondelinquent friends. Although the empirical correlation between delinquency and having delinquent friends is well established, its interpretation is a matter of dispute. One school of thought holds that delinquent friends cause or exacerbate an individual's own delinquency through modeling and social reinforcement for delinquent behavior (Akers 1998). In contrast, conforming friends model and reinforce prosocial behavior. The other school of thought holds that the correlation between delinquency and delinquent friends is spurious and results from a self-selection process in which teenagers with delinquent propensities tend to seek out other teenagers with similar propensities (Gottfredson and Hirschi 1990).

As teenagers move into young adulthood, they may become subject to new forms of informal social control. These new forms of control include employment and marriage. Individuals who are lucky enough to find good jobs or enter good marriages or both become subjugated to new sources of informal social control. One of the most important sources is marriage. Marriage entails new responsibilities and ideally a strong emotional commitment to another person (Laub et al. 1998). To the extent that a young man is emotionally attached to his wife and his children, he is likely to curb his criminal inclinations out of respect for his wife's wishes and a desire not to jeopardize his family life. Marriage may influence involvement in crime in another way by limiting a person's access to deviant peers and to opportunities to engage in deviant activities. The time that a man spends with his spouse or children at home is time that he cannot spend with criminal peers on the street. Time spent working to provide a living for one's spouse and children is also time taken away from the risky attractions of street life. Getting married reduces one's contact with friends in general and delinquent friends in particular. Reduced exposure to delinquent friends leads to reduced involvement in crime and deviance (Warr 1998).

Qualitative and quantitative evidence suggests that, for some men, becoming involved with a woman and getting married can be a route out of crime (Shover 1985). What appears to matter

most is the development of high-quality marital bonds (Sampson and Laub 1993; Sampson et al. 2006). However, the effect of a good marriage on criminal offending does not occur all at once. Rather, it is a gradual process that cumulates over time. The longer time an offender invests in a good marriage, the more likely that marriage is to have a preventive effect on his criminal trajectory (Laub et al. 1998). Offenders who enter into good marriages gradually become committed to their spouses. This commitment appears to function as a source of informal social control that grows slowly over time and leads toward desistance.

Crime and Self-control

According to Gottfredson and Hirschi (1990), at the individual level, the most important causal factor in crime is a behavioral propensity called low self-control. Low self-control refers to the relative ability (or inability) of an individual to avoid taking advantage of an opportunity to satisfy his or her immediate wants and desires by engaging in some form of criminal or deviant behavior. Other monikers for this concept include self-regulation, effortful control, self-discipline, and will power. People with low self-control are thought to be impulsive, self-centered, not concerned with long-term consequences, physically active, and risk-loving. Self-control is conceived to be an individual trait that varies over individuals but remains stable for any given person over time (Gottfredson and Hirschi 1990). Thus, people who have low self-control relative to others early in life will also have low self-control relative to others later in life, while conversely people with high self-control at one point in time will tend to have high self-control at other times. People with low self-control can be distinguished from people with high self-control by their tendency to pursue "short-term gratification with little consideration for the long term consequences of their acts" and with little sympathy for the rights or feelings of others (Hirschi and Gottfredson 1987a, p. 959). Low self-control is not always manifested in criminal behavior. Depending on the situation, it may be expressed in other deviant, risky, or disreputable

ways, such as reckless driving, alcohol and drug abuse, promiscuous sex, and job quitting (Hirschi and Gottfredson 1987a, b). In short, people with low self-control are impulsive, insensitive to others, and almost always interested in pursuing their own personal pleasures in the quickest way possible. They are attracted to crime, deviance, and analogous acts because they tend to provide quick rewards and easy gratification.

Where does low self-control come from? According to Gottfredson and Hirschi (1990), the answer to this question can be found in the homes and families of delinquents. The parents of delinquents are responsible for the development of low self-control. In their theory, self-control is something that parents have to instill in their children early on by engaging in specific parenting practices. If parents fail to engage in these practices, then their children will not develop a strong sense of self-control, or to put it the other way round, then the children will inevitably have low self-control. In other words, self-control is not something that people are born with or that develops naturally as they grow up. Rather, it is something that must be instilled in children by their parents. To instill self-control, parents must do four things in regard to their children. They must (1) care for their children, (2) monitor their behavior, (3) recognize their wrongdoing, and (4) discipline or correct their wrongdoing when they observe it. If parents fail to employ this suite of practices, then their children will naturally develop with low self-control, and they will do so very early in life, typically before ages 11 or 12. After that, a person's relative level of self-control is set and cannot be developed or augmented later in life.

Current Issues and Controversies

Informal Social Control and Desistance

The causal role of informal social controls in the desistance process is disputed. Some scholars who take a life course perspective believe that change is an ever present possibility and that even serious long-term offenders can and do desist at later stages in the life course (Maruna 2001;

Maruna et al. 2009; Vesey et al. 2009). From this perspective, desistance is seen as a process that involves several mutually interacting factors, including turning points, informal controls, and the structuring effects of routine activities. In theory, the path to desistance begins when an individual experiences some sort of turning point, such as marriage, military service, moving to a new location, or finding a fulfilling job. Regardless what it is, this event sets in motion a series of changes in the offender's life that may over time lead to "desistance by default." Desistance by default expresses the idea that offenders do not necessarily make explicit and conscious decisions to change their lives and "go straight" as a result of some sort of epiphany or a single transformative event. Rather, their behavior and sense of identity change gradually and perhaps largely unconsciously as a result of experiencing a turning point that exposes them to new sources of informal social control and that also changes their routine activities.

Getting married, joining the military, or finding a job can serve to knife the offender off from his previous life style and expose him or her to greater informal social controls. Turning points also impose the structuring effects of conventional routine activities on offenders. Military service, marriage, and employment, all require time commitments. They require that one be at certain places at certain times. They impose duties to fulfill and obligations to keep. In short, individuals become caught up in a series of routine activities in which they spend time with conventional others doing conventional ordinary things. The time available to hang out with criminal peers and engage in deviant activities is gradually reduced until it reaches the point where an offender has for all intents and purposes desisted.

Another school of thought, however, holds that desistance is not caused by turning points and increased exposure to informal social controls. This school of thought, called the developmental perspective, holds that there are distinct developmental trajectories that people get locked into. Developmental trajectories are stable in the sense that once people enter a particular

trajectory or developmental group, their future progress is more or less determined. For example, Gottfredson and Hirschi (1990) argue that criminal behavior results from an underlying behavioral propensity. At an early point in the life course, each individual's propensity toward crime and deviance is established when their level of self-control is developed. From that point on, one's level of self-control drives behavior independent of other factors and is largely resistant to change. Thus, the presence or absence of desistance in adolescence reflects only variation in this underlying propensity toward crime. Another developmental theorist, Moffitt (1993), argues that there are two distinct trajectories – life-course-persistent and adolescent limited. Individuals on the life-course-persistent trajectory are unlikely to experience turning points such as marriage, or to react badly if they do experience them, because of certain innate characteristics that are established early in life. These individuals never really desist from deviant or criminal behavior, even though the form in which their deviance is manifested may change as they age. For example, the teenage mugger becomes the adult wife-beater. Adolescent limited offenders engage in crime and delinquency for a few years during their teenage years and then desist largely for maturational reasons, not because they are exposed to new forms of informal social controls.

Developmental theorists do not deny that there is an association between exposure to informal social controls and reduced involvement in crime and deviance, but they believe that this association is spurious and not causal. Reduced involvement in crime and deviance may be associated with participation in social institutions such as marriage and work, but it is not caused by them. Rather, participation in these institutions is a matter of self-selection. It is, of course, possible and some research suggests that both social causation and self-selection are involved in the desistance process (Wright et al. 1999).

Origins of Self-control

Even though low self-control is an individual trait that would appear to bear strong similarities to

other personality traits, such as impulsiveness, aggressiveness, and insensitivity, the originators of the term explicitly rejected the notion that it is biologically or genetically based (Gottfredson and Hirschi 1990). Rather, in their view, low self-control is caused solely by inadequate parenting early in a child's life. Recent advances in behavioral genetics suggest that this conclusion is either not correct or greatly oversimplified. Insofar as genetic researchers have been able to determine, it appears that all individual traits are influenced to some degree by heredity, that is, by the genes that parents pass on to their children (Rutter 2007). Height, weight, body type, eye color, hair color, and facial appearance are some of the more obvious physical traits for which it often is easy to see physical resemblances between parents and their offspring. However, in addition to these gross anatomical characteristics, children also inherit psychological characteristics and behavioral tendencies from their parents (Plomin et al. 1990). For example, aggressive parents are more likely to have aggressive children than nonaggressive parents (Huesmann et al. 1984). A growing body of evidence now suggests that genetic factors also are implicated in the development of low self-control (Beaver 2011; Wright et al. 2008a; Wright and Beaver 2005).

Genetic factors may influence the development of low self-control in several different ways. First, low self-control appears to be a heritable trait or characteristic that parents pass on to their children. Thus, parental genotype directly influences offspring genotype, including the genes related to low self-control. Second, parents with low self-control tend to create home environments that are not conducive to the development of high self-control (Moffitt 1997). Indeed, their parental management techniques may be so inadequate that they actually foster the development of low self-control (Beaver 2011). In this case, the genetic effect is indirect in that the parent's genotype is implicated in the creation of a particular home environment that then stimulates the development of low self-control in the child. Finally, children born with genetically influenced

personality characteristics that make them difficult to deal with, such as lack of impulse control, aggressiveness, and resistance to parental authority, may provoke negative parenting practices that further exacerbate behavioral tendencies (O'Connor et al. 1998). This type of effect is an evocative gene environment interaction, meaning that the child's genotype in a sense creates an environment that fosters the expression of behavioral tendencies inherent in the genotype.

Stability of Self-Control

One of the most provocative and controversial aspects of low self-control theory concerns the stability of this characteristic over the life course. As noted above, the originators of the term contended that relative differences between people in self-control remain stable as people age. Note, this contention does not mean that people with low self-control never change the way they behave as they age. Rather the idea of relative stability means that even though persons with low self-control may exhibit more moderate behavior as older adults than they did as teenagers, they will still be more antisocial, self-centered, and short-sighted than their contemporaries who have adequate levels of self-control. Indeed, research does indicate that absolute levels of self-control increase with age for most people.

At this point in time, however, research on the relative stability hypothesis has produced mixed results. While it is clear that there is substantial continuity in antisocial behavior over the life course, it is also clear that the level of continuity is by no means perfect. Evidence for stability comes from a number of longitudinal studies. For example, in a study of 205 boys aged 10–16, researchers asked the mothers to recall how easy or difficult it had been to get along with their sons when they were 1–5 years old. Five years later, when the boys were 15–21 years old, those who had been characterized by their mothers as “difficult” were twice as likely to have an official record of delinquency as those rated “easy.” The difficult boys also self-reported committing delinquent acts at a higher rate than the easy boys (Loeber et al. 1991). Similarly, kindergarten children who are rated by their

teachers as having high levels of hyperactivity are more likely to engage in delinquency during the transition from childhood to adolescence (ages 10–13) than children who do not exhibit hyperactivity (Tremblay et al. 2003). Finally, in a study using the Dunedin, New Zealand, data, White et al. (1990) found considerable continuity in antisocial and delinquent behavior from age 3 through the early teen years. Children who scored high on measures of disobedient and aggressive behavior at age 3 were more likely to exhibit other conduct disorders later in childhood and to be arrested by the police in their early teen years than children who scored low in disobedience and aggression (White et al. 1990). The fact that some children begin to show signs of abnormal conduct so early in life suggests that some forms of antisocial behavior reflect a general temperament that may persist over time (White et al. 1990). This study provides impressive evidence for continuity in antisocial behavior because it is a prospective study of a normal population (Wright et al. 2008b).

Thus, evidence for continuity in antisocial behavior is plentiful, but nevertheless, it must be interpreted carefully (Thornberry and Krohn 2003). Conduct problems at a young age do predict delinquency in the teenage years, but the success rate of these predictions is sometimes not very good. Most antisocial preschoolers do not go on to become antisocial adults or even antisocial juveniles. In the Dunedin study, for example, there was a high false-positive rate when antisocial behavior at age 3 alone was used as a predictor of antisocial behavior at age 11. Indeed, the predictions were wrong almost 9 times out of 10. The researchers note: “Of the 209 children predicted to have antisocial outcomes at age 11, 84.7 % did not develop stable and pervasive antisocial behavior” (White et al. 1990, p. 521). Other studies have had better results in the sense that the prediction rates are better. For example, Campbell and colleagues studied children from age 3 to age 13 with assessments at ages 4, 6, and 9 in between. Of the children who were rated as hard to manage by their parents at age 3, fully 48 % met formal criteria for having externalizing disorders (attention deficit, conduct

disorder, or oppositional disorder) by age 9. In contrast, only 16 % of the children in a control group displayed externalizing disorders at age 9 (Campbell 1995). In statistical terms, the difference between 48 % and 16 % is large and represents a strong effect. However, in practical terms, one could say that the study actually shows that discontinuity is more likely than continuity as less than half of the problem group exhibited problem behaviors at both times. At this point, the most accurate summary of the research would be that continuity in behavior is indeed a widespread feature of human development, but it is not a universal feature.

Although research on continuity in antisocial behavior suggests that traits such as low self-control may be stable over the life course, they do not directly assess that stability. To date, there are few studies that directly investigate stability in self-control itself over time. The few studies that are available, however, suggest that like antisocial behavior, relative levels of self-control can vary over the life course. That is, instability in individual rankings on self-control over time has been observed, suggesting that self-control may not be an immutable trait. The degree to which self-control is malleable, however, is still open to debate with some researchers finding more stability than others (Burt et al. 2006; Hay and Forrest 2006).

Future Directions

A large body of research now clearly indicates that both social control and self-control are important factors in the etiology of crime and behavior in general across the life course. Furthermore, it is becoming increasingly apparent that within the life course, self-control and social control have interactive effects on each other. For example, an individual's level of self-control may influence how he or she reacts when confronted with a turning point, such as marriage, which would in turn affect the individual's relative exposure to informal social controls. Likewise, it is possible that exposure to increased informal social controls may eventually lead to transformations in the individual that increase his or her level of self-control

(Giordano et al. 2002; Laub and Sampson 2003; Sampson et al. 2006). Future research on self-control and selection effects is needed, as well as, research on the transformative power of informal social controls in regards to developing increased self-control in adulthood.

Research on the stability of self-control over the life course has largely ignored whether self-control is malleable in response to direct officially imposed rehabilitative or treatment interventions. This shortcoming is unfortunate, because, as Currie has noted, research showing continuity in antisocial behavior over time reveals only what happens when no effort is made to change how individuals develop (Currie 1998). Yet, a voluminous body of research suggests that a multitude of rehabilitative programs work in terms of reducing recidivism among offenders at all stages of the life course (MacKenzie 2006). If delinquent and criminal behaviors are an expression of low self-control, then it follows by implication that reductions in offending caused by treatment interventions may also be associated with changes, indeed increases, in self-control. Future research is needed that directly assesses whether and how measured levels of self-control can be changed in response to different forms of intervention and treatment.

Related Entries

- ▶ [Control Theory](#)
- ▶ [Genes, Crime, and Antisocial Behaviors](#)
- ▶ [Genetic Basis to Self-Control](#)
- ▶ [Informal Social Control](#)
- ▶ [Pathways to Delinquency](#)
- ▶ [Risk factors for Adolescent Sexual Offending](#)

Recommended Reading and References

- Akers R (1998) Social learning and social structure: a general theory of crime and deviance. Northeastern University Press, Boston
- Beaver KM (2011) The effects of genetics, the environment, and low self-control on perceived maternal and paternal socialization: results from a longitudinal sample of twins. *J Quant Criminol* 27:85–105

- Burt CH, Simons RL, Simons LG (2006) A longitudinal test of the effects of parenting and the stability of self-control: negative evidence for the general theory of crime. *Criminology* 44:353–396
- Campbell SB (1995) Behavior problems in preschool children: a review of recent research. *J Child Psychol Psychiatry* 36:113–149
- Currie E (1998) *Crime and punishment in America*. Metropolitan Books, New York
- Giordano PC, Cernkovich SA, Rudolph JL (2002) Gender, crime, and desistance: toward a theory of cognitive transformation. *Am J Sociol* 107:990–1064
- Gottfredson MR, Hirschi T (1990) *A general theory of crime*. Stanford University Press, Stanford
- Hay C, Forrest W (2006) The development of self-control: examining self-control theory's stability thesis. *Criminology* 44:739–774
- Hirschi T (1969) *Causes of delinquency*. University of California Press, Berkeley
- Hirschi T, Gottfredson M (1987a) Causes of white-collar crime. *Criminology* 25(4):949–974
- Hirschi T, Gottfredson M (1987b) Toward a general theory of crime. In: Buikhuisen W, Mednick S (eds), *Crime & capitalism*. Brill, Leiden, pp 8–26
- Huesmann LR, Eron LD, Lefkowitz MM, Walder LO (1984) Stability of aggression over time and generations. *Dev Psychol* 20:1120–1134
- Janowitz M (1975) Sociological theory and social control. *Am J Sociol* 81:82–108
- Laub JH, Nagin DS, Sampson RJ (1998) Trajectories of change in criminal offending: Good marriages and the desistance process. *Amer Sociol Rev* 63(2):225–238
- Laub JH, Sampson RJ (2003) *Shared beginnings, divergent lives: delinquent boys to age 70*. Harvard University Press, Cambridge
- Loeber R, Stouthamer-Loeber M, Green SM (1991) Age at onset of problem behavior in boys and later disruptive and delinquent behaviors. *Crim Behav Ment Health* 1:229–246
- Maruna S (2001) Making good: how ex-convicts reform and rebuild their lives. *American Psychological Association*, Washington, DC
- Maruna S, LeBel TP, Naples M, Mitchell N (2009) Looking-glass identity transformation: pygmalion and golem in the rehabilitation process. In: Veysey BM, Christian J, Martinez DJ (eds) *How offenders transform their lives*. Willan Publishing, Portland, pp 30–56
- MacKenzie DL (2006) *What works in corrections: reducing the criminal activities of offenders and delinquents*. Cambridge University Press, New York
- Moffitt TE (1993) Adolescence-limited and life-course persistent antisocial behavior: a developmental taxonomy. *Psychol Rev* 100:674–701
- Moffitt TE (1997) Adolescence-limited and life-course-persistent offending: a complementary pair of developmental theories. In: Thornberry TP (ed) *Developmental theories of crime and delinquency*. Transaction Publishers, New Brunswick, pp 11–54
- O'Connor T, Deater-Deckard K, Fulker D, Rutter M, Plomin R (1998) Genotype-environment correlations in late childhood and early adolescence: antisocial behavioral problems and coercive parenting. *Dev Psychol* 34:970–981
- Patterson GR, Reid JB, Dishion TJ (1992) *Antisocial boys*. Castalia Publishing Co., Eugene
- Plomin R, Chipuer HM, Loehlin JC (1990) Behavioral genetics and personality. In: Pervin LA (eds) *Handbook of personality theory and research*. Guilford Press, New York, pp 225–243
- Rutter M (2007) Gene-environment interdependence. *Dev Sci* 10:12–18
- Sampson RJ, Laub JH (1993) *Crime in the making: pathways and turning points through life*. Harvard University Press, Cambridge
- Sampson RJ, Laub JH, Wimer C (2006) Does marriage reduce crime? a counterfactual approach to within-individual causal effects*. *Criminology* 44(3):465–508
- Shover N (1985) *Aging criminals*. Sage, Beverly Hills
- Thornberry TP (1987) Toward an interactional theory of delinquency. *Criminology* 25:863–892
- Thornberry TP, Krohn MD (2003) *Taking stock of delinquency: an overview of findings from contemporary longitudinal studies*. Kluwer Academic/Plenum Publishers, New York
- Tremblay RE, Frank V, Daniel N, Linda P, Seguin JR (2003) The Montreal longitudinal and experimental study: rediscovering the power of descriptions. In: Thornberry TP, Krohn MD (eds) *Taking stock of delinquency: an overview of findings from contemporary longitudinal studies*. Kluwer Academic/Plenum Publishers, New York, pp 205–254
- Vesey BM, Christian J, Martinez DJ (2009) In: Veysey BM, Christian J, Martinez DJ (eds) *How offenders transform their lives*. Willan Pub, Cullompton; Portland
- Warr M (1998) Life-course transitions and desistance from crime. *Criminology* 36(2):183–216
- Warr M (2002) *Companions in crime: the social aspects of criminal conduct*. Cambridge University Press, Cambridge, UK/New York
- White J, Moffitt TE, Earls F, Robins LN, Silva PA (1990) How early can we tell? preschool predictors of boys' conduct disorders and delinquency. *Criminology* 28:507–533
- Wright JP, Beaver KM (2005) Do parents matter in creating self-control in their children? A genetically informed test of Gottfredson's and Hirschi's theory of low self-control. *Criminology* 43:1169–1202
- Wright BRE, Caspi A, Moffitt TE, Silva PA (1999) Low self-control, social bonds, and crime: social causation, social selection, or both? *Criminology* 37:479–514
- Wright J, Beaver K, Delisi M, Vaughn M (2008a) Evidence of negligible parenting influences on self-control, delinquent peers, and delinquency in a sample of twins. *Justice Quarterly* 25:544–569
- Wright JP, Tibbetts SG, Daigle LE (2008b) *Criminals in the making: criminality across the life course*. Sage, Los Angeles

Social Control Theory of Sexual Homicide Offending

Matt DeLisi¹ and John Paul Wright^{2,3}

¹Criminal Justice Studies, Iowa State University, Ames, IA, USA

²School of Criminal Justice, University of Cincinnati, Cincinnati, OH, USA

³King Abdulaziz University, Jeddah, Saudi Arabia

Introduction

For decades, the study of offenders who commit sexually based offenses generally and who commit predatory sexual homicide specifically has come almost entirely from psychiatry, forensics, and applied criminal justice. Mainstream theoretical criminology has mostly abstained from studying predatory sexual homicide offenders for a variety of reasons that will be explored. A negative consequence of this omission is that criminological theories are often poorly suited to explain the most pathological forms of criminal violence and, by extension, poorly suited to realistically explain the most pathological offenders. This is a major limitation.

This current entry seeks to redress this by exploring the utility of self-control theory (Gottfredson and Hirschi 1990) as a conceptual vehicle to understand sexual homicide offending. Although these offenders evince lifestyles that are consistent with the theory, their instrumental, methodical approach to sexual homicide is also in many respects directly contrary to the tenets of the theory. In addition, the current authors suggest that forensic typologies that seek to characterize sexual homicide offenders as organized or disorganized (Ressler et al. 1988) can be fruitfully understood within a self-control theoretical framework.

Criminology and Sexual Homicide

Two overlapping reasons explain the presumed incompatibility of studying sexual homicide

offenders within the parameters of mainstream criminological theory. The first relates to the infrequency of sexual homicide and the likelihood that offenders in traditional criminological datasets contain such offenders. If murder has a low base prevalence, then the base rate of an even more pathological crime is of course lower still. For instance, the prevalence of serial murder has been estimated to constitute less than 1 % of the total murders in society (McNamara and Morton 2004). This means that community samples and even epidemiological samples simply will not contain predatorily violent offenders such as sexual murderers. Prior studies have found that some of the most frequently used datasets in criminology have approximately zero homicide offenders in them (DeLisi 2001) and instead are comprised of less severe/violent offenders. Due in part to data accessibility issues, mainstream criminology simply declined to understand the most severe offenders in favor of explanations that are compatible with more common, normative offenders.

A second reason that criminological theory has ignored sexual homicide offenders is plain condescension toward a topic that some academics view as too sensationalistic to be taken seriously. Although this might seem like a harsh accusation, the literature is littered with examples of it. For instance, Edwin Sutherland (1950), the patriarch of American academic criminology, argued that sexual psychopath laws – which were designed to socially control instrumentally violent criminals such as sexual murderers – were based on false, questionable knowledge based in popular literature as opposed to science. The tenor of Sutherland's differential association theory is to repudiate individual-level pathologies of the offender in favor of social learning processes. Thus, sexual homicide offenders and the nosology surrounding them (e.g., psychopathy, sexual sadism) were portrayed as antithetical to more appropriate criminological theorizing.

More recently, Ray Surette advanced that serial murderers are media icons, the purpose of which is to “generate fear, degrade social networks, increase reliance on the media, and foster social isolation and polarization” (1994, p. 147).

In a review of a book that was at best tangential to the topic of sexual homicide, Shadd Maruna (2006) opined:

My only substantial complaint though is the embarrassing decision to include the dated, non-academic, non-theoretical chapter by Ressler, Burgess, and Douglas on serial killers in the collection. Talk about letting down the side. With one, cheap attempt to grab the attention of the 'I signed up for criminology because I like 'Silence of the Lambs' market of undergraduate readers. Cromwell almost completely undermines all of the remarkable achievements of the real criminologists pushing the boundaries of criminological research. (2006, p. 275)

Indeed, the indices of major criminological theoretical works and anthologies of criminology theory contain little to no references to sexual homicide offending suggesting that the topic is largely unrelated to mainstream criminological thought.

Fortunately, not all criminologists have been so dismissive of the most extreme and violent criminals (see, e.g., Vaughn et al. 2009). And not all criminologists have ignored potential linkages between criminological theory and homicide offending. For example, Castle and Hensley (2002) surmised that social learning theory could be a useful theoretical framework toward understanding serial murder by exploring experiences from military service and exposure to killing as potential precursors to multiple homicide offending. Social learning theory was also used to understand the prospective linkages between juvenile fire setting and subsequent serial murder (Singer and Hensley 2004). Within the literature, there is also increasing acknowledgement that criminology has for too long overlooked the study of the most violent offenders. To illustrate, DeLisi (2006) noted, "For some time, unfortunately, mainstream criminology tended to focus on more normative offenders and less extreme forms of antisocial behavior. As a result, the study of murder and other extreme forms of violent crime were important but demonstrably peripheral to the discipline. Since the general public is fascinated with more extreme forms of criminal violence, academic criminology clearly missed an opportunity to

connect with public audiences, including practitioners and policy makers" (pp. 172–173). Unlike their criminologist peers, forensic psychologists, neuroscientists, historians, and journalists have focused considerable scholarly energy toward understanding the etiology of sexual homicide, and it is from these disciplines where the sexual homicide epistemology largely resides.

The current entry has an explicit aim: to explore Gottfredson and Hirschi's (1990) self-control theory – a dominant content area in criminological theory and research – and its relation to sexual homicide offending. Sexual homicide offending is defined as the predatory perpetration of murder in conjunction with some sexually oriented offense, such as rape, sodomy, or other type of sexual assault. In some ways, self-control theory provides a useful framework for understanding the lives of predatory sexual murderers particularly those who leave disorganized crime scenes. However, in other ways, the offending behavior of these offenders is characterized by extremely high level of self-control and self-regulation particularly those who leave organized crime scenes. Thus, a self-control paradox exists when studying the apex of violent criminals.

Self-Control and Violence

To date, no study has formally examined self-control theory vis-à-vis sexual homicide although a range of studies have explored the association between low self-control – characterized by Gottfredson and Hirschi as being self-centered, impulsive, poorly tempered, and action oriented as opposed to cognitively oriented, preferring simple tasks as opposed to ones requiring tenacity, and having poor gratification delay – and violence. DeLisi and Vaughn (2008) reported that self-control was a strong indicator of career criminality, a status that is disproportionately responsible for the most serious crimes in a society. Using data from a statewide sample of institutionalized delinquents, they found that those scoring 1 SD above the mean on a low self-control scale were over five times more likely to

become a career criminal and disproportionately likely to commit violent crimes. The association between low self-control and violence is not limited to offender samples. Drawing on data from students, Bouffard (2010) found that self-control was inversely related to both sexual entitlement and self-reported perpetration of sexual aggression. Self-control is a potentially important moderator of criminal careers over time because persons with very low self-control are theorized to continue offending often in violent ways.

Given the nature of most violent disputes, it is unsurprising that the self-control construct would be associated with reactive violence, which frequently typifies intimate-partner violence and generalized assaults (e.g., bar fights). By definition, having a quick, volatile temper means that interpersonal conflicts are not handled internally and ignored; they are handled externally and instantaneously settled. Inherent in this statement, moreover, is the salience of impulsivity and low gratification delay to understanding violence. Consistent with Gottfredson and Hirschi's (1990) theory, violence is an easy option to quickly end the aggravating stimuli created by another person. According to the theory, those with low self-control do not handle frustration and the day-to-day adversities imposed by adult responsibilities and thus will use a quick, easy method to deal with that frustration.

To use a clearly unscientific yet accurate descriptor, it could be said that offenders characterized by self-control deficits are generally *sloppy*. They are sloppy in their school habits and attention to detail which contributes to adverse encounters with teachers and other students and generally negative academic performance outcomes (Houts et al. 2010). They are sloppy in their attention to the emotional and time commitments of interpersonal relationships, and those relationships suffer. Almost an extension of their school difficulties, they are sloppy in their dedication and commitment to work responsibilities and, predictably, experience frequent periods of unemployment. Given the syndromic nature of self-control (DeLisi 2011), it is probable that offender sloppiness is also gleaned from their murder scenes. This issue is explored within the framework of self-control theory next.

In an important work that sought to differentiate the murder scenes of sexual homicides and potentially draw inferences from the crime scenes to profile the offender, Ressler, Burgess, and Douglas (1988) advanced a typology of sexual homicide offending. On one hand was a disorganized sexual murderer whose offense was spontaneous, whose victim was known but depersonalized, and who engaged in minimal conservation during the attack. The disorganized killer's murder scene was random and sloppy, characterized by sudden violence with minimal use of restraints, and sexual acts after death. The victim's body was left in view at the death scene, and importantly, weapons and other forensic evidence were often present at the scene.

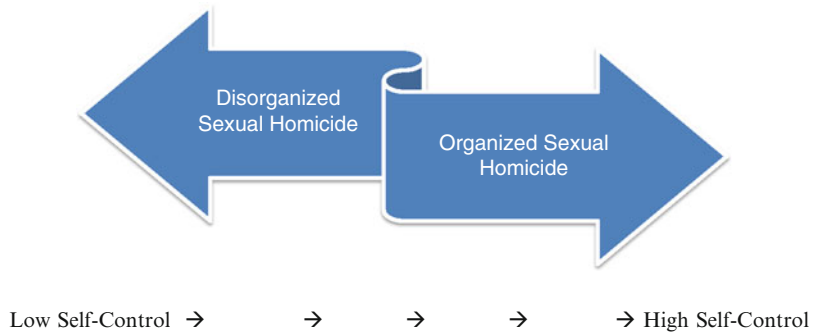
In contrast, the organized sexual murderer committed a planned killing of a targeted stranger who was personalized. There was controlled conservation during the killing, use of restraints, aggressive acts before death, and demand for submission. The entire crime scene of the organized killer reflected control. The victim's body was hidden, often transported, and there was an absence of weapons or other forensic physical evidence. Taken together, the Ressler et al. (1988) typology painted a picture of two very different perpetrators in terms of their functioning, their competence, their mastery over their murder victim, and (semantically most important) their *control* over their murder scene.

As shown in Fig. 1, the disorganized-organized sexual homicide offender typology can effectively be understood as a continuous distribution of self-control. In many respects, the disorganized offender's behavior contemporaneous to their homicide event is consistent with the conduct of the modal criminal offender. The murder is unplanned and hasty and represents a poorly contemplated, rash action to satisfy an underlying desire (recall that the victim is known to the offender). The crime scene of the disorganized sexual homicide offender is, simply, sloppy.

In contrast, the organized sexual homicide offender seemingly exudes self-control in certain respects. The crime scene evinces self-control: forensic evidence is cleaned up, and weapons

Social Control Theory of Sexual Homicide Offending,

Fig. 1 Self-control and sexual homicide organization



and other physical evidence are removed from the scene. The murder itself is premeditated and reflects considerable planning and control; it is a well-executed plan. Planning is not the strong suit of offenders with low self-control. All of these conditions of the organized sexual homicide offender intimate a paradoxical relationship between self-control and violence, one that could not be anticipated by self-control theory.

A sense of higher self-control has also been shown among other types of multiple homicide offenders. In their study of 34 subjects who perpetrated 27 mass murders between 1958 and 1999, Meloy et al. (2001) found that adolescent mass murderers are more often predatorily rather than affectively violent and often do not display sudden, highly emotional warning signs. In other words, their violence represents considerable control evidenced by their selection of victims.

Recasting the organized-disorganized typology as a continuous measure of self-control allows researchers to utilize a criminological theory to conceptually understand sexual murderers. Empirical research is needed to access the threshold where high levels of offender self-control are associated with extremely violent conduct. Additional research is also needed to access the linearity of this relationship. It is possible that the relationship between self-control and sexual homicide is curvilinear. It is also possible that high self-control evinced by sexual homicide offenders is comorbid with other important constructs, particularly sexual sadism and psychopathy. These correlates of sexual homicide offenders are examined next.

In his landmark *Psychopathia Sexualis*, Richard von Krafft-Ebing (1886) introduced the construct of sadism which involved a melding of sexual gratification of the offender with the infliction of pain and humiliation of the victim. In its original conceptualization, sadism was referred to as lust murder and correlated with an absence of guilty (today, known as psychopathy) and even the attempted consumption of the victim. A recent summary review of studies that examined the prevalence of sexual sadism among forensic populations produced a variety of important findings. Chief among them was the finding that the prevalence of sexual sadism was very high, about 30–40 % among offender samples. However, when the sample contained exclusively homicide offenders, the prevalence of sexual sadism was in the 70–80 % range (Krueger 2010; also see, Berner et al. 2003).

An important comparative study also shows the salience of sexual sadism to homicide offending. Yarvis (1995) interviewed and surveyed 78 murderers, 92 rapists, and ten rape-murderers to assess the prevalence of sexual sadism. None of the murderers and just 6.5 % of the rapists met the criteria for sexual sadism; meanwhile, the prevalence among rape-murderers (or, in the parlance of the current entry, sexual homicide offenders or sexual murderers) was 30 %. The criminal careers and life circumstances of sexual homicide offenders are typically severe and noteworthy for their psychopathology even compared to other serious criminals. For instance, Trojan and Salfati (2010) recently compared 137 single-victim homicide cases drawn from the Cincinnati, Ohio, Police

Department and 17 closed serial murder cases obtained from the FBI. Whereas the single homicide offenders averaged 13 arrests and eight convictions, serial murderers averaged five arrests and three convictions. Although this seems comparatively less severe, five or more arrests is a standard measure of habitual or chronic criminality (the most prolific serial killer had 25 prior arrests). The criminal histories of serial murderers also demonstrated more instrumental offenses suggesting a specific targeting of a victim that is consistent with the organized typology described earlier.

Hill et al. (2008) studied the forensic reports of 139 sexual homicide offenders in Germany who had served time in prison and 90 of whom had been released. The mean number of sexual homicide victims per offender was 3.7. Twenty years after their release, more than 23 % of offenders recommitted sexual offenses, and more than 18 % committed nonsexual violent offenses, such as armed robbery and aggravated assault. Three men or more than 3 % were subsequently convicted of attempted or completed murder. They found that the earlier that an offender committed his first sexual homicide, the more likely he was to be convicted of a sexual crime, which included extremely violent offenses such as sexual homicide, rape, sexual abuse of a child, and other sexual assault and any violent crime. Nearly one in five defendants convicted of sexual homicide has prior convictions for homicide, and more than half have other violent crimes in their criminal history (Häkkinen-Nyholm et al. 2009).

Langevin and his colleagues (Langevin 2003; Langevin et al. 1988) compared the antisocial histories of 33 sexual homicide offenders to 80 sexually aggressive offenders, 23 sexual sadists, and 611 general sex offenders. All these men had been seen for psychiatric assessments either for pretrial or as part of parole evaluation. Overall, the criminal careers and antisocial pasts of sexual homicide offenders were significantly worse than even these other severe risk groups. More than 27 % of the sexual homicide offenders had been committed to reform school, more than 21 % were gang members, over 15 % had been expelled from school for behavioral problems,

and greater than 30 % had prior history of animal cruelty, vandalism, and/or fire setting.

Dietz et al. (1990) conducted a descriptive study of 30 sexually sadistic criminals. All of these men intentionally tortured their victims for sexual arousal. Their crimes often involved careful planning, the selection of strangers as victims, approaching the victim under a pretext, participation of a partner, beating victims, restraining victims and holding them captive, sexual bondage, anal rape, forced fellatio, vaginal rape, foreign object penetration, telling victims to speak particular words in a degrading manner, murder or serial killings (most often by strangulation), concealing victims' corpses, recording offenses, and keeping personal items belonging to victims. Beauregard and Proulx (2002) compared the offending process of two groups of non-serial sexual homicide offenders, 20 who were characterized or driven by anger and 16 who were motivated by sadistic pathology. Sadistic murderers were significantly more likely than non-sadistic killers to premeditate their crimes, humiliate their victims, mutilate their victims, and dismember the corpse of the victim. Sadistic murderers took longer to complete their offenses – nearly 90 % took longer than 30 min and were significantly more likely to be apprehended. They also experienced greater positive affect after the commission of their murder. Sadistic killers were also less likely to surrender to police after their crimes, less likely to admit to their crimes, less likely to admit all of the acts committed during their homicides, and less likely than other murderers to admit responsibility.

Research based on analyses of crime scene behaviors of sexual murder and rapes differentiates these forms of sexual violence. In cases of rape, offense behaviors include blindfolding and binding of victims, ripping the clothing of victims, possession and display of weapons, theft, and vaginal penetration. In cases of sexual homicide, there were more severe and depraved indicators of violence including forensic awareness on the part of the offender (e.g., cleaning up or destroying evidence at crime scene), multiple infliction of wounds, and multiple forms of penetration of the victim including objects (Salfati and Taylor 2006).

Across these studies, there is interesting evidence for considerable self-control on the part of sexual murderers. It is not known currently whether these offenders are demonstrating self-control in the Gottfredson and Hirschi (1990) sense or whether they are sadistic psychopaths whose proclivities for extreme violence are guiding their behavior. What is known, however, is that the violence of these offenders at least at the culmination of their murders does not comport with the sloppy profile of the modal criminal offender, one whose life is rife with self-control deficits.

Discussion

For the purposes of theory building, the extreme violence of sexual homicide offenders provides an opportunity to stretch the empirical veracity of constructs. At times, the behaviors of the most severe offenders cast well-known criminological constructs in a new light. For instance, cognitive functioning and intelligence are inversely associated with antisocial behavior, yet some sexual homicide offenders (most famously Ted Bundy) were highly intelligent. This does not impugn the intelligence-crime relationship but does offer opportunities to rethink theory and the putative associations between variables at different points on the offending distribution.

Gottfredson and Hirschi (1990) advanced a theoretical construct that was advertised to be *the cause* of crime, and empirically their theoretical ideas have fared well (see, DeLisi 2011). The current authors believe their theory also has utility for understanding sexual homicide offending and that the range of self-control can be superimposed on existing typologies of sexual homicide. But at another level, the pathology manifest by organized sexual murderers is contrary to the tenets of the theory. They are highly controlled. Based on prior studies of sexual homicide offenders, it is possible that what appears to be high self-control could be spurious and obscured by the

effects of psychopathy and sexual sadism. Indeed, Meloy (1999) has previously characterized organized sexual murderers as compulsive (disorganized sexual murderers were referred to as catathymic in his model), which was characterized among other clinical features as having diagnoses for sexual sadism and psychopathy. Recent research has found that sadism and psychopathy are distinct factors in the psychopathology of sex offenders (Mokros et al. 2011); perhaps self-control is distinct as well.

Related Entries

- ▶ [FBI Influence on State and Local Police](#)
- ▶ [Forensic Science](#)
- ▶ [Specialization and Sexual Offending](#)

Recommended Reading and References

- Beauregard E, Proulx J (2002) Profiles in the offending process of nonserial sexual murderers. *Int J Offender Ther* 46:386–399
- Berner W, Berger P, Hill A (2003) Sexual sadism. *Int J Offender Ther* 47:383–395
- Bouffard LA (2010) Exploring the utility of entitlement in understanding sexual aggression. *J Crim Justice* 38:870–879
- Castle T, Hensley C (2002) Serial killers with military experience: applying learning theory to serial murder. *Int J Offender Ther* 46:453–465
- DeLisi M (2001) Extreme career criminals. *Am J Crim Justice* 25:239–252
- DeLisi M (2006) Book review: encyclopedia of murder and violent crime. *Crim Justice Rev* 31:172–173
- DeLisi M (2011) Self-control theory: the Tyrannosaurus rex of criminology is poised to devour criminal justice. *J Crim Justice* 39:103–105
- DeLisi M, Vaughn MG (2008) The Gottfredson-Hirschi critiques revisited: reconciling self-control theory, criminal careers, and career criminals. *Int J Offender Ther* 52:520–537
- Dietz PE, Hazelwood RR, Warren J (1990) The sexually sadistic criminal and his offenses. *Bull Am Acad Psychiatry Law* 18:163–178
- Gottfredson MR, Hirschi T (1990) A general theory of crime. Stanford University Press, Stanford

- Häkkinen-Nyholm H, Repo-Tiihonen E, Lindberg N, Salenius S, Weizmann-Henelius G (2009) Finnish sexual homicides: offence and offender characteristics. *Forensic Sci Int* 188:125–130
- Hill A, Habermann N, Klusmann D, Berner W, Briken P (2008) Criminal recidivism in sexual homicide perpetrators. *Int J Offender Ther* 52:5–20
- Houts RM, Caspi A, Pianta RC, Arsenault L, Moffitt TE (2010) The challenging pupil in the classroom: the effect of the child on the teacher. *Psychol Sci* 21:1802–1810
- Krafft-Ebing R (1886) *Psychopathia sexualis*. Enke, Stuttgart
- Krueger RB (2010) The DSM diagnostic criteria for sexual sadism. *Arch Sex Behav* 39:325–345
- Langevin R (2003) A study of the psychosexual characteristics of sex killers: can we identify them before it is too late? *Int J Offender Ther* 47:366–382
- Langevin R, Ben-Aron MH, Wright P, Marchese V, Handy L (1988) The sex killer. *Ann Sex Res* 1:263–301
- Maruna S (2006) Review of Paul Cromwell's in their own words: criminals on crime. *Aust N Z J Criminol* 39:274–275
- McNamara JJ, Morton RJ (2004) Frequency of serial sexual homicide victimization in Virginia for a ten-year period. *J Forensic Sci* 49:529–533
- Meloy JR (1999) The nature and dynamics of sexual homicide: an integrative review. *Aggress Violent Behav* 5:1–22
- Meloy JR, Hempel AG, Mohandie K, Shiva AA, Gray BT (2001) Offender and offense characteristics of a nonrandom sample of adolescent mass murderers. *J Am Acad Child Adolesc Psychiatry* 40:719–728
- Mokros A, Osterheider M, Hucker SJ, Nitschke J (2011) Psychopathy and sexual sadism. *Law Hum Behav* 35 (3):188–199.
- Ressler R, Burgess A, Douglas JE (1988) *Sexual homicide: patterns and motives*. D. C. Heath, Lexington
- Salfati CG, Taylor P (2006) Differentiating sexual violence: a comparison of sexual homicide and rape. *Psychol Crime Law* 12:107–125
- Singer SD, Hensley C (2004) Applying social learning theory to childhood and adolescent firesetting: can it lead to serial murder? *Int J Offender Ther* 48:461–476
- Surette R (1994) Predator criminals as media icons. In: Barak G (ed) *Media, process, and the social construction of crime: studies in newsmaking criminology*. Garland Publishing, New York, pp 131–158
- Sutherland EH (1950) The sexual psychopath laws. *J Crim Law Criminol* 40:543–554
- Trojan C, Salfati CG (2010) A multidimensional analysis of criminal specialization among single-victim and serial homicide offenders. *Homicide Stud* 14:107–131
- Vaughn MG, DeLisi M, Beaver KM, Howard MO (2009) Multiple murder and criminal careers: a latent class analysis of multiple homicide offenders. *Forensic Sci Int* 183:67–73
- Yarvis RM (1995) Diagnostic patterns among three violent offender types. *Bull Am Acad Psychiatry Law* 23:411–419

Social Disorder and Physical Disorder at Places

Sue-Ming Yang

Department of Criminology, National Chung-Cheng University, Min-Hsiung, Chia-Yi, Taiwan

Overview

Urban disorder has now become a center of public policy. Sennett (1970) viewed so-called disorder as a manifestation of the social diversity of a place where people come from different racial, ethnic, social, and economic backgrounds. Thus, disorder is considered a byproduct that comes with the modernization of urban cities. Sennett argued that disorder is actually “useful” for urban dwellers, as a disorderly environment “forces” people to get to know one another and therefore, enhances familiarity among heterogeneous population and reinforces racial integration. On the contrary, disorders are often considered as social negativities by criminologists. For example, James Q. Wilson (1975) pointed out that daily hassles such as street people, panhandlers, rowdy youths, or ‘hey honey’ calls trouble urban residents as much as crime does, if not more. Garofalo and Laub (1978) also argued that these urban characteristics of the modern society (disorder is one of them) are the main sources of the “fear” of residents, rather than the true fear of “crime.” Wilson and Kelling (1982) further suggested that disorder has criminogenic effects with untended disorder eventually leading to crime problems. These contrasting viewpoints illustrate the subjective nature of what disorder is and its corresponding social meanings. Browsing through the studies of disorder, it is difficult to identify one single definition that everyone agrees upon. Disorder from one person’s view might represent the norm to another person. In this entry, various definitions about disorder will be reviewed. Two major types of disorder, social and physical disorder, and their association with crime will be reviewed and discussed.

A city isn't just a place to live, to shop, to go out and have kids play. It's a place that implicates how one derives one's ethics, how one develops a sense of justice, how one learns to talk with and learn from people who are unlike oneself, which is how a human being becomes human.

– “*The Civitas of Seeing*” Richard Sennett (1989)

The Sources of Disorder

In the view of the Chicago School researchers, a city is not merely an environmental setting but is rather an organism that grows, changes, and evolves over time (Wirth 1962). Urban areas are usually characterized with diversity in the demographic composition of people who live and work within them. From this perspective, a city looks like a “mosaic of segregated people,” as referred to by Robert Park (1928). In cities, different people demonstrate their cultural heritage, live in different lifestyles, yet cohabit in the same areas. The close contact of these “differences” naturally leads to tension and conflicts among different groups and the lack of cohesion results in social disorganization. Within the disorganized environment, each social or racial group competes for resources and actively seeks to preserve their peculiar cultural forms in order to sustain their own conception of life (Park 1928).

With the expansion of a city, the influx of population brings different cultures and values into a city. For urban inhabitants, therefore, the increase in exposure to different cultures and other value systems and lifestyles tend to create uneasy feelings for individuals. Due to the unfamiliarity of other cultures, an increased level of diversity might be seen as a sign of disorder by mainstream society (Sennett 1989).

However, disorder is by no means a clearly defined concept in either research or practice. The meaning of disorder, just like all other social norms, changes along with the development of a city. In other words, the definition of disorder might vary by society and represents the social norms separating between approval and disapproval behavior (see Durkheim’s Moral Education, 1961).

The definitions of social norms are always the result of competition and the transaction of different cultures and value systems. This process is critical as the social meaning of a behavior often determines how people react to the behavior. In other words, when a society defines certain behaviors as unwanted or deviant, people will start viewing them as a menace or danger to society. Conflict criminologists have long argued that crime is socially constructed and is used by the powerful groups in the society to control the disadvantaged underclass. This ambiguous nature of the social norm is even more salient when dealing with less severe social phenomena such as disorder. Similarly, the defining process of disorder is not equally influenced by all groups in the society—the powerful groups often enjoy the privilege of determining the “norms” (Durkheim 1961; Sennett 1970).

The Definition of Disorder

Disorder is not crime, and usually represents minor violations of social norms that do not directly harm other people. Unlike crime, there is no commonly agreeable definition to classify disorder. Skogan referred to it as a “slippery concept” (Skogan 1990, p. 4). He argued that disorders are behaviors that are not prohibited by the criminal law; or disorders in isolation constitute relatively minor offences. The ambiguity nature of disorder leaves room for different ways defining disorder.

In prior literature, various occurrences and behaviors have been identified as disorder, ranging from deteriorated buildings to teenagers hanging out on the street corner. Even the term “disorder” is not always used to describe these minor violations. Past research has referred to this particular phenomenon as “signs of crime” (Skogan and Maxfield 1981), “early signs of danger” (Stinchcombe et al. 1980), “urban unease” (Wilson 1968), “broken windows” (Wilson and Kelling 1982), “soft crime” (Reiss 1985), “public moral offenses” (Weisburd and Mazerolle 2000), “cues to danger” (Warr 1984), “incivilities” (Taylor 1999), “sign of disrespect”

(Bannister et al. 2006), “warning signals” (Innes 2004), and “disorder” (Skogan 1990). Finally, most scholars have now settled on the term of either *disorder* or *incivilities* (LaGrange et al. 1992). While disorder has been discussed and theorized prior to 1990 (see Wilson and Kelling 1982), Skogan was the first to clearly articulate disorder and its impacts on neighborhoods (Skogan 1990). In his book *Disorder and Decline*, Skogan points out that the concept of disorder is ambiguous and can represent a variety of meanings. Thus, he notes that it is important to differentiate between the friendly “active uses of the environment” (not disorder) and disorderly behavior which bothers the residents of a neighborhood, as the former may suggest a very well-connected neighborhood while the latter may represent a disoriented neighborhood.

In addition to how research defines disorder, it is also important to know how law enforcement determines what phenomena are considered disorderly. In Kelling and Sousa’s evaluation of New York City’s Quality of Life Policing effort, they used misdemeanor arrests to represent police efforts without providing a clear list of targeted disorders. In the four scenarios illustrated in the report, the enforcement was mainly targeted at public urination or public drinking (open container). Sure enough, based on the broken window thesis drafted by Wilson and Kelling, the actual practice probably included more types of “minor offense” including things like panhandlers, drunks, addicts, rowdy teenagers, prostitutes, loiterers, and the mentally disturbed. On the other hand, Thacher (2004) summarized order maintenance police practices and noted four main categories of “disorders” that were the focus of police: (1) obstructing or lying down in the subway, (2) public urination, (3) panhandling, and (4) youth or student parties (page 396).

What kind of phenomenon is considered disorderly? After reviewing the various definitions and findings from studies, Kubrin (2008) concluded the following “Definitions of disorder used by researchers and officials studying and practicing broken windows policing are not necessarily consistent with residents’ perceptions in

their own communities.” (page 206). From the previous literature review, it is apparent that disorder is indeed a “slippery concept” in both research and practice.

The Disorder and Crime Association

Recently, disorder has become a new center of attention for criminologists and criminal justice practitioners. From the Broken Windows Thesis proposed by Wilson and Kelling (1982), fear of crime research (see Garofalo and Laub 1978), to broken windows policing (or called the quality of life policing in New York City), disorder has been formulated as the core element in the equation of inquiry.

Despite the popularity of the broken windows based ideas among practitioners, the association between disorder and crime has been a much debatable topic (see Taylor 1999 and Sampson and Raudenbush 2004). Wilson and Kelling (1982) asserted that disorders are the root cause of crime. They suggest that signs of disorder lead to a sense of low social control perceived by those in the community, which triggers residents’ fear. Fearful residents may further withdraw from the community, and subsequently more serious types of crimes will occur. Therefore, targeting crime would not solve the fundamental problem, but focusing on disorder intervention will result in a greater crime prevention benefit.

Similarly, Skogan believes that incivilities will eventually result in neighborhood decline because they cause a range of psychological, social-psychological and behavioral outcomes in neighborhoods. Beginning with the appearance of incivilities, the decline of a neighborhood is a gradual process (Skogan 1990, p. 65). First, signs of incivilities cause social withdrawal of residents from participating community activities. It also discourages people from cooperating with their neighbors. Second, the existence of incivilities weakens neighborhood morale. Residents become concerned about their personal safety after a number of incidents happen in their neighborhood. As a result, residents lose trust in each other. Third, from a practical

standpoint, incivilities undermine the housing market. It directly affects the willingness of people to invest in the area. Due to a lack of investment, the housing market is suppressed and neighborhood decline becomes an unavoidable outcome. To test his hypothesis, Skogan (1990) collected information from 40 neighborhoods in six different cities during 1977–1983. The simple association between disorder and robbery victimization in those neighborhoods was + .80. However, when other neighborhood factors were taken into account, the association dropped to + .54, though still fairly substantial. Based on these findings, Skogan concluded that with the small number of cases at hand, it is hard to tell “*whether they have either separate ‘causes’ or separate ‘effects’ at the area level*” (1990, 73).

Keizer et al. (2008) conducted a series of field experiments on physical disorder and found that the presence of disorder, such as graffiti, trash etc. increased the likelihoods of passersby to commit minor disorderly behavior. Though this study provides some direct support of the negative effects of disorder, the study did not examine the effects of disorder on crime. Moreover, their study focuses on physical disorder only, rather than both types of disorder. The association between disorder and crime was further confirmed by Weisburd et al. (2012) longitudinal study of crime trends in Seattle, WA. Controlling for opportunity variables as well as characteristics derived from social disorganization theory at street level, physical disorder was found to be one of the strongest indicators predicting crime trends. Specifically, streets with more physical disorder incidents were much more likely to be in the chronic crime pattern instead of crime-free pattern. Also, increases in physical disorder over time were related to increased crime trends. Though the findings lend confidence to the disorder-crime association, only physical disorder measure was included in the study.

While both Wilson and Kelling and Skogan believe disorder in the neighborhoods causes crime and other social negativities, other scholars hold different opinions on the nature of their association. In a study of drug hot spots in Jersey City, Weisburd and Mazerolle (2000) found that

serious crime and disorder tend to cluster together in drug hot spots. In other words, disorder and crime are positively correlated at these drug hot spots. In contemporary social disorganization theory, both Bursik and Sampson treat disorder as a consequence of the lack of social control or as a product of structural variables. Thus, disorder is only related to crime through its associations with structural variables. That is, crime and disorder should be positively correlated but the relationship is not *causal*.

Corman and Mocan (2000) examined the association between disorder and crime rates in NYC over time. However, they did not collect directly measured disorder data; rather, they used misdemeanor arrests as the proxy of disorder. They found that misdemeanor arrests in NYC from 1974–1999 were significantly, negatively related to robbery, motor vehicle theft and grand larceny after controlling for economic conditions and deterrence, but were not significantly related to the other four index crimes. Despite the empirical disagreement, within all the theoretical arguments about disorder/incivilities, most of them assume a positive association between disorder and crime. Meaning, wherever there is disorder, there tends to also have crime problem. The issue is, whether the association is correlational or causal. Also focusing on New York City, Geller (2007) examined housing conditions and the relationship to crime rates in NYC. She tested whether deteriorated housing conditions leads to crime increases as expected by the broken windows thesis. The results of her study did not support such a hypothesis. However, in the study, only the effects of physical disorder were tested. This is an important distinction as some studies have suggested that social and physical disorder may have different relationships with crime.

St. Jean (2007) conducted an ethnographic study in a high crime neighborhood in Chicago and found that disorder and crime tended to coexist in the same areas. However, he pointed out the differential association of social and physical disorder with crime. Yang (2010) in a longitudinal study of City of Seattle, WA, also found a significant correlation between disorder and crime. The direction of causation, however,

was opposed to what was suggested by the broken windows thesis. The results from Granger causality tests generally showed no causal relationship between disorder and violence, and in a few places the causality appeared to run from violent crime to disorder. Additionally, Yang shared St. Jean's findings and argued that research should make a distinction between social disorder and physical disorder. They are not only qualitatively different phenomena; moreover, these two types of disorder have different associations with crime.

Separating Social Disorder and Physical Disorder

Activities under the classification of disorders, or incivilities, are not all homogeneous. Under the big umbrella of disorder, there are two commonly used subcategories: physical disorder and social disorder (Skogan 1990; Hancock 2001; Sampson and Raudenbush 1999). Later, Taylor (1999) proposed very specific operational definitions of social and physical disorder (incivilities). Social incivilities include behaviors such as public drinking, drunkenness, rowdy and unsupervised teens, sexual harassment, arguing and fighting, open prostitution, and public drug sales. While physical incivilities include things like abandoned buildings, graffiti, litter, vacant and trash-filled lots, unkempt yards and housing exteriors, abandoned cars and the conversion of houses and apartments to drug-selling locations.

From a research standpoint, social disorder and physical disorder are qualitatively different. Sampson and Raudenbush (1999) clearly pointed this out in the beginning of their article: “[s]ocial disorder, we refer to behaviors involving strangers and considered threatening. . . [p]hysical disorder, we refer to the deterioration of urban landscapes (p. 603–604).” From this statement, it is obvious that social disorder requires the involvement of human behaviors, which impose threats to others in the neighborhood. Thus, the key element defining social disorder depends on the presence of actors who perform the offensive actions. St. Jean (2007)

also believes that neighborhood disorder exists in two forms: physical and social disorder. He argues that physical disorder refers to “unpleasant neighborhood characteristics” created by messy environment such as litter, trash, abandoned tires, empty lots, and broken windows. On the contrary, social disorder is a result of the “unpleasant and perhaps intimidating social interactions among people in public space.” As such, social disorder includes social behaviors like panhandling, public drinking, fighting, and loitering (St. Jean 2007, p. 196).

In addition, Hancock (2001) suggested another less general way of classifying disorder by looking at its duration. Based on its temporal length, he classified disorder into two groups: the episodic disorder which occurs at no fixed time, such as public drinking, and the on-going type of disorder like abandoned buildings which are always there, unless someone takes initiative to change the condition. Putting these ideas together, social disorder is usually an episodic behavior, which only lasts for a short duration. Conversely, physical disorder represents an objective condition that might last for a long period of time unless some actions are taken to change it. The focus of the latter is on the physical conditions rather than on individuals acting within the conditions.

Many of the past studies proposed clear definitions of social and physical disorder along with characteristics included within each category (Sampson and Raudenbush 1999; Skogan and Maxfield 1981; Skogan 1990; Taylor 1999; LaGrange et al. 1992). However, when doing analysis, people often combined these two categories together into one total disorder measure – assuming that they represent the same underlying construct. For instance, Skogan and Maxfield (1981) used a general term of “sign of disorder” to represent the violations of “people’s expectations about fit and proper conditions and conducts.” Within this “sign of disorder” they included a wide range of circumstances like unsupervised teens hanging out on the streets, abandoned buildings, illegal drug use, and vandalism. Sampson and Raudenbush (1999), though proposed clear definitions about social and physical disorder separate, did not examine their

individual effects in their test of the broken window thesis. In their analysis stage, they made a statement that “*the results are so similar for physical and social disorder, and because the two scales are highly correlated ($r = .71$), we combine them into a summary index of disorder...*” (p. 626). As such, the distinction between social and physical disorder and their associations to crime were not fully examined in the study.

In a separate study, Raudenbush and Sampson (1999) tested methodological properties of social and physical disorder measures using item response modeling with hierarchically constructed data. They found that social disorder measures were recorded less often and thus, less reliable compared to physical disorder measures, due to the rarity of the social disorder observations.

Steenbeek and Hipp (2011) also followed the item response modeling approach to explore the reciprocal effects of disorder on social structures and social control under the social disorganization framework. Using data collected in the Netherlands, they argued that the correlations between social disorder and physical disorder items were too strong. Thus, they constructed a disorder measure at the neighborhood level combining all social and physical disorder items rather than separating the two constructs.

In short, despite the fact that many prior studies have acknowledged the distinction between social and physical disorder, their separate effects to crime and other social problems are still understudied empirically.

Social Disorder, Physical Disorder, and Crime

Overall, the qualitative differences between social and physical disorder are rarely addressed. Mixing the concepts of social and physical disorder hinders us from disentangling the intertwined relationship between disorder and crime. From theoretical perspectives, social disorder and physical disorder are different in many ways. First, social disorder involves actors and current actions; thus, the presence of social disorder perhaps provides a pool of potential targets,

motivated offenders or both for violent offenses to occur. Physical disorder, however, does not necessarily involve actors. Most of the time, we do not see/know the actors who dump trash or break glass in the next block. We recognize physical disorder when we see it, even without seeing the creator(s) of it. As such, physical disorder provides unmistakable visual cues to the users of the space. Therefore, residents’ perception of physical disorder should be more consistent than their perceptions of social disorder as the latter involves an individual’s value judgment.

Compared to social disorder, physical disorder is generally considered less criminogenic for crime. An abandoned place could attract would-be offenders to engage in criminal activities. However, physical disorder itself, without the presence of potential offender/victims, should be related to crime to a lesser extent than social disorder. It is possible that there exists a sequential relationship between social disorder and physical disorder. This possibility will be examined later using the Granger causality test to see if there is a directional relationship between the two types of disorder.

The distinction between social and physical disorder also carries importance on another equally important subject—fear of crime. In addition to crime, fear is another phenomenon that has long been connected to disorder (see Garofalo and Laub 1978; Kelling and Coles 1996). Kelling and Coles (1996) argued that fear mediates the effects of disorder on crime in a community. However, the differential effects of social and physical disorder were not discussed by Kelling and Coles. Based on Sampson and Raudenbush’s phrases cited earlier, it is expected that social disorder and physical disorder should have different impacts on fear. Social disorder invokes a feeling of fear to residents, while physical disorder provides a more neutral image that might still be bothersome to some residents. As such, the emotional reactions related to social and physical disorder are not necessarily the same.

Empirical studies also support that social disorder and physical disorder are different and so are their associations to crime. The empirical evidence from Taylor’s (1999) observations in

Baltimore provides a good example of how different those two types of incivilities can be. In the study, physical disorder is found to be correlated with changes in rates of aggravated assault, burglary, and motor vehicle theft while social disorder is correlated only with changes in the rate of rape (all of the correlations are significant at .10 level, Taylor 1999, p. 5). St. Jean (2007) analyzed crime data and interviewed residents in poor neighborhoods in Chicago to study disorder and crime. From cross-tabulation analyses, he found that places with high collective efficacy and low social disorder have fewer crime problems. Places with low social disorder tend to have low street drug dealing. However, high physical disorder had no significant association with drug dealing, robberies and batteries. While the finding seems to support the broken windows thesis, he argued that places with high levels of social disorder do not necessarily have more crime problems. The data at best could argue that in the absence of social disorder, it is also less likely for the places to have drug dealing problems.

Nonetheless, St. Jean sought for a more thorough understanding about the underlying mechanism linking disorder and crime by conducting in-depth interviews with street criminals and other individuals who live in the crime ridden areas to get their perspectives. The presence of physical disorder alone does not attract drug dealers to the area; rather, it is the location and the activities occurring within and nearby the area that determine whether it is going to become a drug hot spot. Though the dilapidation of buildings does show that the government did not invest in the area and perhaps there is a demand for drugs in the place from people who could not find any decent jobs. The association between robberies and physical disorder is even weaker. The robbers pick places to commit crime based on accessibility and anonymity, not because of the abandoned buildings or trash lying around in the areas. As for batteries, he found that places with high levels of physical disorder tend not to have battery problems. It is reasonable because places with physical disorder tend to be occupied by abandoned lots or buildings and therefore, are less likely to have domestic disputes due to fewer people living in the area.

From the qualitative interviews, St. Jean found that social disorder has strong and significant associations with narcotic violations, robberies, and batteries. But a closer examination of the mechanism behind the association revealed the complexity in the phenomena. Places with low social disorder tend not to have drug dealings; however, places with high social disorder may or may not attract drug dealers to the areas. Moreover, the association between social disorder and narcotic violations is sometimes reciprocal. That is, people who use drugs might be involved in social disorder while they are waiting for their needs to be met. As for robberies and social disorder, the relationship is positive and significant. But further inquiry shows that the locations and actors in the places matter more than whether the areas have social disorder. Sometimes, having more people hanging around a place provides a potential pool for robbery targets. However, when the people do not seem to be vulnerable or do not possess enough monetary goods, they would not be selected as targets. The tight coupling between social disorder and batteries shows that places with more disorderly people are also more likely to have domestic conflicts. Again, St. Jean found that the association is only supported in places occupied by high social disorder and disruptive families. In places that have high levels social disorder but are occupied mainly by non-disruptive families the association diminishes. Overall, the findings from the ethnographic portion of this study were consistent with what was found in the statistical analyses. That is, physical disorder does not matter in predicting crime places. Moreover, social disorder does not matter in offenders' decisions in choosing a place as their base; what really matters to them is where the place is located and who are hanging around the place. St. Jean's study provides a comprehensive picture illustrating the relationships among crime, social, and physical disorder. Nonetheless, the analyses in the study were all based on a short time period and limited the selected poor neighborhoods in one high-crime police beat in Chicago. Therefore, it is hard to know whether the relationships remain the same over time in other places.

Finally, in another study using 16 years of data collected at the census-block-group-level in Seattle, Yang (2010) studied the long-term associations between social disorder, physical disorder, and violent crime. The dual-trajectory analysis findings suggested that a block group with a low level of disorder (for both social and physical disorder) is less likely to experience violence problems. If a block group is assigned to the low rate social disorder trajectory, then there is an 88.1 % of chance that this block group will also be found in the no (or negligible) violent trajectory. Similarly, 82 % of block groups with negligible physical disorder problems are also relatively free of violence.

However, block groups that have high disorder problems are not necessarily also plagued by violence. For block groups within the high social disorder trajectory, there is only about a 30.6 % of chance of also being classified into the high rate violent trajectory group. Also, having been assigned to the high physical disorder trajectory only corresponds to a 30 % of chance of having a high level of violence.

This means that having no disorder can be seen as a powerful protective factor for block groups in preventing future violence – as there is less than 1 % of a chance that a block group will have a high level of violence if it has negligible disorder problems. Although having high disorder can be a risk factor to increase the chance that a block group experiences violence, only about 30 % of block groups that showed a high level of disorder also showed a high level of violence. This means that by using disorder as a risk factor, we will have a 70 % chance of making an incorrect prediction of where high violence places would be located. Perhaps lack of disorder and having high-disorder are not the two sides of the same coin. Or perhaps there is not a linear relationship between disorder and violence.

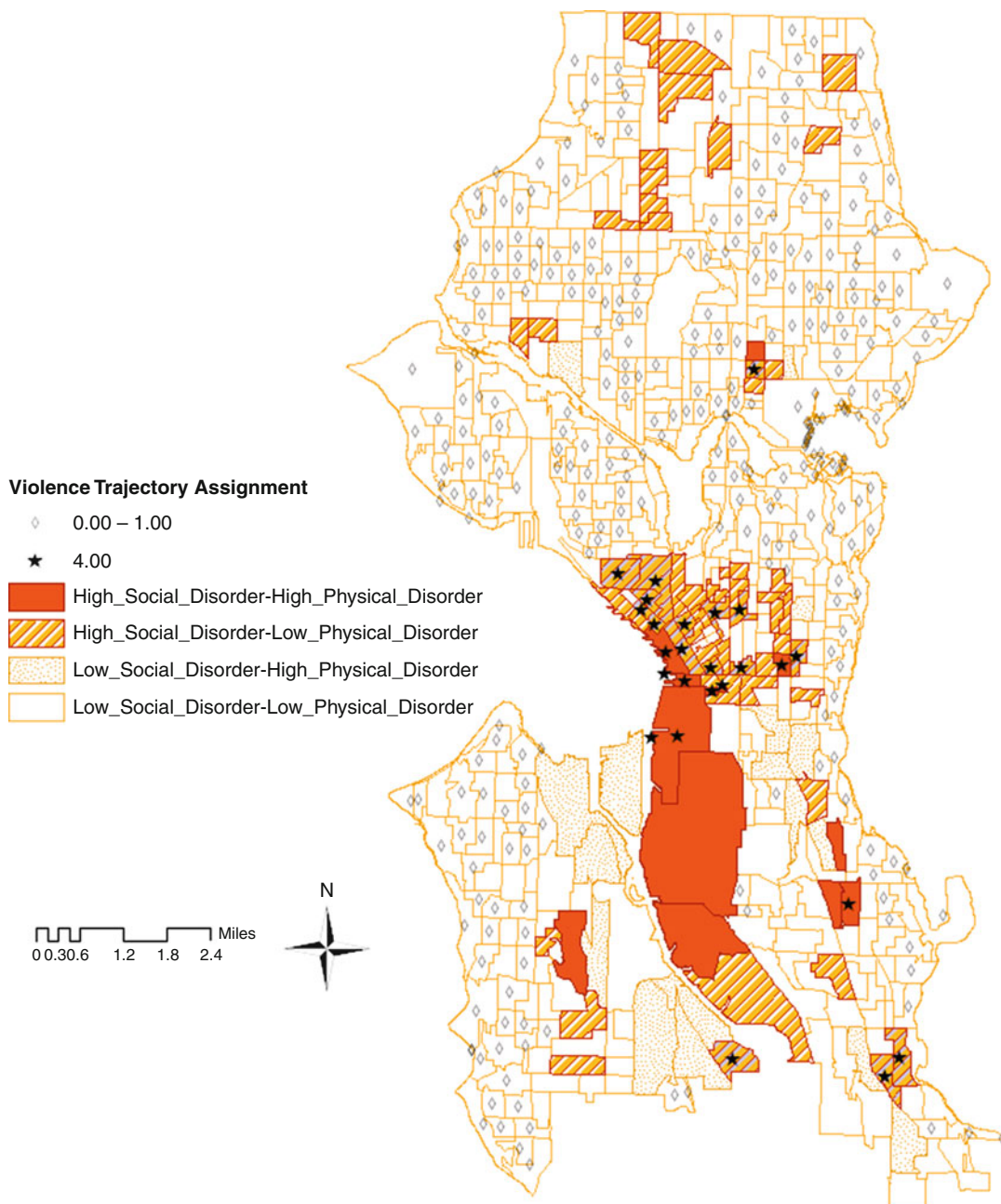
Furthermore, Yang also conducted geographic analysis on the selected block groups with high levels of social disorder, physical disorder, and violence (see Fig. 1). The visual illustrations on social disorder-physical disorder dyad and violence further confirm the previous findings. First, block groups with high violence trajectory

are always located within the areas with a high level of social disorder. However, places having high social disorder do not necessarily have high violence rates. In other words, to locate a high violence block group, social disorder provides more accurate information than physical disorder. This finding echoes St. Jean's study results from Chicago. Second, places that are low in both social disorder and physical disorder are also free of violence problems. When a block group is free from disorder problems, it is also more likely to be free from violence problem. On the other hand, when a block group starts to show any signs of a disorder problem, it is very unlikely that it will exhibit a low rate of violence. High physical disorder block groups, though not informative of where the high violence block groups are located, are predictive of where the *low* violence block groups will *not* be. The spatial concordance between social disorder and violence points out a possibility that there may exist a connecting mechanism that governs both disorder and violence at places. The studies of St. Jean and Yang both show that having more broken windows does not necessarily predict what the level of crime will be at a given place.

The Next Step for Disorder Research

Overall, the empirical findings and theoretical arguments point to the need for studying these two types of disorder separately. Social disorder is found to be more relevant to crime than physical disorder. Perhaps it is because people involved in disorderly behaviors are also more likely involved in violence, regardless of the environmental factors. Or, perhaps it is also possible that these people might provide visible targets for motivated perpetrators or they might be the instigators of violent behavior.

Another possible source of variation is the type of measures used in the study. One of the major obstacles on studying the linkage between disorder and crime is whether residents could actually distinguish between disorder and crime (Gau and Pratt 2008). This is crucial as the mechanism outlined in the broken windows thesis to be effective; residents need to be aware of the



Social Disorder and Physical Disorder at Places, Fig. 1 Geographic Locations of the 15 Selected Block Groups

existence of disorder. As such, some scholars started to research the impacts of perceptual disorder and the factors that could lead to perceived disorder. Clearly, observed levels of disorder play a role in this process, but perceptions are more critical. Generally, it is found that the perceptions

of crime and disorder are highly related and do not have discriminant validity (see Gau and Pratt 2008; Worrall 2006). In other words, disorder and crime might be manifestation of the same underlying construct with different level of intensity. Thus, police strategies targeting disorder, as opposed to

crime, might not lead to either crime or fear reduction effects as intended.

Innes (2004) also proposed a potential mechanism about how disorder affects residents' perceptions about crime and disorder. What matters most is not the nature of disorder; rather, it is the interpretation of disorder that leads to fear and negative emotions. The signs of disorder provide a framing function that helps people to give meanings to the quality of places. As such, social disorder and physical disorder might send out different signals to people. Physical disorder might reveal the lack of investment of the places by both government and residents. On the other hand, social disorder can be indicative of the type of people who hangout in the areas and the strengths of social ties in the area.

Hinkle and Yang (in progress) further point out that it is important to evaluate whether residents' perceptions of disorder matches with the observed level of disorder measured by researchers through systematic social observations. They further argue that social disorder and physical disorder might have differential conditional effects on residents' perception about the disorder problem, fear of crime, and satisfaction of the residential environment. Thus, scholars who are interested in disorder-related research need to examine social disorder and physical disorder separately. Moreover, it is also important to control the methods of measurement used to gauge levels of disorder, be it the actual level or the perception of the problem.

Disorder, unlike crime, is a common phenomenon and exists in many people's daily lives. It has also become an important issue for research in many disciplines. Recent studies have shown the importance of evaluating social disorder and physical disorder separately. These two types of disorder are not only qualitatively different, but they also relate to crime in very different ways. Moreover, the presence of social disorder and physical disorder send very different signals to passersby, residents and potential offenders and possibly lead to different perceptions about places. Future research needs to pay closer attention to the distinction between the two types of disorder and their implications for social policy.

Related Entries

- ▶ [Broken Windows Thesis](#)
- ▶ [Defining Disorder](#)
- ▶ [Discriminant Validity of Disorder and Crime](#)
- ▶ [Disorder: Observational and Perceptual Measures](#)
- ▶ [Order Maintenance Policing](#)

Recommended Reading and References

- Bannister J, Fyfe N, Kearns A (2006) Respectable or respectful? (in) civility and the city. *Urban Studies* 43:919–937
- Corman H, Mocan N (2000) A time-series analysis of crime, deterrence, and drug abuse in New York city. *Am Econ Rev* 90:584–604
- Durkheim E (1961) *Moral education*. The Free Press, New York
- Garofalo J, Laub JH (1978) The fear of crime: broadening our perspective. *Victimology* 3:242–253
- Gau JM, Pratt TC (2008) Broken windows or window dressing? Citizens' (in) ability to tell the difference between disorder and crime. *Criminol Public Policy* 7:163–194
- Geller A (2007) *Neighborhood disorder and crime: an analysis of broken windows in New York City*. Unpublished Doctoral Dissertation, New York: Columbia University
- Hancock L (2001) *Community, crime and disorder: safety and regeneration in neighborhoods*. Palgrave, New York
- Innes M (2004) Signal crimes and signal disorders: Notes on deviance as communicative action. *The British Journal of Sociology*, 55(3):335–355
- Keizer K, Lindenberg S, Steg L (2008) The spreading of disorder. *Science* 12:1681–1685
- Kelling GL, Coles C (1996) *Fixing broken windows: restoring order and reducing crime in American cities*. Free Press, New York
- Kubrin CE (2008) Making order of disorder: a call for conceptual clarity. *Criminol Public Policy* 7:203–214
- LaGrange RL, Ferraro KP, Supancic M (1992) Perceived risk and fear of crime: role of social and physical incivilities. *J Res Crime Delinq* 29:311–334
- Park RE (1928) Forward. In: Wirth L (ed) *The ghetto*. The University of Chicago Press, Chicago, IL, pp vii–ix
- Raudenbush S, Sampson RJ (1999) *Ecometrics: toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods*. *Sociol Methodol* 29:1–41
- Reiss AJ (1985) *Policing a city's central district: the Oakland story*. U.S. Department of Justice, National Institute of Justice, Washington, DC (March)

- Sampson RJ, Raudenbush SW (1999) Systematic social observation of public spaces: a new look at disorder in urban neighborhoods. *Am J Sociol* 105:603–651
- Sampson RJ, Raudenbush SW (2004) Seeing disorder: Neighborhood stigma and the social construction of “broken windows.” *Social Psychology Quarterly*, 67, 319–342
- Sennett R (1970) *The uses of disorder: personal identity and city life*. W.W. Norton and Company, Inc., New York
- Sennett R (1989) The civitas of seeing. *Places* 5(4):82–84
- Skogan WG (1990) Disorder and decline: crime and the spiral of decline in American neighborhoods. Free Press, New York
- Skogan WG, Maxfield MG (1981) *Coping with crime: individual and neighborhood differences*. Free Press, New York
- St. Jean PKB (2007) Pockets of crime: broken windows, collective efficacy, and the criminal point of view. The University of Chicago Press, Chicago, IL
- Steenbeek W, Hipp JR (2011) A longitudinal test of social disorganization theory: feedback effects among cohesion, social control, and disorder. *Criminology* 49(3):833–871
- Stinchcombe AL, Adams CA, Heimer KL, Scheppelle TW, Taylor DG (1980) *Crime and punishment – changing attitudes in America*. Jossey-Bass, San Francisco
- Taylor RB (1999) Responses to disorder: relative impacts of neighborhood structure, crime, and physical deterioration on residents and business personnel. National Institute of Justice, Washington, DC
- Thacher D (2004) Order maintenance reconsidered: moving beyond strong causal reasoning. *J Crim Law Criminol* 94(2):381–414
- Warr M (1984) Fear of victimization: why are women and the elderly more afraid? *Soc Sci Quart* 65:681–702
- Weisburd D, Mazerolle LG (2000) Crime and disorder in drug hot spots: implications for theory and practice in policing. *Police Q* 3(3):331–349
- Weisburd D, Groff ER, Yang S-M (2012) *The criminology of place: street segments and our understanding of the crime problem*. Oxford University Press, New York
- Wilson JQ (1968) *Varieties of police behavior*. Harvard University Press, Cambridge
- Wilson JQ (1975) *Thinking about crime*. Basic Books, New York
- Wilson JQ, Kelling GL (1982) Broken windows: the police and neighborhood safety. *Atlantic Monthly* 211:29–38
- Wirth L (1962) *The ghetto*, 7th edn. The University of Chicago Press, Chicago
- Worrall JL (2006) The discriminant validity of perceptual incivility measures. *Justice Q* 23:360–383
- Yang S (2010) Assessing the spatial-temporal relationship between disorder and violence. *J Quant Criminol* 26(1):139–163

Social Disorganization and Terrorism

Susan Fahey

Criminal Justice Program, Richard Stockton College of New Jersey, Galloway, NJ, USA

Synonyms

Anomie; Breakdown; Collective action; Normlessness; Political action; Violent

Overview

One would expect more terrorism incidents to originate in a socially disorganized environment, that is, an environment that has experienced a rapid social change that has destroyed the regular rules and bounds on human behavior. The notion that one’s surroundings may influence behavior is not a new one. In particular, the idea that the organization of the social environment, such as neighborhoods, families, and communities, may stimulate violence has been around since the origins of the social sciences. Emile Durkheim was a pioneer in describing the anticipated effects of the social environment on violent behavior.

Anomic Behavior

To understand how a socially disorganized environment influences one’s behavior according to Durkheim (1930 [1951]), it is important to understand how he viewed human nature. To Durkheim, humans lacked the ability to regulate their desires and wants. In fact, humans were like other animals in that they were controlled by their needs and desires, particularly food, shelter, and the sexual drive. Unlike animals, though, humans have the power of reflection, the ability to imagine more than they have, and it is because of this ability to imagine more that they require an

external force to limit their desires and needs. In normal times, the social environment provides this external limiting force; that is, one's behavior is regulated by the social rules and norms that are ordinarily in force. For example, if an individual sees that another has a very nice and desired vehicle, under most circumstances, that individual will not simply go and take that vehicle by force. Under most circumstances, the individual will refrain from taking the vehicle, because the individual is kept in the bounds of legal and moral behavior by the traditional rules of society. The fear of arrest and the disappointment expected by one's family and community as well as other social consequences keep individuals from simply taking the object of their desires.

However, Durkheim (1930 [1951]) says that if the society is experiencing temporary but rapid social changes, normal circumstances do not apply. In fact, during these times, what are called anomie or socially disorganized times, it is as if the rules of society have been stripped off and individuals are left in a world with no normal ways of doing things and no rules about wrong and right. In addition, the power of reflection, the power to imagine more, better, and different will come into play. In this world, individuals may become so disoriented by the loss of rules and controls that they may start to want everything they see. In many cases, individuals will not be able to satisfy all of their wants and desires and will become unsettled. Because Durkheim studied what led people to commit suicide, he states that such persons are more likely to commit suicide. Compared to well-regulated societies, those which experience rapid social changes that create anomie are more likely to experience increases in suicides.

Durkheim's (1930 [1951]) writings on the effects of anomie on behavior can be and have been applied to crime and terrorism. An example of a rapid social change sufficient to bring on a state of anomie in the society might be a civil war between an insurgent force and the government. In such a society, where an insurgent force is violently attempting to overthrow the government and its rights to make and enforce rules about legal and illegal conduct, the rules that

dictate right and wrong in everyday life will not be in place. This constitutes anomie, a state of deregulation of behavior.

An individual living in such a state will be unsure as to what the new rules and regulations are in this different, anomic world. When the old rules were forcefully removed by the outbreak of civil war and the new rules have not been made clear or taken root yet, individuals will either need to restrain their desires and needs themselves or they will find this new anomic world without rules intolerable. The inability to tolerate this anomic world may make it more likely for some people to do behavior that is against the rules and norms of behavior that used to be in force in the society. If such individuals are sufficiently motivated, perhaps by a grievance against the government (Gurr 1976), such individuals may become involved in extreme acts, including terrorism. Anomie, according to Durkheim, is not a permanent condition. Over time, a new set of rules to define what is normal in society will develop, and the society will regain its equilibrium. This new set of rules will make living more tolerable for individuals again, and this will decrease the rate at which individuals engage in problematic behavior.

It is important to remember that this theory of human behavior is probabilistic; it suggests that the presence of such rapid social change makes it *more likely* that *some* individuals will engage in actions that are problematic. It is not a theory that predicts that all individuals will engage in problematic behavior nor is it definite that any specific individuals will be so engaged.

How Much Rapid Social Change Is Needed?

To result in anomie, a rapid social change must involve a large-scale structural challenge to the ordering and functioning of societal institutions (Piven and Cloward 1977). In other words, the rapid social change has to be serious and broad. It is also important that the rapid social change be large enough to stimulate changes in the structure and routines of everyday life. If a rapid social

change occurs that does not interrupt the structure and routine of everyday life for most people, it is unlikely to change the behavior of those individuals whose routines continue uninterrupted. For example, when individuals are released from the routine of going to work every day as well as the sustenance provided by work, this changes the pattern and rhythm of their lives. If the interruption continues for a period, it is likely sufficient to produce anomie.

How Do Anomic Conditions Affect Behavior?

A socially disorganized environment likely affects behavior through its effects on the structure and routine of everyday life (Snow et al. 1998). People are more likely to take action to preserve what they know and their current routines rather than to seek something new and better. The known routines they will act to save are the things they do every day (behaviors) and the way they think about these things (cognition). The behaviors involve habitualized patterns of action. The cognition involves attitudes individuals adopt when approaching their daily routines. These attitudes chiefly involve an unquestioning, unreflective routine way of getting through the day.

These behavioral and cognitive routines of everyday life are what constitutes the *quotidian*, which is derived from the French word for daily. For example, if a rapid social change occurs that does not affect the daily routine for most people, such as the actions as well as thoughts and ideas of getting up every day to go to work, come home from work, and taking care of children and home, then it is unlikely to stimulate people to do something about the rapid social change. It is when the behavioral and cognitive routines of the everyday are disrupted that individuals are likely to be moved to act. Action is particularly likely if the routine of work is disrupted. This is because work provides both monetary sustenance that is needed for survival and dominates the waking hours.

The Great Depression is one rapid social change that severely disrupted the *quotidian* (Piven and Cloward 1977). In particular, many

individuals lost their jobs or livelihoods. As these individuals sought work and could not find it in their communities, many left their communities in search of it elsewhere. The Great Depression vastly affected the *quotidian*. The key connection between a socially disorganized environment and individuals taking action is whether the anomic conditions truly affect the individuals' lives by disrupting their everyday routines.

Once the *quotidian* is disrupted, the present world of individuals becomes a problem that begs to be fixed. This disruption makes everyday life unbearably uncertain until such time that the old routines can be reclaimed or new routines can be adopted and adjusted to. Inherently, individuals are spurred to action by the need to get back their old routines, thus reducing the uncertainty in their lives. It should be noted, though, that there may be a threshold for the loss of the *quotidian*, past which action may be unlikely. This threshold may lie at the complete destruction of the individual's way of life, such as for refugees from wars or victims surviving genocidal campaigns.

How Do Social Disorganization and the Quotidian Relate to One Another?

When Durkheim and the *quotidian* theorists were writing, they were not trying to explain the occurrence of terrorism. However, these works have been used to explain the occurrence of nonroutine collective action (Useem 1998). Nonroutine collective actions seriously contravene social norms, particularly those against violence. Terrorist attacks constitute nonroutine collective action. The acts themselves are born of the interaction and planning of a group of individuals and involve violence against people or property and, thus, are appropriately described as nonroutine collective action.

What, then, is the process of social disorganization influencing terrorism? First, the rapid social change must occur. It must occur rapidly, and it must have effects that broadly affect the lives of many people in society. An example would be the onset of a civil war. The society must not be able to absorb the negative effects of

this change. Individuals will lose their integration into the rhythm and networks of life; they will feel cut off. In addition, they will only weakly be linked to the collective identity of those in their society. These individuals are set adrift.

Second, the social change must influence the everyday lives of people. In fact, it must have substantial effects on the behavioral and cognitive routines of the everyday lives of most people. For example, if a civil war breaks out, perhaps the fighting itself has made it too dangerous for citizens to travel to work every day or perhaps the factory in which they work has been shut down due to damage from nearby bombings. Losing the routine of waking up and going to work every day as well as the thought process that goes along with such a routine may be profoundly disorienting to individuals.

Further, along with the loss of everyday routines, individuals in such societies may not know how the rules and norms of life have changed. What is acceptable and normal may have changed. Both the loss of the routine and the previously known rules and norms have effects on people. With no routine and no norms, these individuals will be poorly integrated into society. They will be included in fewer networks and activities than before the loss of the quotidian; they should also have less identification with the collective identity. The severing of networks and the collective identity and the destruction of the previous norms and rules are what classically constitutes social disorganization. This social disorganization should make it more likely that individuals will do nonroutine collective actions. They may take to the streets to protest or riot, they may commit ordinary crime, and they may engage in political violence. This political violence may be what scholars consider terrorism.

How to Define Terrorism?

The next question must be what constitutes terrorism. Definitions of terrorism abound, and the debate continues on which definition is the best. Often, the definition of terrorism chosen will reflect the main interest of the definer. For example, the United States' Federal Bureau of

Investigation, which often conducts domestic terrorism investigations, has a definition of terrorism some have called narrow and legalistic (Martin 2011). This definition is likely precise and legalistic, because it is used to investigate and enforce the law. On the other end of the spectrum, academics and researchers who perform studies of terrorism often use definitions that are broader or more detailed in order to facilitate the collection of data or analysis of data.

No specific definition of terrorism is adopted here, but some guidelines might be useful. These guidelines include non-state actors using unconventional force illegally against civilian or other noncombatant targets for political motives to influence an audience (Martin 2011). Noncombatant targets are often interpreted to include military personnel when off duty. Political motives are often interpreted rather broadly and may include acts of violence that are aimed at instigating social or religious change as well as strictly political motives. In addition, groups can have multiple or overlapping motives in that they may desire political autonomy for their social or ethnic group as well as religious change. An example of this would be the group Hamas, which seeks an independent Palestinian state that would be governed by Islamic law. Further, terrorist groups seek to influence a larger audience beyond the immediate victims of their attacks. For example, the Popular Front for the Liberation of Palestine (PFLP) undertook a campaign of dramatic aerial hijackings for the express purpose of publicizing the plight of the Palestinian people to the world. They chose aerial hijackings in order to grab the world's attention and to direct it to understanding Palestinian grievances. Similarly, the group Black September is thought to have undertaken the attack on the Israeli athletes during the 1972 Olympic Games in Munich, Germany, as a way of airing their grievances to a wider audience.

Is Breakdown a Reasonable Explanation of Terrorism?

In order for the Durkheimian model of societal breakdown to be a reasonable explanation of

terrorism, three things should be evident. First, there should be a rapid social change of sufficient magnitude and scope that it cuts individuals off from the social ties and bonds they had to the community before the rapid social change. Second, this loss of ties (breakdown) should stimulate discontent in individuals in that society. This discontent may stem from being cut off from society in general or from the quotidian. Third, it is important to demonstrate whether this discontent must make nonroutine collective action more likely to occur. If all of these qualifications can be met, then breakdown theory has passed a crucial test with respect to whether terrorism is more likely to occur in an environment of social disorganization.

Alternate Hypotheses

It is entirely possible that breakdown theory does not explain terrorism particularly well. What then are the alternate explanations? The major competing explanation is known as resource mobilization and is best exemplified by the work of Charles Tilly and his colleagues. Where breakdown posits that individuals who are cut off from the rules and norms of society by the loss of the quotidian will be more likely to engage in nonroutine collective action, the resource mobilization scholars argue that there is no connection between breakdown and nonroutine collective action. In fact, resource mobilization (RM) scholars do not recognize a distinction between routine and nonroutine collective action; rather, they assert that both routine and nonroutine actions spring from the same mechanisms (Tilly et al. 1975). They also contend that crime and suicide are not collective actions, but instead, they constitute social pathologies and ought not to be considered for explanation by the RM perspective.

Next, RM theorists argue that collective action is explained by dense social networks and a robust collective identity. That is, individuals who are well integrated into the society through their primary and secondary social networks are more likely to engage in collective action.

Primary networks are made up of family bonds and the informal social controls provided by this type of network. Secondary social networks are informal social bonds in work, school, and church. Individuals with strong primary and secondary social networks are more likely to engage in collective action according to this perspective. That is, individuals who are well connected to the people in their lives and who identify with the society's collective identity are more likely to be involved in collective action, such as social movements or protests.

More specifically, RM theorists suggest that rather than preexisting social isolation, weak social networks, and weak collective identification, collective action grows out of preexisting social organization. Social organization can include formal and informal social bonds and resources. Individuals who are well integrated into social networks, particularly secondary social networks, are a wellspring of resources, civic-minded attitude, and connections to other people. This type of connectedness makes it *easier* rather than harder to motivate collective action. In fact, individual connections to a whole group of individuals may make bloc mobilization possible, such as mobilizing church members to protest outside of abortion clinics (Useem 1998).

Taking Sides

How, then, to decide which perspective is most useful for explaining terrorism and other forms of crime? This question can be examined using several types of studies. The first type of study would be to interview individuals who have participated in collective action. This would involve collecting their personal information, such as whether they were members of secondary groups. It would also be helpful to ask whether they feel well connected to society or cut off. It would be important to examine those who engaged in routine collective action separately from those who engaged in nonroutine collective action. For example, one set of researchers examined the characteristics of African Americans in 15 cities

during the 1960s, a period of great social unrest, including both nonviolent protests and violent race riots (Miller et al. 1976). They compared those who had engaged only in nonviolent protest (routine collective action) to those who expressed attitudes supportive of rioting (nonroutine collective action). They found that those who had engaged in nonviolent protesting were more likely to be older, married, and of higher occupational and educational levels and were less likely to have been raised in “broken” homes than those individuals who had attitudes supportive of rioting. Although many of these individuals may not have participated in the riots, this study may demonstrate that those who engage in routine collective action are more likely to be well integrated in society and less likely to be socially isolated than those who engage in nonroutine collective action. This study supports the breakdown position with respect to the role of social integration and organization.

In contrast, other researchers examined the social movement against drunk driving, known as Mothers Against Drunk Driving (MADD). The movement was founded by two married mothers who had lost daughters to drunk drivers with multiple prior offenses. Interestingly, the movement seemed to take off from the beginning, with MADD’s first press conference being held months after its founding on Capitol Hill with three Congressional representatives present. MADD conducts routine collective actions, including tree plantings, gatherings at the US Capitol, and political lobbying, including lobbying US President Ronald Reagan to sign the federal legislation that would effectively raise the drinking age to 21 (madd.org 2011). One study found that on average, MADD participants were socially involved, married, middle aged, college educated, and held privileged jobs (Weed 1987). Further, another study found that MADD chapters were more likely to be founded in counties with wealthier and more educated residents (McCarthy et al. 1988). Given that MADD has only been involved in routine collective actions, this movement, the types of collective actions in which they engage, and the status characteristics of those who participate in it seem to indicate that

routine collective actions may require some social organization and participants who are enmeshed in social networks.

The second way of studying the connection between social disorganization and collective action is to examine a particular movement which involved both routine and nonroutine collective actions for evidence that the presence of social disorganization increased the discontent of citizens, whether this discontent made it more likely that these individuals would participate in collective action and, finally, whether those individuals who were well bonded to their families and their communities and participated in community activities before the rapid social change were more likely to engage in collective action (Useem 1980). This case study examined the antibusing movement which grew out of the Supreme Court decision that mandated the racial desegregation of public schools by busing students to racially integrate previously segregated schools. The study involved interviewing individuals who lived in parts of Boston which were likely to be affected by the integration of schools. They were directly asked about whether they had engaged in any collective action, their levels of discontent, and their ties to family and community. This case study directly examined the key questions about breakdown and preexisting community engagement on collective action.

The rapid social change of the antibusing movement was represented by the order to desegregate public schools. The antibusing movement in Boston involved both violent and nonviolent forms of collective action, including boycotts, demonstrations which involved violence, the formation of private schools to circumvent the desegregation order, and the formation of community organizations to represent the antibusing interests. This case study found that those who were closely tied to their communities and those who were members of community organizations were more likely to be discontented with the desegregation decree. That is, socially isolated individuals were less likely to be discontented. This supports the RM position. Those who were discontented, in turn, were more likely to participate in collective action against busing. This

supported the breakdown position. It seems possible that those who were more connected to their communities were better informed about the order to desegregate or were more likely to be affected by the order (e.g., had school-aged children) and, thus, had more of a stake in trying to protest or circumvent the desegregation order. This type of rapid social change did not affect all community members equally; those with school-aged children would be more likely to be discontented by it. Nevertheless, this study provided support for both breakdown theory and the resource mobilization theorists.

The third way of studying this question is to examine environments that have endured the effects of rapid social change that resulted in social disorganization. For example, we could also compare the levels of nonroutine collective action in states or countries that have experienced rapid social change and social disorganization to those countries or states that have not experienced rapid social change and social disorganization. This could allow researchers to examine many different countries around the world and over long periods of time. Or alternatively, we could examine the same environment over a long time period, including periods of rapid social change as well as relatively calm periods. In this way, we could examine the effects of rapid social change on collective action while holding the country environment constant.

One study examined patterns of crime and collective action in France between 1830 and 1931 (Lodhi and Tilly 1972). For collective actions, the authors included riots, strikes, and demonstrations. For the rapid social change, they examined the effects of urbanization (i.e., increases or decreases in the proportion of people in France living in cities with 10,000 or more people). Urbanization constitutes a classic rapid social change, because it generally involves the migration of rural residents to the cities. These rural residents often moved to the cities en masse in search of work, because they were unable to sustain a living in farming or because of an increase in the availability of jobs in the cities due to industrialization. The males and occasional whole families who migrated to the cities

may be predisposed to difficulties in adjusting to the new urban environment. These adjustment difficulties may have resulted in social disorganization. From the breakdown perspective, these difficulties were likely to involve the isolation of these individuals from their family and community ties, leading them to be cut off from society. This would make them more likely to engage in nonroutine collective action. From the RM perspective, urbanization itself may have no actual effects on the individuals who migrated to the cities. Rather, it is the *experience* of living in a city that makes individuals more prone to collective action. This is called urbanity, which is the proportion of people who live in cities with 10,000 or more people relative to the population of the entire country.

Thus, this research set up a direct test between breakdown and RM. The main question was whether urbanization (breakdown) or urbanity (RM) would drive either crime, collective violence, or both. During the major period of urbanization in France, the number of property crimes declined, while crimes against the person did not appear to be driven by either urbanity or urbanization. These findings most clearly contradict the breakdown perspective, which suggests that urbanization should drive both person and property crimes upward. On the other hand, urbanity seemed to predict property crime and some of the forms of collective violence. That is, the experience of living in a city was more clearly responsible for increases in property crime and some forms of collective action than the social disorganization experienced by a society which has en masse migrations of its citizens to the cities.

Another way of testing which theory better explains how collective action is produced is to study one social environment over a long period of time to assess whether crime and collective action occur during the same periods of time. The RM theorists suggest that crime is not a collective action but rather a social pathology, and thus, it should have no relationship with routine collective action. In contrast, breakdown theorists say that crime and nonroutine collective action ought to occur at similar times, because they are driven by similar processes (social

disorganization). However, routine collective action and crime should not occur alongside one another because they are not driven by similar processes.

LaFree and Drass (1997) looked at this question before, during, and after the civil rights movement in the United States. Unfortunately, the study did not separate routine and nonroutine collective actions. The authors compared homicide, robbery, and burglary arrest rates for blacks and whites to collective actions. These collective actions were primarily a result of the African American civil rights movement. Thus, they included such actions as sit-in rallies, marches, boycotts, protests, demonstrations, and civil-rights-related riots.

The idea is to examine whether crime and collective action follow similar patterns over time. If they do, it can be inferred that they are being caused by the same or similar processes, and this would support the breakdown perspective, because if they are predicted by the same processes, then, it is likely that collective action does not require preexisting community ties and membership in collective groups (Lodhi and Tilly 1972). If they do not follow the same pattern over time, increasing and decreasing at similar times, then, it is unlikely that they are driven by the same processes. This type of finding supports the RM model.

Upon examination, crime and collective action did follow a similar path, increasing and decreasing together from 1955 until the early 1970s for both blacks and whites. However, collective action decreased dramatically at this point as the civil rights movement, the rapid social change, came to its conclusion. However, arrest rates stayed high and continued to follow their own path from the early 1970s until 1991. These findings do not clearly support either breakdown or RM. It is possible that if nonroutine collective actions, such as riots, and routine collective actions had been examined separately and each compared to crime that the findings would have been more clearly supportive of one side or the other.

Overall, there is no clear winner in the debate between breakdown and RM. It is likely that both theoretical explanations work to explain different

segments of collective action. For example, it is likely that preexisting social organization, such as membership in community groups and having personal ties to the community, makes routine collective action more likely, such as the MADD movement against drunk driving. Individuals who have a stake in the community are probably more likely to act to defend that community using legal and nonviolent methods when they perceive a change, like in the antibusing movement. However, nonroutine collective action likely does not require such organization. Individuals who are more likely to engage in nonroutine collective action, such as collective violence or riots, appear to be more socially isolated individuals who are freed to act in response to the socially disorganized environment which surrounds them. Thus, each of these theoretical explanations appears to explain different domains of collective action.

Is Social Disorganization Felt by Citizens?

It is likely that the social disorganization created by the rapid social change is felt by those who live in the society and are affected by the changing environment. This may be felt as a change of or removal of the norms, a feeling that the old rules no longer apply (Durkheim 1930 [1951]). This may also be felt from the loss of the everyday routine (Snow et al. 1998), which may be experienced as a feeling of restlessness or rootlessness. Further, this may also be felt as impending danger or a loss of a sense of safety (Useem 1998).

In an interesting but simple test of this notion of a feeling of impending danger, Useem (1998) examined the relationship between homicide rates, domestic handgun production, and rioting activity in the United States. The time period was 1964 until 1994, a time which included the African American civil rights movement as well as race-related riots in major cities. The thinking was that handgun production would increase as the handguns previously manufactured were purchased; handguns are not usually used for

hunting, unlike long guns. Handgun production rates closely tracked the homicide rate over the time period, except for those years in which there were increases in riots. During these years, handgun production rates jumped up more than would be expected given the danger presented by the homicide rate. Thus, individuals appeared to be frightened by the occurrence of these riots and purchased handguns to protect themselves even if the riot did not occur geographically close to them. This research supports the notion that individuals can feel social disorganization as a sense of impending danger.

Does the Evidence Show a Relationship Between Social Disorganization and Terrorism?

The final question is whether terrorism is more likely to occur in socially disorganized environments. One interesting research study examined this question in the context of breakdown theory (Fahey 2010). In this study, the socially disorganized environment resulted from the occurrence of political instability within the country. The instabilities included the outbreak of war, genocide, adverse regime change, which was a move toward a more autocratic government, and a combination of these events occurring within a short time period. The occurrence of these instability types can be seen as measures of rapid social change. In addition, these types of events are also likely to affect the everyday lives of individuals. If the security situation in a particular country deteriorates rapidly, as in the case of war and genocide, or the political environment has suddenly become far less open, as in the case of adverse regime change, citizens will likely be less able to continue on with the rhythm and routine of their normal everyday life, such as work, leisure, and home time. In extreme cases, they may be forced out of their homes. In addition, the political instability may create a breakdown in the informal and formal social ties that bind individuals to society; the breakdown in the ties may make it more likely that individuals living in a society which has

experienced this political instability will participate in nonroutine collective action. This will be due to the idea that the informal and formal ties can no longer function to restrain the actions of these individuals. Further, the loss of the everyday routine and the loss of the controlling power of social ties may work together to increase the likelihood of nonroutine collective action.

Fahey (2010) examined whether political instability was related to the occurrence of terrorism in 147 countries from 1970 to 2005. She found that when political instability occurred within a state, increases in terrorism incidents were more likely. One of the strengths of this study is that it examined the occurrence of instability worldwide over 35 years. This helps to ensure that the results are more stable, that they do not apply only to one country or situation, and that they do not apply to only one decade or time period. Political instability does seem to constitute a rapid social change in the Durkheimian model, and it likely affects the regulatory processes in society, that is, social disorganization, as well as the everyday routines of the citizens in that society.

Conclusions

The purpose of this entry was to discuss and explain the rapid social changes that may lead to social disorganization, the two dominant theoretical hypotheses to explain the association between social disorganization and collective action, and some of the evidence that has been used to support either or both of the theoretical expectations. More evidence is needed to come to a firm conclusion on the relationship between socially disorganized environments and terrorism. However, it is likely that both theoretical explanations have some utility. Specifically, it is likely that some individuals and groups are spurred to terrorism due to the loss of prevailing norms in a newly socially disorganized environment as well as the loss of identification with the collective identity. Further, it is also likely that those who had connections to others, particularly through community groups, before the onset of

the social disorganization will be in a better situation to join, recruit, and mobilize for and engage in terrorism through a terrorist group. Bloc mobilization may be particularly important for terrorist groups. What seems to be clear is that social disorganization may be one possible explanation of terrorism incidents. For terrorism that occurs after a rapid social change that produces social disorganization, it is likely that the disorganization provides at least a partial explanation.

Recommended Reading and References

- Durkheim E (1930 [1951]) *Suicide: a study in sociology* translated by George Simpson and John A. Spaulding. The Free Press, New York
- Fahey S (2010) *State instability and terrorism*. Doctoral dissertation, University of Maryland, College Park
- Gurr TR (1976) *Rogues, rebels and reformers*. Sage, Thousand Oaks
- Hehir A (2007) The myth of the failed state and the war on terror: a challenge to the conventional wisdom. *J Int Statebuilding* 1:307–332
- Helman GB, Ratner S (1992) Saving failed states. *Foreign Policy* 89:3–20
- LaFree G, Dugan L, Fahey S (2008) Global terrorism and failed states. In: Hewitt JJ, Wilkenfeld J, Gurr TR (eds) *Peace and conflict. Paradigm*, Boulder
- LaFree G, Drass K (1997) African American collective action and crime, 1955–1991. *Soc Forces* 75:835–853
- Lodhi AQ, Tilly C (1972) Urbanization, crime and collective violence in 19th century France. *Am J Sociol* 79:296–318
- Madd.org (2011) 25 years of saving lives, by Laurie Davies <http://www.madd.org/about-us/history/madd25thhistory.pdf>. Accessed 21 July 2002
- Martin G (2011) *Essentials of terrorism: concepts and controversies*, 2nd edn. Sage, Thousand Oaks
- McCarthy JD, Wolfson M, Baker DP, Mosakowski E (1988) The founding of social movement organizations: local citizens' groups opposing drunk driving. In: Carroll GR (ed) *Ecological models of organizations*. Ballinger, Cambridge, MA, pp 71–84
- Miller AH, Bolce LH, Halligan MR (1976) The new urban blacks. *Ethnicity* 3:338–367
- Newman E (2007) Weak states, state failure and terrorism. *Terror Polit Violenc* 19:463–488
- Piazza J (2007) Draining the swamp: democracy promotion, state failure, and terrorism in 19 Middle Eastern countries. *Stud Conflict Terrorism* 30:521–539
- Piazza J (2008) Incubators of terror: do failed and failing states promote transnational terrorism? *Int Stud Q* 52:469–488
- Piven FF, Cloward R (1977) *Poor people's movements. Why they succeed, how they fail*. Pantheon, New York
- Snow DA, Cress DM, Downey L, Jones AW (1998) Disrupting the quotidian: reconceptualizing the relationship between breakdown and the emergence of collective action. *Mobilization* 3:1–20
- Tikuisis P (2009) On the relationship between weak states and terrorism. *Behav Sci Terror Polit Aggress* 1:66–78
- Tilly C, Tilly L, Tilly R (1975) *The rebellious century, 1830 to 1930*. Harvard University Press, Cambridge, MA
- Useem B (1980) Solidarity model, breakdown model and the Boston anti-busing movement. *Am Sociol Rev* 45:357–369
- Useem B (1998) Breakdown theories of collective action. *Annu Rev Sociol* 24:215–238
- Von Hippel K (2002) The roots of terrorism: probing the myths. *Polit Quart* 73:1–39
- Weed FJ (1987) Grass roots activism and drunk driving issues: a survey of MADD chapters. *Law Policy* 31:259–278

Social Factors and Third Party Policing

Daniel Brown-Kenyon
Institute for Social Science Research, Policing and Security, The University of Queensland, St. Lucia, QLD, Australia

Overview

The Third-Party Policing model of crime control advocates for the use of partnerships between the police and other key agents of social control – for example, government, business, and community sectors – however, to date, almost no research has examined the social factors that contribute to the success or failure of these partnerships. This entry utilizes a Social Identity approach to analyze some of the key factors that affect the efficacy of these partnerships. Specifically, partnerships are broken down into three distinct phases, and various social factors are considered at each of these phases. This entry also discusses some of the organizational level variables that will be affected if these partnerships are to succeed.

Introduction

Third-Party Policing (TPP) is an approach to dealing with crime problems where police form

partnerships with other public service agencies, business, nongovernment organizations, and the community to deliver more effective crime control strategies (Mazerolle and Ransley 2006). The TPP approach requires identification of “legal” mechanisms that are the domain of other parties to assist police in managing public order. Mazerolle and Ransley (2006) identify a wide range of legal levers used in a variety of different settings, such as partnerships with local councils, business, health and safety authorities, and other law enforcement agencies. The primary aim of the TPP approach is to deliver a more effective police service through a twofold approach: recruiting suitable crime control partners and making use of legal levers otherwise unavailable to the police.

This contribution aims to advance the theoretical understanding of TPP by outlining current research and theory on intergroup relations and applying this to policing partnerships. TPP advocates that partnerships are vital for the success of modern policing practices. However, to date, little to no research has been conducted within the policing arena to examine these partnerships and what drives their success and failures. Drawing on the Social Identity literature (see Brown 2000; Haslam 2004), this entry examines some of the social and organizational factors that interact with the success or failure of these partnerships. This analysis is undertaken in five sections, which are comprised of two conceptually distinct areas, focusing on the social aspects of partnerships and then outlining the organizational factors that affected partnerships. The first section presents an overview of a Social Identity approach to intergroup process. The next section separates partnerships into three core stages, including partnership seeking, partnership development, and effective partnership functioning. Next, the role of information sharing in partnership development and functioning is addressed. Finally, some of the organizational constraints within police organizations that are likely to threaten the success of a partnership approach are addressed. At the end of each section, the arguments are summarized and important relationships between the key factors and likely outcomes are discussed.

A Social Identity Approach to Intergroup Process

Social Identity theory is the preeminent theory of intergroup relations. Social Identity theory explores the social motivations and consequences of group behavior. Specifically, Social Identity theory provides a context for examining individuals’ behavior within groups (Haslam 2004). According to Social Identity theory, individuals possess social representations of themselves, which are comprised of the social networks they maintain, the groups to which they belong, and the organizations for which they work (Ellemers et al. 2002). Social Identity theory argues that social identity is a critical component of self-concept, and as such, individuals seek to view their group memberships in a positive frame. This is achieved through negative comparisons between their own group and other groups. Through these comparisons, individuals are able to boost their self-esteem (Brown 2000) and show preferential treatment to those of their own groups (Brown 2000). Social Identity theory provides a meta-level context to analyze and understand a range of phenomena involved in intergroup processes.

Since the conception of Social Identity theory, there has been substantial development with a range of subtheories having been developed to explain some of the more nuanced aspects of group behavior. This entry applies a Social Identity approach to understanding the three hypothesized stages of partnerships. The first stage, termed “partnership seeking,” helps to understand the internal group structure of the police organization and how these social factors impact on partnership-seeking behavior. Here, key factors of cohesion, leadership, and the norms and values within the organization are all likely to affect the willingness of individuals to seek out partnerships. In the next phase, termed “partnership formation,” the core factors of diversity, distinctiveness, and identification during the formation of a new partnership are considered. In the final stage, “partnership functioning,” intragroup processes and their implication on cooperation are considered. Here, group identification, procedural justice,

and leadership and the impact of these factors on cooperation within a newly formed partnership are discussed.

Partnership Seeking and the Role of Social Identity Theory

This section uses Social Identity theory to examine three internal social factors that police organizations face when seeking TPP partnerships. First, the way that norms and attitudes of a police organization are likely to dictate individual officers' intentions to seek out third-party partnerships is examined. Within this, the role of self-categorization and its impact on conformity with these norms are also outlined. Second, the role of leadership within a police organization in promoting partnership seeking and development of partnerships within the context of TPP is discussed. Finally, this section explores the effect of internal cohesion and how this poses a significant barrier for the police to effectively seek out partnership.

The development of self-categorization theory is an extension of Social Identity research and is particularly useful for understanding partnership-seeking behavior (Hogg and Terry 2000). Self-categorization theory stipulates that individuals categorize their social world by the groups to which they belong, through a process of depersonalization, in which people are reduced to a prototype of the group, thereby embodying the norms, values, and behaviors that personify the group (Hogg and Terry 2000). It is through this process of self-categorization that cognitive groups are formed. Depersonalization results in cognitive perceptions of similarity between individuals and their group. These effects are emphasized for individuals of out-groups, such that when they are perceived as a prototype of the group, the individual differences between group members are reduced and the similarities are enhanced. The outcome of this process is that individuals are no longer perceived as unique, rather, they are perceived to be a representative of their group (Hogg and Terry 2000).

Police are often faced with situations in which they categorize people into groups. This process allows the police to distinguish all members of a group – for example, a child protection agency – as being conceptually the same. Furthermore, the effect of self-categorization will be more salient. That is, the officer's identity as a police officer will be more powerful, when officers are faced with the possibility of sharing aspects of their work that makes them unique, such as crime control. Evidence suggests that a critical component of group discrimination is the dimensions that the groups are compared on (Ashforth and Mael 1989). As managing law and order is a key dimension distinguishing police from other professionals, contexts requiring police to share this role with others will likely result in increased categorization of those others. This effect is likely to result in the police taking steps to demonstrate the inability of their partners to fulfill the role that they have been assigned. Empirical research supports this position, with results consistently demonstrating that self-categorization effects are most pronounced in an intergroup context (Ashforth and Mael 1989; Doosje et al. 2002). An example for policing may include a collaborative taskforce where the police share the responsibilities with a partner, such as a child protective service, who has the same power to act and responsibilities as the police. Therefore, it is possible the police would attempt to exasperate the difference between the two organizations with an emphasis on the positive aspects of police and the negative aspects of the child protection agency. This would likely result in the police demonstrating reluctance to share their responsibilities with this other agency through a distrust of their ability to deliver the required services.

Research also indicates that the degree of identification felt toward an organization is a critical factor in conformity to group norms, attitudes, and behaviors (Tyler and Blader 2001). As individuals seek to enhance both their own status within the group as well as the overall standing of the group, they are highly motivated to perform in a manner that increases their standing in both domains. Conforming to norms and attitudes is one method to achieve this. In doing

so, individuals represent the prototype of the group, while demonstrating active commitment to the group. Conversely, individuals face sanctions for failing to conform. Therefore, the degree of self-categorization and meaningful value that an officer places on their position within the organization will also be predictive of their willingness to conform to the organizational rules, norms, and behaviors.

This suggests that the organizational norms and attitudes of police organizations need to foster partnership-seeking behavior if a police agency is to be successful in institutionalizing TPP. Police organizations may need to undergo a cultural change, with a strong focus on the involvement of partnerships in crime control innovation. In these situations, the self-categorization process will continue to reinforce these norms and attitudes, thereby ensuring that new police officers will also engage in this behavior.

Self-categorization theory argues that group leaders are well positioned to change the culture and develop new attitudes. According to self-categorization theory, leaders are perceived to be members of a group that are *most* prototypical (Hogg 2001), the individual(s) that embody the values, norms, and attitudes of the group (Hogg 2001). Furthermore, self-categorization theory argues that leaders have the consent of their followers to change the cultural attitudes and norms of a group. Leaders within an organization hold the unique position to install change and create a culture of active partnership formation. They hold unique positions where it is legitimate for them to deviate from the accepted cultural norms to install change. Research has demonstrated this critical role of leaders in conveying of norms (Hogg et al. 2006). Likewise, research has also consistently demonstrated the importance of leaders in fostering innovation and supporting change. Therefore, the onus is on police leaderships to create and foster norms and attitudes that will promote active partnership seeking and development among the junior ranks.

An outcome of highly salient group identity and strong leadership can lead to high levels of cohesion within groups (Hogg 1993). Cohesion is

defined as the social connections that tie individuals to their fellow group members, resulting in their desire to behave and act in a manner consistent with the group (Casey-Campbell and Martens 2009). Research has extensively examined cohesive groups, with an emphasis on the cohesion-performance relationship. There is mixed evidence supporting the cohesion-performance relationship. However, there is general support for the notion that highly cohesive groups perform better on tasks, particularly in project-based team environments (Chicchio and Essiembre 2009).

The nature of police work lends itself to high levels of internal cohesion. The police are required to rely on each other for safety and to work together to effectively achieve outcomes. Often they are required to sacrifice their own safety and well-being for the benefit of society. Anecdotally, police organizations are akin to military organizations, where there is a cultural of solidarity and unity. However, when considering a TPP initiative, highly cohesive police organizations pose a significant constraint on the take-up of TPP partnerships. Research indicates that highly cohesive groups are more reluctant to engage with external groups in intergroup contexts (Hardy et al. 2005). Therefore, excessively cohesive police organizations are likely to prevent officers from seeking partnerships.

It is possible that highly cohesive organizations will attempt to develop internal solutions, instead of seeking partnerships. This results in a continued drain on resources, without commiserate efficiency gains. When advocating for TPP partnerships, the role of cohesion and its effects on the success of new initiatives may need to be considered. Evidence for consideration of these effects can be seen from previous investigations into policing practices. For example, one of the key recommendations from an inquiry into police corruption in Australia was that for new recruits to be educated to a minimum of a tertiary level in an attempt to reduce the insular nature of policing (Fitzgerald 1989). This recommendation was made in an attempt to prevent the effects of a highly cohesive organization perpetuating bad practice.

This section has highlighted some of the internal social factors that police organizations face when seeking TPP partnerships. From the research outlined, the following relationships are proposed. First, the norms and attitudes of a police organization are likely to affect officer's intentions to seek out third-party partnerships. Furthermore, it is possible that this relationship will be affected by the role of self-categorization, such that those who are highly identified may demonstrate greater conformity to these norms. Secondly, it is the role of the leaders within an organization to promote active search and development of partnerships. Finally, high levels of cohesion in police organizations are likely to impede effective partnership seeking.

Partnership Formation: Two Organizations Merging and the Role of Diversity and Distinctiveness

Organizational mergers have long been the topic of interest for a range of business and academic communities. This research has largely focused on the outcomes of a newly formed organization (Terry and O'Brien 2001); however, recent research has also considered the social factors that contribute to the success or failure of partnerships (Terry et al. 2001). Policing partnerships are akin to other types of organizational mergers. They require at least two different organizations to work toward a unified goal and deliver a product – in this case public safety. Research by Terry et al. (2001) has investigated the role of social identity during the formation of a new superordinate organization after a merger. This research indicates a range of social factors that should be considered when forming a TPP partnership. Primarily, during the initial merger, the higher-status group members react more negatively toward the merger than lower-status group members (Terry et al. 2001). The implication of this finding is critical to understanding the police motivations to form a partnership. It is arguable that the police will hold the higher-status position within most new partnerships; they will typically be the organization that initiates the partnership and the organization that has the most to gain.

The inverse of the status relationship is also true. There is evidence to suggest that individuals from lower-status organizations are more willing to join higher-status organizations (Terry et al. 2001). However, some results have found that the effect of status can be moderated by legitimacy, such that, if lower-status organizations perceive their status as illegitimate, they will be less likely to join the higher-status organization (Amiot et al. 2007). When considering these findings in the context of a TPP partnership, it is clear that steps need to be taken to mitigate these effects in order for these diverse organizations to work together. This indicates that the police need to ensure that they do not portray themselves as superior to their partner organization as well as approach their partners in a collaborative and transparent manner that avoids invoking perceptions of status differences between the two groups.

Research indicates that lower-status organizations have higher staff turnover rates after merger (Amiot et al. 2007). This research also suggests that lower-status group members have less identification with new organization (Amiot et al. 2007). Together, these two findings are vital to the understanding of how partnerships function. As highlighted in the first section, the key to ensuring that individuals perform for the group, and adhere to the group norms of behavior and attitudes, is increasing the identification that they feel toward the group. Therefore, if lower-status individuals feel less identification toward the partnership, they are less likely to remain in the partnership. If the police believe that they are a higher-status organization and their key partners – such as child protection agencies – are a lower-status organization, then this is likely to result in initial reluctance from the police to form a partnership. However, should this be overcome, issues of status legitimacy, identification with the partnership, and turnover rates are all likely to be significant challenges for both parties.

When a new partnership is formed, the police and their partners come together under the banner of a new group. This partnership is formed to provide a more efficient service and reduce the resource burden from one agency alone.

However, there is likely to be historical issues of status, prestige, and potentially conflict that need to be managed. Furthermore, issues of group diversity pose significant problems to the formation of these partnerships. Diversity within organizations has long been a topic of significant interest to the academic and business community (Brewer 1995). In the context of a TPP partnership, it is important to consider diversity within the context of police and the partner agencies. These partnerships involve individuals from different groups, with different experiences/skills and different organizational cultures. Furthermore, these individuals will come from organizations and skill backgrounds, which have different underlying philosophies and ideologies about how to manage complex crime problems. The research outlines the role of different approaches to problems and the impact of these cultural ideologies on the success of group performance.

Diverse partnerships have the potential to foster conformity to organizational norms and rules, as these become the superordinate guidelines to behavior within the given context (Rink and Ellemers 2007). To enable this, organizational identities need to be highly salient and provide significant emotional value to its members. In contrast, in partnerships where there is a history of conflict and highly charged emotions, the organizational identity will be weak, thereby making the subgroup identities highly salient (Tsui et al. 1992). This results in reduced compliance to the norms and rules of the partnership and a return to the norms and values of their originating organizations (Rink and Ellemers 2007). Forming a successful TPP partnership requires that all parties are aware of organizational diversity both within and across the partnering organizations. Historical conflict between partners poses a significant constraint on the success of these partnerships. Furthermore, by identifying common goals and objectives and reinforcing the mutual commitment of the organizations, it is possible to overcome these issues.

Research in the area of optimal distinctiveness theory helps to explain the conditions under which partnership identification is achieved. Optimal distinctiveness theory (Brewer 1991)

argues that we have two competing drives in social groups – one for inclusion and one for distinctiveness – and that these two drives act in opposition to each other. In the context of a new partnership, individuals will want to be included in the partnership, to feel they are a valued member and that they are able to contribute to the development and direction of the partnership. However, individuals will simultaneously want to maintain connections with their original organization. Therefore, the successful formation of a new partnership is contingent on both of these needs being met at an optimal level (Brewer 1995).

For the police to form a partnership with key partners, a significant risk arises from these organizations being identified under the banner of a large and ambiguous group, typical of “whole of government approaches,” as these groups remove the unique contributions of the individual organizations that comprise the partnership. Likewise, the assimilation of groups under the higher-status group, in this situation the police, poses the same risks. These partnerships need to develop identities that are unbiased and do not represent just one group, yet attempt to encompass the root values and goals of the partnership. Research shows that group size and overly inclusive groups are both negatively related to performance (Badea et al. 2010). According to the optimal distinctiveness literature, this is likely to result in poor functioning and lack of identification at the superordinate level (Badea et al. 2010) and likely to result in high turnover or a quick end to the partnership. Therefore, these partnerships should avoid the use of group structures that inhibit the unique contribution of each member organization.

This section highlighted the role of several key factors that inform the process and the likelihood of police forming partnerships within the context of TPP. The role of group diversity, distinctiveness, and partnership identity have all been examined and discussed. Next several relationships between these variables are examined. First, partnerships where status difference between the police and a partner organization is highly salient are likely to result in poor intergroup cooperation

and weak identification with the partnership. Secondly, diversity of organizational identities in the partnerships will impact on partnership functioning. Specifically, partnerships with historical conflict and competition for resources will not function effectively as a partnership; this effect may be exasperated by differences in cultural ideology about how to approach problems. Finally, the success of a new partnership developing is contingent on the ability for individuals to retain connection with their originating organization.

Promoting Cooperation Within a Partnership: The Role of Identification, Procedurally Fair Decision Making, and Leadership

TPP partnerships require the active and continued cooperation of its members. Significant research has explored the motivations that promote cooperation within groups, with a specific focus on information sharing, extra-role behavior, and ensuring that the long-term goals of the organization are met over the temptation of short-term individual gains. Recently, Tyler (2011) examined the relationship between instrumental factors such as material rewards and role titles, and social factors such as identification and procedurally fair decision making, in fostering cooperation of employees with the organization. The results consistently demonstrated that above and beyond instrumental factors, social motivation was found to be critical in adherence to organizational rules and norms, extra-role behavior, and intragroup cooperation (Tyler 2011). Extra-role behavior is defined as undertaking work and responsibilities that are outside the individual's job description for which they do not receive direct momentary reward.

In this section two critical aspects of cooperation within partnerships are highlighted: identification with the partnership and procedural fairness of decision making within the partnership. As discussed in the earlier sections, when an individual categorizes themselves as a member of an organization, this will lead to increased adherence to the norms and attitudes of the organization (Tyler and Blader 2001). Therefore, increasing the degree of social identification

with the partnership will lead to an individual feeling greater emotional significance about their connection with the group, thereby increasing their drive to enhance the group relative status. TPP partnerships require a significant commitment from both the individuals and the groups connected to them. This commitment can range from undertaking activities that are beyond the typical duties of the role they fulfill to providing extra resources to ensure the success of the partnership. By increasing the degree of identification to the partnership, there is potentially an increase in the adherence toward the goals and aims of the partnership. Furthermore, this is also likely to result in the members undertaking extra-role behavior when faced with impediments to ensure the success of the partnership (Tyler 2011).

According to group engagement model (Tyler and Blader 2003), identification with the organization is enhanced by the decision-making processes that the organization uses. Specifically, the use of procedurally fair processes to arrive at decisions increases the strength of identification that individuals feel toward their organization. Research demonstrates a link between the procedural fairness of an organization and the behavior of its employees (Blader and Tyler 2009). Procedural fairness can be described as the processes and methods that an individual or group enacts to arrive at a decision as being fair (Blader and Tyler 2003). There is consistent evidence within a range of fields that demonstrate the link between procedural fairness and cooperation, acceptance of outcomes, and extra-role behavior (Tyler and Blader 2001). Furthermore, evidence suggests that people are more concerned with the procedural fairness of a decision rather than the outcome of a decision alone. The goal of a new partnership, therefore, is to behave in a procedurally fair manner, as perceived by all members of the group. This is achieved through unbiased decision making, offering all members a voice before a decision is made, making the decision on facts rather than opinions, and using clear and transparent processes. Therefore, the relationship between procedural justice and cooperation in the group is mediated through identification with the group (Tyler and Blader 2003).

For TPP partnerships, this finding implies that it is possible to increase the degree of identification that the individual members feel toward the partnership, by utilizing procedurally fair decision-making process.

The final aspect of effective partnership functioning relates to leadership. As discussed earlier, the role of a leader can be of critical importance in the conveying of norms, attitudes, and appropriate behavior (Hogg 2001). When a leader behaves in a manner that embodies the values of the group and is seen as the most representative member of the group, then it is likely that they will be able to gain the trust and confidence of their followers (De Cremer and Van Knippenberg 2002). Likewise, there is sufficient evidence to suggest that a leader that acts in a procedurally fair manner is also likely to enhance their subordinates' willingness to accept decisions (De Cremer and Van Knippenberg 2002). Extending this, research has indicated that leaders play a vital role in the development and fostering of innovative practices (Hogg et al. 2006). Therefore, it is vital that upon formation of the partnership, a leader who will drive the strategic direction, yet encompass the requirements outlined above, is identified. As outlined in earlier sections, leaders play a vital role in the success of any venture; they provide the strategic direction (Hogg et al. 2006) and can guide the subordinates through challenging times. The success of new partnerships hinges on the effectiveness of the leader.

Three critical aspects of a partnership define its functioning. First, individuals who are part of the partnership must demonstrate high levels of identification with the new partnership, as doing so will ensure compliance and cooperation with the partnership objectives. Second, in order to increase identification, the partnership must demonstrate procedurally fair processes in its decision making, with a particular emphasis on demonstrating collegial and collaborative approach and avoiding unilateral decision-making processes. Finally, the leadership of the partnership must represent all parties within the partnership and actively engage in procedurally fair process.

Organizational Constraints of Third-Party Policing

In their analysis of TPP approaches to policing complex problems, Mazerolle and Ransley (2005) outline that police organizations face organizations' constraints that are likely to impact on the ability of the police to form partnerships. When advocating for partnerships, these factors need to be considered in conjunction with the social factors outlined above. One of the most crucial factors that require consideration relates to the ability of organizations and government to effectively share information. This is vital, as without sharing of intelligence, methods, and policies, the ability of partner organizations to coordinate the delivery of their services becomes restricted. Furthermore, it may lead to redundancy and duplication of work or frustration for lack of action from partners.

The capacity to share information between partner organizations is one of the most complex factors to be overcome to facilitate a productive working partnership. When attempting to find innovative solutions to problems, one of the most important tools that can be employed is sharing information across organizational boundaries. However, when government agencies are involved in these partnerships, the transfer of information between partners can pose significant problems. Most Western governments have laws that restrict the transfer of information (Bellman et al. 2004). These restrictions typically impose conditions that prevent the transfer of information between government agencies. These laws are designed to protect the interests and privacy of individuals within the community. However, they can also prevent the effective coordination and cooperation of core agencies. Therefore, one of the most significant structural challenges that these partnerships face is the need to overcome information sharing hurdles (Lyons 2002). It is important to consider what information needs to be shared, how the information will be used, and who, under relevant laws, is able to hold that information. Addressing these considerations early in the formation of a partnership is critical to the success of a partnership.

Mazerolle and Ransley (2005) identify a number of organizational factors that either facilitate or constrain the adoption of Third-Party Policing. One of the most significant challenges that police agencies face in forming partnerships is the centralization of decision making. They argue that the concentration of decision-making power with the executive is not conducive to the formation of partnerships. The crux of their argument is that officers at all levels of the organization need to have the freedom to identify problems and innovate solutions, and doing so can only occur effectively when there has been a decentralization of decision making (Mazerolle and Ransley 2006). Furthermore, decentralization of decision making offers unique advantages to geographically dispersed police agencies. For example, issues that affect police forces in large metropolitan areas are likely to be significantly different from those that affect police officers in rural or remote areas. Therefore, officers need to be able to identify local solutions to local problems. This argument is not unique to TPP advocates. Problem-oriented policing, arguably one of the most influential developments in policing practice, has advocated for support from the senior executive for officers to implement localized strategies to complex crime problems (Goldstein 2003).

Finally, Mazerolle and Ransley (2006) outline that role specialization has a significant impact on the ability of officers to form partnerships, stating that more specialized roles are conducive to the formation of specific partnerships than roles that are more general. This argument follows the notion that officers and personnel that fulfill highly specialized roles within an organization will have greater insight into the needs from the partners and the ability of the partners to engage in a cooperative approach. Extending this argument, the seniority of officers will also have an impact on their ability to develop and formalize partnerships, with senior officers having the ability to implement formal cross-organizational partnerships and junior officers seeking ad hoc episodic partnerships. Therefore, identification of the right personnel within the police organization and gaining the support of

the senior executive are factors that may well contribute to the success of potential partnerships.

Conclusion

In an era of social change in the way that the police conduct their duties, there is inherent value in a TPP approach to crime control. Through the use of effective partnerships, the police are able to provide a more efficient and effective service that has the potential to develop more holistic outcomes for the targets. This entry examines the partnership aspect of TPP using Social Identity theory to explore the intergroup processes that affect the ability of the police to seek out and develop crime control partnerships. Three critical phases comprise the process of partnership building within TPP: partnership seeking, partnership formation, and partnership effectiveness.

In the partnership-seeking phase, research outlined the role of the internal social structure of the police organization, with an emphasis on the norms and attitudes, to ensure that they are focusing on the ultimate goal of the organization. In conjunction with this, the role of internal cohesion within a police force was considered, with a particular emphasis focusing on what impact this has on the ability and desire of the organization to seek out partnerships. Research also examined the role of leadership in shaping the future direction of police agencies, with an eye toward fostering partnership-seeking behavior. In the formation phase of TPP partnerships, the research outlined indicates that through the consideration of the diversity of ideological backgrounds and organizational culture, it is possible to improve working relationships. Extending this, allowing individuals to join and have meaningful input into the future directions of a partnership while still allowing them to maintain a strong connection with the organization where they originated is another avenue to success. In the final stage, partnership functioning, the research outlined above argues that it is critical to help the individual form a sense of identity with the

partnership in order to foster cooperation. It is possible to achieve this with procedurally fair decision-making process, with a strong emphasis on the use of unbiased, fair, and neutral process. Engaging and developing a strong and consistent leader is also important in fostering a strong sense of identity with the new partnership.

While this entry has largely focused on the importance of social factors in the formation and functioning of a partnership, there are also instrumental factors that contribute to the success and failure of partnerships. Information sharing is a significant challenge to the way that partnerships can function, with the inability of some information to be shared between partners potentially leading to duplication of work and frustration with partners. Likewise, the decision-making ability of individuals within an organization represents challenges for partnerships. In situations where all decisions need to be passed through the senior executive, this has the potential to limit the ability of officers to seek and contribute to partnerships. Finally, officer specialization is likely to impact on the quality and type of partnership, with more specialized officers having greater insight in the needs of their own organization in respect to specific issues but also insight into how partners function their ability to contribute to a coordinated effort.

The issue of partnership formation is not one that is solely the responsibility of the police. All key stakeholders, including government, policy makers, communities, and businesses have a role in overcoming the challenges faced in addressing complex crime. The police face many challenges in attempting to develop their practices in a new era of policing; therefore, drawing attention to these challenges provides opportunities for further research and input from policy makers and practitioners alike. If the police and society more generally are going to deliver a more effective service, then it is incumbent on the community to work collaboratively to develop solutions. As we learn more about how and why crime is committed, it is no longer feasible to assume that insular police organizations will be able to recruit and retain the knowledge, skills, and resources required to combat future criminal behavior.

Related Entries

- ▶ [Civil Remedies](#)
- ▶ [Communities and the Police](#)
- ▶ [Communities That Care](#)
- ▶ [Community Policing](#)
- ▶ [Legal Frameworks for Third-Party Policing](#)
- ▶ [Procedural Justice and Cooperation](#)
- ▶ [Third Party Policing and School Truancy](#)

Recommended Reading and References

- Abrams D, Hogg MA (2004) Metatheory: lessons from social identity research. *Pers Soc Psychol Rev* 8(2):98–106. doi:10.1207/s15327957pspr0802_2
- Amiot CE, Terry DJ, Callan VJ (2007) Status, equity and social identification during an intergroup merger: a longitudinal study. *Br J Soc Psychol* 46(3):557–577. doi:10.1348/014466606x146015
- Ashforth BE, Mael F (1989) Social identity theory and the organization. *Acad Manage Rev* 14(1):20–39
- Badea C, Jetten J, Czukor G, Askevis-Leherpeux F (2010) The bases of identification: when optimal distinctiveness needs face social identity threat. *Br J Soc Psychol* 49(1):21–41. doi:10.1348/000712608x397665
- Bass BM (1997) Does the transactional–transformational leadership paradigm transcend organizational and national boundaries? *Am Psychol* 52(2):130–139. doi:10.1037/0003-066x.52.2.130
- Bellman S, Johnson EJ, Kobrin SJ, Lohse GL (2004) International differences in information privacy concerns: a global survey of consumers. *Inf Soc* 20(5):313–324. doi:10.1080/01972240490507956
- Blader SL, Tyler TR (2003) A four-component model of procedural justice: defining the meaning of a “Fair” process. *Pers Soc Psychol Bull* 29(6):747–758. doi:10.1177/0146167203029006007
- Blader SL, Tyler TR (2009) Testing and extending the group engagement model: linkages between social identity, procedural justice, economic outcomes, and extrarole behavior. *J Appl Psychol* 94(2):445–464. doi:10.1037/a0013935
- Brewer MB (1991) The social self: on being the same and different at the same time. *Pers Soc Psychol Bull* 17(5):475–482. doi:10.1177/0146167291175001
- Brewer MB (1995) *Managing diversity: the role of social identities*. American Psychological Association, Washington, DC
- Brewer MB (1996) When contact is not enough: social identity and intergroup cooperation. *Int J Intercult Relat* 20 (3–4):291–303. doi:10.1016/0147-1767(96)00020-x
- Brown R (2000) Social identity theory: past achievements, current problems and future challenges. *Eur J Soc Psychol* 30(6):745–778. doi:10.1002/1099-0992(200011/12)30:6<745::aid-ejsp24>3.0.co;2-o
- Casey-Campbell M, Martens ML (2009) Sticking it all together: a critical assessment of the group

- cohesion–performance literature. *Int J Manag Rev* 11(2):223–246. doi:10.1111/j.1468-2370.2008.00239.x
- Chiochio FO, Essiembre H (2009) Cohesion and performance. *Small Group Res* 40(4):382–420. doi:10.1177/1046496409335103
- De Cremer D, Leonardelli GJ (2003) Cooperation in social dilemmas and the need to belong: the moderating effect of group size. *Group Dyn Theory Res Pract* 7(2):168–174. doi:10.1037/1089-2699.7.2.168
- De Cremer D, Van Knippenberg D (2002) How do leaders promote cooperation? The effects of charisma and procedural fairness. *J Appl Psychol* 87(5):858–866. doi:10.1037/0021-9010.87.5.858
- Doosje B, Spears R, Ellemers N (2002) Social identity as both cause and effect: the development of group identification in response to anticipated and actual changes in the intergroup status hierarchy. *Br J Soc Psychol* 41(1):57–76. doi:10.1348/014466602165054
- Ellemers N, Spears R, Doosje B (2002) Self and social identity. *Annu Rev Psychol* 53(1):161–186. doi:10.1146/annurev.psych.53.100901.135228
- Fitzgerald GE (1989) Report of a commission of inquiry pursuant to orders in council. Queensland Government, Brisbane
- Goldstein H (2003) On further developing problem-oriented policing: the most critical need, the major impediments, and a proposal. *Crime Prev Stud* 15:13–47
- Hardy J, Eys MA, Carron AV (2005) Exploring the potential disadvantages of high cohesion in sports teams. *Small Group Res* 36(2):166–187. doi:10.1177/1046496404266715
- Haslam SA (2004) Psychology in organizations: the social identity perspective. Sage, London
- Hogg MA (1993) Group cohesiveness: a critical review and some new directions. *Eur Rev Soc Psychol* 4(1):85–111. doi:10.1080/14792779343000031
- Hogg MA (2001) A social identity theory of leadership. *Pers Soc Psychol Rev* 5(3):184–200. doi:10.1207/s15327957pspr0503_1
- Hogg MA, Terry DJ (2000) Social identity and self-categorization processes in organizational contexts. *Acad Manage Rev* 25(1):121–140
- Hogg MA, Fielding KS, Johnson D, Masser B, Russell E, Svensson A (2006) Demographic category membership and leadership in small groups: a social identity analysis. *Leadersh Q* 17(4):335–350. doi:10.1016/j.leaqua.2006.04.007
- Lyons W (2002) Partnerships, information and public safety: community policing in a time of terror. *Polic Int J Police Strateg Manag* 25(3):530–542
- Mazerolle L, Ransley J (2005) *Third Party Policing*. Cambridge University Press, New York
- Mazerolle L, Ransley J (2006) *Third party policing*. Cambridge University Press, Cambridge, UK
- Rink F, Ellemers N (2007) Diversity as a basis for shared organizational identity: the norm congruity principle. *Br J Manag* 18:S17–S27. doi:10.1111/j.1467-8551.2007.00523.x
- Tajfel H, Turner JC (1986) The social identity theory of intergroup behavior. In: Worchel S, Austin WG (eds) *Psychology of intergroup relations*, 2nd edn. Nelson-Hall, Chicago, pp 7–24
- Terry DJ (2000) Inter-group relations and organizational mergers. In: Hogg MA, Terry DJ (eds) *Social identity process in organizational contexts*. Taylor & Francis, New York
- Terry DJ, O'Brien AT (2001) Status, legitimacy, and ingroup bias in the context of an organizational merger. *Group Process Intergroup Relat* 4(3):271–289. doi:10.1177/1368430201004003007
- Terry DJ, Carey CJ, Callan VJ (2001) Employee adjustment to an organizational merger: an intergroup perspective. *Pers Soc Psychol Bull* 27(3):267–280. doi:10.1177/0146167201273001
- Tsui AS, Egan TD, Iii CAOR (1992) Being different: relational demography and organizational attachment. *Adm Sci Q* 37(4):549–579
- Turner JC, Hogg MA, Oakes P, Reicher SD, Wetherell MS (1987) *Rediscovering the social group: a self-categorization theory*. Blackwell, Oxford, UK/Basil
- Tyler TR (1989) The psychology of procedural justice: a test of the group-value model. *J Pers Soc Psychol* 57(5):830–838. doi:10.1037/0022-3514.57.5.830
- Tyler TR, Blader SL (2001) Identity and cooperative behavior in groups. *Group Process Intergroup Relat* 4(3):207–226. doi:10.1177/1368430201004003003
- Tyler TR, Blader SL (2003) The group engagement model: procedural justice, social identity, and cooperative behavior. *Pers Soc Psychol Rev* 7(4):349–361. doi:10.1207/s15327957pspr0704_07
- Tyler T (2011) *Why people cooperate: The role of social motivations*. Princeton University Press, Princeton

Social Harms

- ▶ [International Responses to Victims in Criminal Justice](#)

Social Learning Theory

Kyle J. Thomas and Sally S. Simpson
Department of Criminology and Criminal
Justice, University of Maryland, College Park,
MD, USA

Synonyms

[Behavioral learning](#); [Differential association-reinforcement theory](#); [Learning theory](#)

Overview

In their differential association-reinforcement theory, Burgess and Akers (1966) reformulated, and restated, Edwin Sutherland's principles of differential association (1947) to include the learning principles of modern behavioral learning. Sutherland's theory had been roundly criticized on two main grounds: (1) its failure to make explicit the precise mechanisms of learning and (2) difficulties operationalizing and empirically testing the main ideas of the theory leading some to assert that the theory was unfalsifiable (Nettler 1984). Akers (1973, 1998) subsequently revised differential association-reinforcement theory into a social learning theory of crime and deviance that included four key constructs: differential association, definitions, differential reinforcements, and imitation. In the most general terms, Akers' social learning theory asserts that both conforming and deviant behavior are learned in the same way, that one's likelihood of engaging in antisocial behavior is influenced by his/her prior and anticipated consequences of behavior, or that the probability of engaging in deviant behavior is contingent on the prior and anticipated future reinforcements and punishments one has experience regarding deviance.

Social learning theory remains a dominant perspective in criminology, and the empirical literature testing this theory is arguably one of the largest and most persuasive in the field (see Warr 2002). This entry provides a general overview of the assumptions, constructs, and propositions that are central to social learning theory, summarizes the empirical literature testing this perspective, addresses some of the key criticisms, and concludes with new directions for the theory.

Background

In 1965, C. Ray Jeffrey recommended that Sutherland's differential association theory be replaced with a single statement drawing on the principles of modern behavioral learning. Specifically, Jeffrey (1965, p. 295) noted that criminal

behavior, like all other human behavior, is governed by the principles of differential reinforcement such that it "occurs in an environment in which in the past the actor has been reinforced behaving in this manner, and the aversive consequences attached to the behavior have been of such a nature that they do not control or prevent the response." The importance of Jeffrey's recommendation cannot be overstated, as he correctly pointed out that criminologists from the learning paradigm had neglected to incorporate the significant advancements in behavioral psychology into the understanding of crime since Sutherland's original contributions.

Around the same time, Burgess and Akers were in the process of integrating these modern learning principles into the differential association perspective. They believed that all of the explanatory variables and processes in differential association theory could be subsumed under a more general theory of human behavior that is based on behavioral learning principles. Burgess and Akers accomplished this by revising each of Sutherland's nine statements to form seven new propositions that invoke language from behavioral psychology and modern learning. For instance, the first and eighth propositions in Sutherland's differential association theory which read "Criminal behavior is learned" (p. 6) and "The process of learning criminal behavior by association with criminal and anti-criminal patterns involves all of the mechanisms that are involved in any other learning" (p. 7), respectively, were combined into a single proposition that read "Criminal behavior is learned according to the principles of operant conditioning" (Burgess and Akers 1966, p. 132). Further, the key proposition of Sutherland's theory which read "A person becomes delinquent because of an excess of definitions favorable to violation of law over definitions unfavorable to violations of law" was similarly revised in a manner consistent with behavioral learning, such that "Criminal behavior is a function of norms which are discriminative for criminal behavior, the learning of which takes place when such behavior is more highly reinforced than non-criminal behavior" (p. 143).

The initial formulation of social learning theory in criminology was heavily influenced by the behavioral learning paradigm developing in psychology (Bandura 1969). In this seminal work, Burgess and Akers discussed the importance of things like discriminative stimuli (signals that elicit a particular behavioral response), saturation (when the potency of a reinforcer is maximized and no longer affects behavior), and extinction (the reduction or loss of a behavioral response when reinforcement is withheld) when describing the learning process. The complexity inherent in this perspective, however, led many criminologists to question the utility and testability of this learning perspective. These concerns, along with the movement towards a variable-based criminology (Hirschi 1969), led Akers to revise social learning theory into a more tenable set of constructs and propositions.

Akers' Revised Social Learning Theory

Social learning theory is developed to explain not just the acquisition of criminal behavior but also the maintenance and modification of criminal behavior (Akers 1998). Put simply, the process of social learning offers not just explanations on the motivation to initially engage in criminal behavior but also explanations on how certain variables work to maintain criminal behavior throughout time and how these variables also can maintain and promote conformity over time. Akers (1998, p. 50) argues that “The basic assumption in social learning theory is that the same learning process. . . produces both conformity and deviant behavior. The differences lie in the direction. . . [of] the balance of influences on behavior.” Thus, individuals are at an increased likelihood of engaging in deviant behavior when they differentially associate with others who commit criminal behavior themselves, who define deviance as desirable, and who provide (or have provided) greater reinforcement than punishment for that deviance. Conversely, individuals are more likely to be prosocial to the extent that they differentially associate with

individuals who engage in and espouse norms conducive to conformity.

When developing his revised social learning theory, Akers asserts that the social learning process can be principally captured by four distinct, yet interdependent, constructs: differential association, definitions, differential reinforcement, and imitation. The subsequent sections of this entry will elaborate on these four constructs and discuss how they influence the probability of engaging in criminal and deviant behavior.

Differential Association

Social learning theory begins with the assumption that all learning takes place in a context of structural interactions with others. The most important social context in which mechanisms of learning operate is through the differential associations individuals have. This construct draws heavily on Sutherland's theory, which emphasizes the importance of “intimate personal groups” in the learning process.

Akers asserts that differential association can influence delinquency both through behavioral/interactional dimensions and normative dimensions. The former refers simply to the direct association and interactions individuals have with others and whether the behavioral patterns that these associations expose to an individual are criminal or conforming. In other words, it refers to the behavioral nature of one's associates (e.g., deviant vs. non-deviant). The latter, however, are the different patterns of norms and values that an individual is exposed to through these associations. Sutherland's discussion of differential association was primarily concerned about the transmission of definitions that occurred in these differential interactions, but Akers essentially allows the differential association construct to incorporate a much broader range of deviant influences present when interacting with others, including the transmission of deviant definitions, providing reactions that are favorable to delinquent conduct that alters the perceived rewards/costs of crime and exposing individuals to deviant models from which to imitate.

Like Sutherland, Akers' social learning theory does not view all interactions as being equally

influential in affecting human behavior. Instead, Akers retains the four modalities discussed by Sutherland, asserting that associations, and their subsequent influence on deviant behavior, vary in terms of frequency, duration, priority, and intensity. Frequency is how often an individual interacts with a group or individual. Duration concerns the length and relative amount of time spent in different associations. Priority refers to when in time the association began, in other words, which associations came earliest in time. Finally, intensity has to do with the “significance, saliency, or importance of the association to the individual” (Akers 1998, p. 64). Drawing on these four modalities, Akers (like Sutherland) argues that associations will have a greater impact on shaping an individual’s behavior if they occur more frequently, if they have been interacting for longer periods of time, if they began earlier in the individual’s life, and if they are more meaningful relationships to the individual.

This construct of differential association subsumes many of the major elements of Sutherland’s differential association theory. Sutherland’s recognition that the interactions that individuals have when associating with others is central to the learning process is acknowledged by Akers, and he continues to stress the importance that different modalities have on influencing how these associations shape behavior. In fact, given Akers’ assertion that most learning is shaped by social reinforcement contingencies – rather than nonsocial ones (e.g., intrinsic rewards/punishments) – differential association is a necessary condition in the social learning perspective.

Definitions

Sutherland’s emphasis on the transmission and acquisition of definitions favorable to delinquency when explaining deviant behavior made it difficult to distinguish that construct from differential associations. However, Akers explicitly separates these two constructs in his social learning theory. Despite this, Akers’ characterization of definitions is nearly identical to that described by Sutherland. Definitions can generally be thought of as the normative beliefs,

attitudes, or meanings that individuals attach to any given behavior. These can include the motives, drives, rationalizations, justifications, and attitudes that define the circumstances in which the commission of an act is right or wrong, desirable or undesirable, justified or unjustified (Akers 1998, p. 77). In social learning theory, definitions can be both general and specific. General definitions refer to the broad religious and moral beliefs and norms one holds that favor either criminal or conforming behavior. Specific definitions refer to specific acts and situations in which behavior is, at minimum, justified. The distinction between general and specific definitions is of theoretical importance. Individuals may generally hold the belief that “breaking the law is wrong” but at the same time believe that “smoking marijuana is OK.” Similarly, individuals may hold the general definition that “thou shall not steal” while simultaneously believing “it is morally justifiable to steal when hungry.”

Definitions in Akers’ theory do not *require* deviant behavior, as some scholars have wrongly interpreted in Sutherland’s theory (Kornhauser 1978). Akers (1998) argues that *some* definitions favorable to delinquency can be so strongly held that they require delinquency. More common, however, is that individuals hold definitions that rationalize or justify delinquency in certain situations. In this way, definitions in Akers’ theory are similar to the belief construct in Hirschi’s (1969) social control theory, in that the definitions themselves do not directly provide motivation for delinquency but rather are conventional beliefs that are weakly held and can allow for other motivating factors (e.g., differential reinforcement) to have important and direct effects on antisocial behavior. That is to say, their moral reservations against delinquency are weak enough that it can be bypassed in some situations and thus allow for delinquent behavior when the right circumstances or reinforcement contingencies are present (Warr 2002).

Differential Reinforcement

Arguably the construct of greatest importance in Akers’ social learning theory is differential

reinforcement. Though Sutherland stated that criminal behavior is learned like all forms of human behavior, he spent little time discussing the learning process. Drawing on research in psychology, Akers argues that the process of learning antisocial behavior is best captured through the construct of differential reinforcement. “Differential reinforcement refers to the balance of anticipated or actual rewards and punishments that follow or are consequences of behavior” (Akers 1998, p. 67). If an individual believes that they are likely to be rewarded for engaging in deviance and/or punished for deciding upon conforming behavior, then that individual is likely to engage in deviance. Moreover, when faced with a similar situation in the future, that individual is likely to draw on their prior experiences and view that deviant behavioral choice as more reinforcing and the conforming option more punishing. This ultimately leads to what is arguably the most important proposition in Akers’ theory: “Whether individuals will refrain from or initiate, continue committing, or desist from criminal and deviant acts depends on the relative frequency, amount, and probability of past, present, and anticipated rewards and punishments perceived to be attached to the behavior” (Akers 1998, p. 66).

From the social learning perspective, human behavior is governed by both reinforcements and punishments, whereby the former *increases* the probability that a behavior occurs and the latter *decreases* it. Both reinforcements and punishments can be either positive or negative, which references the fact that some consequences influence behavior by adding stimuli (positive) while others influence it by removing stimuli (negative). Accordingly, positive reinforcements increase the likelihood of engaging in or repeating deviance by providing a positive outcome or reaction to it (e.g., friends’ approval, money, or intrinsic feelings of pleasure). Negative reinforcements increase the likelihood of deviance by removing an aversive or unpleasant feeling or event (e.g., eliminating the ridicule of classmates). Meanwhile, positive punishments decrease the likelihood of criminal behavior by providing a painful or unpleasant consequence

associated with a behavior (e.g., an arrest), and negative punishments affect behavior by removing pleasant consequences or rewards (e.g., parental respect). These reinforcements and punishments influence the likelihood of engaging in deviance by altering an individual’s perceived anticipated consequences when faced with similar situations in the future, which in turn affect one’s decision-making calculus at the time of the offense (Akers 1990, 1998).

Just as with differential associations, reinforcement and punishment contingency are not all weighted equally but are influenced by the amount, frequency, and probability of their occurrence. The greater the amount or value placed on the reinforcement contingencies, the more frequently it is reinforced, and the greater the likelihood of experiencing the reinforcement all increase the likelihood of engaging in deviant behavior. Thus, for instance, if an individual places great value on the relationship she shares with her friend who frequently rewards deviant behavior, and she is certain that her friend will reward her for stealing, that individual will have a high probability of engaging in deviant behavior. However, the reinforcements and punishments that are provided by strangers for whom the individual has no personal relationship with will likely have little effect both on the present and on future perceptions of rewards and punishments.

Though the examples provided above have focused largely on social reinforcements and punishments (i.e., peers and parents), Wood et al. (1997) note that reinforcement contingencies can be nonsocial. For instance, the physiological effects of drugs and alcohol can be rewarding to some, which in turn increases the likelihood of repeated use. Moreover, some individuals may view deviant behavior itself as intrinsically rewarding. This would occur if the behavior itself, regardless of social reactions, is viewed as “fun” or “thrilling.” Wood et al. provide support for this premise: positive intrinsic sensations from crime are significantly related to the maintenance of criminal behavior. Interestingly, Akers vests little in the idea that nonsocial factors are inherently reinforcing. For instance,

he argues that individuals learn that the effects of drugs are enjoyable or unpleasant *through social reinforcement*. Similarly, Akers also argues that some individuals are prone to be risk-seeking but that they learn to become risk-seeking, and learn to see delinquency as “fun,” through social interactions with others. In other words, individuals may experience physiological and intrinsic rewards from drug use and crime but primarily only do so if they are predisposed by a *social learning history* to do so.

Akers believes that the primary source of learning occurs through social reinforcements and punishments. When individuals value the relationships they share with others, the prior, anticipated, and actual reactions that they attach to deviant behavior are likely to be highly influential in affecting the likelihood that an individual offends. Importantly, this is not the same as having peers pressuring individuals to offend, but rather this process refers to the influence that these friends have by altering the perceived consequences of different behavioral choices. Social learning theory recognizes that humans are social beings who are concerned with the reactions and opinions of others and that our perceptions on what the anticipated consequences of behavior entail are necessarily a function of prior experiences in similar situations. Even tangible and material items, such as money and wealth, are valued by individuals because of the social rewards and prestige that they provide. Put simply, Akers argues that tangible and material rewards from crime derive their power of influence because they are symbolic of social rewards (i.e., prestige and status) and that their value (and thus their influence on behavior) is of little importance when the social rewards are disassociated from them.

In sum, the inclusion of differential reinforcement into the learning paradigm of criminology filled the important gap as to *how* criminal behavior is learned (i.e., a focus on the “reception” and not the “transmission” process). Though this contribution by Akers is arguably the most important aspect in social learning theory, it also remains one of the most misunderstood. The idea that individuals are affected by reinforcement

contingencies does not mean that humans are passively conditioned by their environment, simply reacting to stimuli like Pavlovian dogs. On the contrary, social learning theory views that individuals are mostly actively engaged and attuned to the anticipated consequences of behavioral actions (Akers 1990), which presupposes that individuals employ some cognition before engaging in delinquency. If an individual determines that the relative balance of the consequences of deviance provides greater reinforcements than punishments, then that individual is at an increased likelihood of engaging in deviant behavior. Moreover, social learning theory argues that the actual consequence attached to the behavior will affect their future likelihood of offending: if an individual engages in antisocial behavior and they are rewarded for it, they are at an increased likelihood of repeating the action when faced with a similar situation in the future. Conversely, if individuals engage in deviance but experience a relatively greater level of negative consequences, they are likely to avoid deviant behavior in the future. Simply put, social learning theory argues that an individual’s prior experiences influence how they interpret anticipated consequences in the future and that these anticipated consequences are influential when deciding on a behavioral action choice (Akers 1998).

Imitation

Imitation refers to the modeling of behavior after similar behavior is observed in others. The classic example of imitation derives from Bandura’s (1961) Bobo doll experiments, where the behavior of children was observed after watching a model aggress on an inflatable doll. As is evident in these experiments, not all behavior that is observed is imitated by others. Of particular importance is the observed consequence of the behavior: if one observes an individual engage in a behavior and that individual is subsequently rewarded for the behavior, it is more likely that that behavior will be modeled. This is known as *vicarious reinforcement*. Conversely, if one observes behavior that is subsequently punished, it is less likely that the behavior will

be modeled – what is known as *vicarious punishment*.

Another factor that influences whether an observed behavior is modeled is the characteristic of the model. In this way, imitation shares some of the modalities as the other constructs in social learning theory. The most salient models are those persons “with whom one is in direct contact in primary groups” (Akers 1998, pp. 76–77). This most obviously means that individuals are more likely to imitate the behavior of close, personal associates than that of strangers. But this statement also implies that individuals are more likely to model behavior observed in direct interaction than behavior observed through media forums. Akers (1998) has argued that technological and social changes have increased the relative importance of modeling behavior on television, movies, and video games. More specifically, given that the relative time spent watching television and movies and playing video games has increased over the last several decades, the modalities of learning have also increased (frequency, duration, priority, and intensity), making media models more salient.

Intuitively, the imitation of others’ behavior is most influential in the acquisition and performance of novel behavior but plays little role in the maintenance of deviant behavior. The continuation of the behavior is more affected by direct reinforcements and punishments. For example, an individual may learn initially that violence is rewarded when one’s status is challenged, but whether they continue to behave in that matter will be contingent on their own experiences with such behavior. However, imitation may still play a role in the acquisition of criminal skills that can make crime easier and less risky.

Though these constructs are treated and discussed independently, there is considerable interdependency among them. Indeed, social learning is itself a process that incorporates all of these constructs in a complex and recursive manner. Initial deviance is influenced by the balance of definitions one has been socialized to hold, by the modeling of others’ behavior, and/or by the anticipated rewards and punishments that derive from a criminal event. Immediately after

this event, the actual consequences of the action play an important role – if the individual experiences some reward, then he/she is likely to believe there will be positive consequences in the future. Further, if one does experience reinforcement, this will likely alter the definitions they hold and, in turn, could potentially alter the individual’s criminal and non-criminal associates (i.e., self-selection). Indeed, social learning theory does not claim that associations are random but predicts that individuals do self-select into peer networks. However, social learning theory does predict that more often than not delinquent associations precede delinquency.

Research on Social Learning Theory

Few theories have been tested empirically to the extent that social learning theory has. These tests most often assess the relationship and direction between variables from the social learning perspective – differential association, definitions, differential reinforcement, and imitation – and some deviance outcome measure (i.e., general delinquency, drug use, violence). The overwhelming majority of these studies have found strong support for social learning predictions, with very few studies finding negative support for the theory.

One of the earliest and most seminal tests of the theory was conducted by Akers and his colleagues in 1979. Using substance use as an outcome, the results found that differential association (i.e., delinquent friends), definitions (how right/wrong substance use is), and differential reinforcement (how parents and friends would react to respondent drug use) were all strongly and statistically related to substance use, providing support for the social learning predictions. Imitation, however, was only weakly and inconsistently related to substance use, though Akers et al. (1979) note that this finding is not unexpected, as imitation is only predicted to have strong effects on the initiation of delinquent behavior. Though the research findings were strongly supportive of the theory, the clear operationalization of the theory into testable

measures and predictions within the article is an equally important contribution. Since this work, many empirical studies have assessed the empirical validity of social learning, with most supportive of the perspective (Pratt et al. 2010).

Another common manner in which research has inferred the empirical validity of social learning theory is through an assessment of the relationship between delinquent peers and offending. As would be predicted by social learning theory, the relationship between peers and delinquency is one of the strongest and most consistent findings in all criminological research (Warr 2002). More frequent, longer-term, and closer associations with peers who engage in delinquent behavior is consistently related to the delinquency or respondents, while associating with prosocial peers is consistently related to conformity. This strong relationship has generally held using various samples and methodological specifications. Ultimately, these findings have led some scholars to conclude that delinquent peers are one of, if not, the most important causes of adolescent delinquency (Warr 2002) and provide considerable support for both the differential association and social learning perspectives.

Perhaps the most telling support for social learning theory comes in two forms: its relative superiority when faced with theoretical competition and its ability to explain known correlates of offending. Regarding the former, many empirical works in the field of criminology have sought to pit variables from one theoretical perspective against those from others. Most generally, this involves putting learning variables in statistical models and comparing the size of the effects against variables from control and strain theories. When this is done, the variables from learning theory tend to explain more of the variance in delinquency than variables from other perspectives (Warr 2002). Learning variables also do an impressive job explaining some of the well-known correlates of offending. For instance, Warr (1993, 1998) has demonstrated that associating with delinquent peers can explain both the age and marriage effect. Similarly, Mears et al. (1998) have shown that learning theory also

explains a substantial portion of both race and gender differences in crime, respectively.

The empirical support for social learning theory has most recently been summarized in a meta-analysis conducted by Pratt and colleagues (2010). This study reviewed the published extant tests of the perspective to provide a synthesized estimate of the effect size of the social learning theory constructs. In reviewing these studies, Pratt et al. conclude that the empirical support for social learning theory is stronger than many other criminological perspectives and is at least as comparable as the most strongly supported theories (see Pratt and Cullen 2000). Interestingly, however, the different constructs of social learning theory appear to have varying levels of explanatory power. The constructs most strongly related to deviance are differential associations and definitions, the two constructs already captured in Sutherland's theory. Differential reinforcements and imitation are only modestly related to delinquency. If indeed differential reinforcement is the most important construct in the social learning perspective, these findings raise some questions about Akers' perspective and provide researchers with an important avenue to further assess the interdependence of the four constructs of social learning theory.

An extension of social learning occurred in 1998 when Akers (1998) suggested that an integrated theory (social structure and social learning) could explain structural variations in crime rates across communities. Although this premise has yet to receive much empirical attention, the basic argument is that social structure is linked to individual behavior – and in turn, crime rates – through its effects on social learning variables, namely, differential association, differential reinforcement, definitions, and imitation. Structural characteristics provide the contexts within which social learning variables operate: it causes individuals to differentially associate with delinquents and exposes individuals to a greater ratio of deviant reinforcement contingencies and to definitions favorable to delinquency, as well as a greater number of delinquent models to imitate.

Four dimensions make up the SSSL perspective. *Differential social organization*

refers to the distribution of structural correlates of crime (age, race, gender, SES) in the community or society that affect rates of delinquency. *Differential location in the social structure* refers to one's own social and demographic location in the social structure that affects social learning variables (e.g., younger individuals are more likely to be exposed to other younger individuals who, in turn, are more crime prone). *Theoretically defined criminogenic aspects of social structure* draws on anomie, social disorganization, and conflict theories which, again, Akers argues affects one's exposure to criminal associates, definitions, models, and reinforcements that can promote deviant behavior. Finally, *differential social location in primary, secondary, and reference groups* refers to the more immediate personal networks that filter the larger social environment to either promote or discourage criminal behavior. In sum, the general premise of SSSL is that social learning variables fully mediate social structural characteristics correlated with high rates of crime. To date, however, this perspective has not fully been tested by researchers, and the ability of social learning theory to explain differences in crime rates remains in question.

Criticisms

Pratt and colleagues' (2010) finding that differential reinforcement and imitation had relatively smaller effects in predicting delinquency when compared to differential association and definitions can be seen as challenging to Akers' contributions to the learning paradigm, but several other criticisms also plague the theory. Most notable, several scholars have challenged the supposed causal relationship between peers and delinquency that is hypothesized by learning theorists. The argument is that the strong correlation between delinquent peers and delinquency is not evidence that differential associations cause delinquency but instead reflects the tendency for "birds of a feather to flock together" (Gottfredson and

Hirschi 1990). This "selection" argument has more recently been championed most fervently by scholars from the control perspective who assert that both delinquency and the association with delinquent peers are both outcomes of similar causal factor (i.e., weak attachments to society or low self-control). The criticism is that delinquent associations are not random, but instead individuals self-select into friendship networks based on delinquent tendencies. Thus, proponents of the selection perspective argue that the relationship between peers and delinquency is spurious, as delinquent behavior is thought to actually precede delinquent peer associations.

Though often seen as one of the greatest challenges to social learning theory, this criticism falls short for both theoretical and empirical reasons. First, social learning theory does not predict that peer associations are randomly constructed but instead explicitly acknowledges that individuals self-select into peer groups based on their prior learning. Unlike control theories, however, social learning theory predicts that once this selection does occur, these associates are likely to have a continued influence on delinquent behavior through differential reinforcement and other learning mechanisms. The empirical research is supportive of the learning predictions: even when statistically controlling for competing variables that are thought to render the peer-delinquency relationship spurious (Gottfredson and Hirschi 1990), the association between peers and delinquency remains (Matsueda and Anderson 1998). Associations with delinquent peers most often precede delinquent behavior (Elliott and Menard 1996), which is supportive of the predictions of social learning theory. Still, this selection versus socialization debate is far from settled (see Weerman 2011). Future research using experimental and quasi-experimental designs will likely shed more light on this issue.

Another criticism of social learning theory concerns the types of deviant behaviors in which the theory is able to explain. Many tests of social learning theory have assessed whether Akers' theory can explain variation in substance

use and other forms of minor delinquency. Much less research has examined how well social learning theory explains more serious forms of delinquent behavior, and the research that does finds that learning theory does a poorer job explaining these more serious forms of delinquency. The reason behind this remains relatively unexplored, but one possibility is that individuals who engage in more serious forms of violence have such high levels of criminal propensity (i.e., high in impulsivity) that they do not require learning mechanisms to engage in delinquent behavior. In any case, the finding that social learning has more trouble explaining more serious forms of delinquency is problematic for the theory, as Akers' perspective is designed to explain all forms of crime and deviance.

Finally, following Kornhauser's critique (1978, p. 34) of differential association/social learning theory assumptions that "man has no nature, socialization is perfectly successful, and cultural variability is unlimited," social learning theory has been interpreted by some as a theoretical perspective unable to explain individual differences and only applicable to group differences in crime where subcultural definitions adhere to deviance (Akers 1996, p. 230). A corollary is that the theory focuses entirely on the development of novel criminal behavior that, once learned, *requires* deviance. In essence, critics see little to differentiate differential association and social learning from cultural deviance theories. Akers (1996), in an exchange with Travis Hirschi on this subject (Hirschi 1996), suggests that these criticisms stem from ambiguities associate with and misrepresentations of Sutherland's statement of the differential association – problems that Akers (1996, p. 238) believes have been rectified in his specification of social learning principles by allowing for variations in both deviant definitions held by individuals and the social reinforcement they receive. "These deviations can develop within the same normative or cultural system. They do not require the existence or participation in an organized deviant subculture."

New Directions

Despite these criticisms, social learning theory remains one of the preeminent theories of crime, and the considerable empirical support for this perspective seems to justify this prominence. Nevertheless, there are some important questions deriving from social learning theory that have been left unexplored.

Saturation Effects and Discriminative Stimuli. In describing a social learning theory of crime and deviance, Akers (1998; see also Burgess and Akers 1966) invokes language from psychology that describes how learning differs across time and situations. Most notably are the concepts of saturation and discriminative stimuli. Saturation occurs when the potency of a reinforcer has been maximized, and that reinforcer no longer has an effect on behavior. This has clear implications for the structure of peer networks: some individuals may no longer be influenced by associating with delinquent associates, as their prior learning and already existing associates may have already hit a saturation point. Put simply, the relationship between delinquent associates and delinquency may be nonlinear, increasing at a decreasing rate. Zimmerman and Vasquez (2011) have assessed this saturation hypothesis at a neighborhood level. Their results provide support for the saturation effect: delinquent peers have a strong effect for individuals residing in non-disadvantaged neighborhoods but a diminishing effect in disadvantaged ones. Still, there is considerable research left unexplored on the saturation effect at the more proximal friendship network level and over the life course. How does the inclusion of a delinquent peer change the delinquency of an individual in a relatively nondelinquent network versus a relatively highly delinquent network? Further, do peers have the same effect over the life course? Or do individuals become saturated from delinquent peers leading to a declining peer effect as one ages?

The discriminative stimuli effect has been relatively neglected in criminological research (however, see McGloin et al. 2011), despite the direct implications it has for theory. Discriminant

stimuli refer to stimuli that influence how an individual will act in a given situation. Discriminant stimuli are most important when they signal which situations and circumstances will yield reinforcement. To be clear, it is highly unlikely that individuals view that their friends will reward violence, theft, or other forms of deviance in all situations. Rather, it is more common that an individual's delinquent friends will reward it in some situations and under some circumstances. Discriminative stimuli serve as the cues as to when these situations and circumstances occur. The same logic applies in understanding one's definitions favorable/unfavorable to delinquency. As noted above, most delinquents do not hold the universal values that "deviance is right" but rather hold definitions that justify delinquency in some situations (Akers 1998). Unfortunately, most measures tapping into peer reactions and individual definitions use global statements on the rightness and wrongness of antisocial behavior. Considerable theoretical insight can be gained by beginning to incorporate this discriminative stimuli logic into learning measures and assessing what situations and circumstances allow individuals to justify deviant behavior. The overall point is that there remains important elements from the more general behavioral learning paradigm that have not been explored in criminological research that can shed some important light on the understanding of crime.

Differential Susceptibility: How and When Learning Variables Matters. Just as learning variables may not have ubiquitous effects across all situations, there may also be variability among individuals in this regard. For generations the primary question surrounding social learning theory has been the central question of causality: do peers and other associates (e.g., family) hold causal role in the facilitation of delinquency? As noted above, the overwhelming evidence supports the predictions from social learning theory. At this point, it seems worthwhile to move beyond simple assessments of *if* learning variables influence delinquency and begin to assess if they matter differently for different people.

Take, for instance, the relationship between delinquent peers and delinquency. Akers suggests that delinquent peers influence delinquent behavior by providing positive reactions to such behavior. He goes on to specifically acknowledge that the influence of such reactions is necessarily dependent on the ability of individuals to consider the longer-term consequences of behavior; in other words, it requires individuals to reflectively consider how their peers would react before deciding to offend (Akers 1990). If this is the case, then some individuals may not be susceptible to the risk of associating with delinquent peers. In particular, individuals high in impulsivity, who are characterized by their tendency to make decisions based almost entirely on situational stimuli, may not need or consider the reactions of their peers when deciding to offend. Instead, these individuals will be most susceptible to the situational risk of unstructured socializing with peers (Osgood et al. 1996). This hypothesis was supported by Thomas and McGloin (2013) using three different large-scale data sets: delinquent peers have their strongest effects on individuals low in impulsivity but only weak (and nonsignificant) effects on individuals high in impulsivity. Further, the situational risk of unstructured socializing with peers has its strongest effects on this high in impulsivity, but a weak (and nonsignificant) effect on this low in impulsivity. Thus, some individuals may be particularly susceptible to the reactions of their peers, while others may not need such reactions to motivate them to deviate. The larger point is that researchers should begin to think more critically about the mechanisms by which learning variables facilitate delinquency. These more nuanced considerations can shed light on important considerations as to how and when peers and other associates facilitate delinquent behavior. Do all individuals rely on delinquent associates to deviate? How and when do individuals consider their own definitions before deciding to offend? Indeed, the question as to if social learning variables are related to crime and deviance appears to have been answered, and researchers should begin to assess whether these variables matter differently for different people.

Objective versus Perceptual Measures. An area of growing interest among criminologists deals with the most appropriate way to measure variables of theoretical interest, and social learning theory has been particularly in the crosshairs of this discussion. Several scholars have expressed concerns regarding the adequacy of traditional perceptual measures of delinquent peers, whereby respondents are asked to report on the behavior of their friends. Most critically, critics have suggested that perceptual measures have had a tendency to overestimate the delinquent peer effect, because individuals are not accurately reporting the behavior of their friends, but instead are merely projecting their own behavior to them (Gottfredson and Hirschi 1990). The limited research that has been conducted is supportive of this claim: when objective measures of peer deviance are used, the relationship between peers and delinquency is reduced, though still substantively large. These findings have led many scholars of crime to encourage the use of objective measures over the use of perceptual ones while outright rejecting the utility of the latter measures because they reflect an individual's misperceptions of their friends' behavior and not the reality of it.

It is interesting that this movement has occurred during the same time period in which rational choice and deterrence theorists have advocated the use of perceptual measures over objective ones when assessing criminal decision making. After decades of empirical work challenged the core propositions of rational choice theory when using objective measures, scholars began to realize that sanction threats can only influence behavior if they are perceived by the individual. The crux of their argument was that these choice theories are perceptual in nature and, thus, require scholars to assess one's perceptions of their beliefs of risk rather than utilizing the objective reality of that risk. The positive research using these perceptual measures has led to some resurgence in the rational choice perspective (Nagin 1998). These positive findings has led Nagin (1998) to argue that scholars interested in the choice perspective should move beyond assessing if risk perceptions

influence behavior and focus on how individuals actually formulate their perceptions of risk, in other words, what information (or "signals") go into the development of subjective risk perceptions.

Interestingly, social learning theory is also an inherently perceptual theory of crime. The core proposition of the theory is that the perceived reinforcements and punishments associated with behavior influence the present and future likelihood of offending. Given this strong emphasis on perceived reinforcements and punishments, and the robust relationship between these perceptions and offending, one future direction for social learning scholars is to heed Nagin's (1998) call and end the discussion on the superiority of objective measures over perceptual ones and begin to consider what factors go into one's perceptions of social reinforcements and punishments. To be clear, a recent study conducted by Young et al. (2011) found that individuals have a tendency to misperceive their own friends' delinquent behavior. Of course, this finding could be used as more ammunition for the objective-measure-superiority debate, but it also can be used to ask more interesting questions regarding individual perceptions of reinforcements: what causes individuals to misperceive their friends' behavior? Or more generally, how do individuals formulate their subjective perceptions of social reinforcements? The answer to these questions has considerably more theoretical utility and can shed important light on the learning process, in general.

Conclusion

Social learning theory begins with the assumption that criminal behavior, like all human behavior, is learned in the interaction with others. This theory went beyond Sutherland's perspective by more clearly specifying what this learning process entailed. Simply, social learning theory asserts that the probability an individual commits deviant acts depends on the relative frequency, amount, and probability of past, present, and anticipated rewards and punishments perceived

to be attached to the behavior. The empirical literature supportive of the theory is one of the most impressive in the field of criminology. Nevertheless, considerable avenues of ongoing and future research can lead to more theoretical clarity.

Related Entries

- ▶ [Behavioral Management in Probation](#)
- ▶ [Cognitive/Information Processing Theories of Aggression and Crime](#)
- ▶ [Co-offending and Offender Decision-Making](#)
- ▶ [Cultural Criminology](#)
- ▶ [Deterrence](#)
- ▶ [Differential Association Theory](#)
- ▶ [Informal Social Control](#)
- ▶ [Punishment as Rehabilitation](#)
- ▶ [Rational Choice Theory](#)
- ▶ [Sex Offender Treatment](#)

Recommended Reading and References

- Akers RL (1973) *Deviant behavior: a social learning approach*. Wadsworth, Belmont
- Akers RL (1990) Rational choice, deterrence, and social learning theory: the path not taken. *J Crim Law Criminol* 81:653–676
- Akers RL (1996) Is differential association/Social learning cultural deviance theory? *Criminology* 34:229–247
- Akers RL (1998) *Social learning and social structure: a general theory of crime and deviance*. Northeastern University Press, Boston
- Akers RL, Krohn MD, Kaduce LL, Radosevich M (1979) Social learning and deviant behavior: a specific test of a general theory. *Am Sociol Rev* 44:635–655
- Bandura A (1969) *Principles of behavior modification*. Holt Rinehart and Winston, New York
- Bandura A, Ross D, Ross SA (1961) Transmission of aggression through imitation of aggressive models. *J Abnorm Soc Psychol* 63:575–582
- Burgess RL, Akers RL (1966) A differential association-reinforcement theory of criminal behavior. *Soc Probl* 14:128–147
- Elliott DS, Menard S (1996) *Delinquent friends and delinquent behavior: temporal and developmental patterns*. In: Hawkins JD (ed) *Delinquency and crime: current theories*. Cambridge University Press, New York, pp 28–67
- Gottfredson M, Hirschi T (1990) *A general theory of crime*. Stanford University Press, Palo Alto
- Hirschi T (1969) *Causes of delinquency*. University of California Press, Berkeley
- Hirschi T (1996) Theory without ideas: reply to Akers. *Criminology* 34:249–256
- Jeffrey CR (1965) Criminal behavior and learning theory. *J Crim Law Criminol Police Sci* 56:294–300
- Kornhauser R (1978) *Social sources of delinquency*. University of Chicago Press, Chicago
- Matsueda RL, Anderson K (1998) The dynamics of delinquent peers and delinquent behavior. *Criminology* 36:269–308
- McGloin JM, Schreck CJ, Stewart EA, Ousey GC (2011) Predicting the violent offender: the discriminant validity of the subculture of violence. *Criminology* 49:767–794
- Mears DP, Ploeger M, Warr M (1998) Explaining the gender gap in delinquency: peer influence and moral evaluations of behavior. *J Res Crime Delinq* 35:251–266
- Nagin DS (1998) *Criminal deterrence research at the outset of the twenty-first century*. *Crime and justice: a review of research*, vol 23. University of Chicago Press, Chicago, IL
- Nettler G (1984) *Explaining crime*, 2nd edn. McGraw Hill, Boston
- Osgood DW, Wilson JK, O'Malley PM, Bachman JG, Johnson LD (1996) Routine activities and individual deviant behavior. *Am Sociol Rev* 61:635–655
- Pratt TC, Cullen FT (2000) The empirical status of Gottfredson and Hirschi's general theory of crime: A meta-analysis. *Criminology* 38:931–964
- Pratt TC, Cullen FT, Sellers CS, Winfree LT, Madensen T, Daigle LE, Fearn N (2010) The empirical status of social learning theory: a meta-analysis. *Justice Q* 27:765–802
- Sutherland EH (1947) *Principles of criminology*, 4th edn. Lippincott, Philadelphia
- Thomas KJ, McGloin JM (2013) A dual-systems approach to understanding differential susceptibility to different mechanisms of peer influence. *Criminology* (Forthcoming)
- Warr M (1993) Age, peers, and delinquency. *Criminology* 31:17–40
- Warr M (1998) Life-course transitions and desistance from crime. *Criminology* 36:183–216
- Warr M (2002) *Companions in crime: the social aspect of criminal conduct*. Cambridge University Press, Cambridge
- Weerman FM (2011) Delinquent peers in context: A longitudinal network analysis of selection and influence effects. *Criminology* 49:253–286
- Wood PB, Gove WR, Wilson JA, Cochran JK (1997) Non-social reinforcement and habitual criminal conduct: an extension of learning theory. *Criminology* 35:335–366
- Young JTN, Barnes JC, Meldrum RC, Weerman FM (2011) Assessing and explaining misperceptions of peer delinquency. *Criminology* 49(2):599–630
- Zimmerman GM, Vasquez BE (2011) Decomposing the peer effect on adolescent substance use: Mediation, nonlinearity, and differential nonlinearity. *Criminology* 49:1235–1273

Social Network Analysis and the Measurement of Neighborhoods

Rick Grannis

Department of Sociology, University of California, Irvine, Irvine, CA, USA

Synonyms

[Neighborhood effects](#); [Neighborhoods](#); [Neighboring](#); [Social networks](#)

Overview

Neighborhoods shape a variety of outcomes for children, families, and residents in general, influencing behavior, attitudes, and values as well. While some neighborhoods foster continuous patterns of criminal activity, others develop collective efficacy, the shared understanding that their constituent members have social capital resources which they are mutually able and willing to use to achieve collective outcomes. Neighborhoods cluster outcomes, some of which cannot be accounted for in terms of the characteristics of the individuals or households currently residing in them; they prove to be real communities with enduring characteristic patterns that survive the replacement of their constituent members. A useful neighborhood definition would be one that helped us better understand these neighborhood communities and their effects. These neighborhood communities are not only geographically meaningful but geographically identifiable as well because the networks of interactions among neighboring residents which produce them, which translate neighboring interactions into neighborhood communities and their effects, are constrained by predictable urban geographic substrates. New research has proposed behaviorally oriented definitions of neighborhoods, defining them in terms of their potential for interaction among residents. Defining neighborhoods in this way provides a lens to focus more closely on neighborhoods

as effect-generating communities emerging from the networked interactions of their constituent residents.

Fundamentals/Key Issues

Why Neighborhoods Matter in the Twenty-First Century: The Continual Emergence of Neighborhood Effects

Neighborhoods: Social Capital, Collective Efficacy, and Crime There appears to be a continually increasing interest in the role of neighborhoods in shaping a variety of outcomes for children, families, and neighborhood residents in general (for an overview, see Brooks-Gunn et al. 1997a, b). These “effects” have included a vast array of phenomena ranging from child and adolescent development (e.g., abuse and maltreatment, school completion and achievement, drug use, deviant peer affiliation, delinquency and gangs, adolescent sexual activity and pregnancy and childbearing, parenting behaviors) to concentrated disadvantage and its many corollaries (economic attainment and labor market success, crime and violence, physical disorder, the perpetuation of racism, to name just a few). An overwhelming conclusion reached by all of these studies of neighborhood effects is that neighborhoods influence our behavior, attitudes, and values. They shape the types of people we will become and expose us to or shield us from early hazards that would seriously restrict the opportunities available to us later in life. After our homes, and in conjunction with them, they are where we first learn whether the world is safe and cooperative or inchoate and menacing.

Not all neighborhoods are alike, however. Some neighborhoods are characterized by high levels of effective community. They clearly offer social capital to their residents, a social organization which facilitates and coordinates cooperative action for mutual benefit, which allows residents to deal with daily life, seize opportunities, reduce uncertainties, and achieve ends that would not otherwise have been possible. This social organization is a resource which is not individually attainable because social capital is not

a characteristic of individuals; it is a supra-individual property of social structure and it seems to be particularly well grounded in neighborhood communities. These sources of social capital tied to the neighborhood community context are analytically distinct from and as consequential as the more proximate family processes and relationships occurring in the home. Some neighborhoods develop a further layer, mutual trust and shared norms, values, and expectations, beyond the resource potential of neighbor networks, which allows them to utilize these networks to achieve desired outcomes. Collective efficacy occurs when members of a collectivity, with social capital resources, believe they are mutually able and willing to use these resources to achieve an intended outcome (Morenoff et al. 2001). The distinction is a subtle, but important, one. A neighborhood may have social capital resources available for its constituent members to utilize but they may not trust the willingness or ability of their fellow residents to use these networked resources for the collective good or they may not even be certain that they agree as to what the collective good might be.

From a less positive perspective, neighborhoods show remarkable continuities in patterns of criminal activity as well. For decades, criminological research in the ecological tradition has confirmed the concentration of interpersonal violence in certain neighborhoods, especially those characterized by poverty, the racial segregation of minority groups, and the concentration of single-parent families. Even in neighborhoods with less socioeconomic or racial isolation, crime rates persist despite the demographic replacement of neighborhood populations (Brantingham and Brantingham 1993). In addition, neighborhoods determine not only one's exposure to crime and violence but also a host of less tangible deleterious factors which contribute to the development of an urban underclass, signs of social disorder which lead residents to perceive their neighbors as threats rather than as sources of support or assistance (Massey and Denton 1993).

Neighborhoods: Geography and (Potential) Effects These neighborhood effects, both the

community enhancing and the community degrading effects, necessarily involve a geographic context. Thus, to analyze and understand them, they necessarily require a geographic equivalent of a neighborhood and a geographic definition of one. Researchers have utilized a variety of such definitions. In fact, they have used so many that Galster (2001, p. 2111) argued that "urban social scientists have treated 'neighborhood' in much the same way as courts of law have treated pornography: a term that is hard to define precisely, but everyone knows it when they see it." Apparently, however, researchers often don't know it when they see it. The modifiable area unit problem (MAUP) is a statistical bias affecting areal unit summary values (e.g., totals, rates, proportions) when they prove "arbitrary, modifiable, and subject to the whims and fancies of whoever is doing, or did, the aggregating" (Openshaw 1984, p. 3). Miller's (1999) survey suggests that, when the spatial units used to study a phenomenon are not clearly defined by theory, any conclusions derived about the studied phenomenon may be hopelessly prejudiced by the arbitrary, at least from a theoretical perspective, choice of spatial unit.

While many statistical techniques and error modeling approaches have been used to try and counteract, reduce, or remove the effects of MAUP, Miller perceives that the ultimate solution has to involve a more behavioral-oriented definition of neighborhood. One needs better intuitions about the general nature of neighborhoods, not better statistical methods. The very existence of the modifiable areal unit problem (MAUP) evidences that theory has taken a back seat. Researchers who have developed methods for creating optimal analytic units with respect to predefined objective functions note correctly that MAUP would be irrelevant if neighborhood equivalents were chosen for theoretical reasons rather than administrative convenience (Alvanides et al. 2001).

Despite this, however, when a geographic definition of neighborhood is required for the purpose of quantitative analysis, "most social scientists and virtually all studies of

neighborhoods . . . rely on geographic boundaries defined by the Census Bureau or other administrative agencies. . . [which] offer imperfect operational definitions of neighborhoods for research and policy” (Sampson et al. 2002, p.445). Administratively defined units such as census tracts and block groups do not directly measure, nor were they designed to measure, the potential for interaction among resident members, the primary process hypothesized to produce neighborhood communities and their effects. For example, the Census Bureau appears to theoretically focus on segregation type effects, defining tracts as “relatively homogeneous units with respect to population characteristics, economic status, and living conditions.” In most cases, however, the sheer ubiquity of data gathered by the Census Bureau or other administrative agencies (e.g., school districts, police districts) proves to be an overwhelming temptation for most researchers. Theory succumbs to the preponderance of data. The very existence of the modifiable area unit problem (MAUP), however, evidences that census geography is not measuring what researchers are studying.

Undoubtedly, neighborhood effects involve a geographic context. Neighborhood effects, however, are not produced by neighborhood geography. Neither are neighborhood effects, at least not a lot of them, merely a by-product or spurious confound of the geographic co-location of residents with particular demographic characteristics or psychological profiles. Neighborhood effects hypothesize that there exists a thing, a social entity, a neighborhood community, which has effects. Neighborhood effects are the product of these neighborhood communities. Neighborhood communities and their effects emerge from neighboring interactions among their constituent members. Neighborhood communities are geographically constrained because the interactions which produce them are geographically constrained. Neighborhood communities are both geographically identifiable and have effects which persist through the replacement of their residents because the networks of interactions which produce them, which translate neighboring interactions into neighborhood

communities, are constrained by predictable urban geographic substrates.

When we think about geography having an effect on community, it is because we believe that something about physical space affects something about how individuals interact within that space. Neighborhoods are more than colored boxes on a map or sets of geo-referenced variables for use in a geographic information system (GIS). A focus on maps, especially maps based on census or administrative geography, emphasizes those aspects of neighborhoods and their residents which can be effectively displayed or referenced to administratively defined polygons and ignores those which cannot. To understand the social-interactional aspect of neighborhoods, we may not necessarily have to think outside the box, but we do have to think about what’s inside of it, residents potentially interacting with each other as neighbors.

Neighboring and Geographic Neighborhoods While it may seem obvious, it is worth highlighting that, at its most fundamental level, neighboring is a proximity-dependent relation. When we say that someone is our neighbor, we are making a statement about them being proximal to us. Neighbors must, by definition, live close to each other; but what constitutes the geographic proximity or availability that defines neighboring?

Many studies have called attention to the strong role of extremely short distances in neighborly contacts. At least from an individual household’s perspective, the distances associated with neighboring are often effectively measured in feet and yards (Festinger et al. 1950). Residential propinquity’s influence on social interaction is typically limited to those who live within a few households away. What is most important is who lives next door, not who lives in the same census tract; who lives a few houses away, not who lives a few blocks away. Neighbors have also been defined to be people who live within walking distance (Grannis 2009). Walking distance, of course, varies by person, being much greater for some than others due to their age and physical fitness and even changing seasonally in some areas. No matter how far walking distance is,

however, for any particular person, the probability that someone will be identified as a neighbor declines rapidly with increased distance from one's home.

Because neighboring is so close at hand, it depends upon very subtle geographic features. Besides focusing on the number of houses or yards separating two households, a natural division, in both cognition and behavior, occurs at the face block. The face block includes all of the dwellings that front on the same street and are situated between only two cross streets (an exception would be cul-de-sacs which are face blocks delimited by only one cross street). The face block has been found to be an important socio-spatial unit (Suttles 1972). At one level, the face block includes virtually all neighbors who live either next door to each other or directly across the street from each other and most of those who live within a few house lengths; therefore, it could be viewed as simply a reflection of the more general effects of proximity. However, studies have shown that residents have more interaction with those on the same face block than they do with residents beyond an intersection, even if they were spatially closer to the latter group (Greenbaum 1982).

Not all face blocks, however, are oriented towards the pedestrian nature of neighboring. Some front on large arterial streets devoted to providing access for travelers while others front on smaller streets more devoted to local living space. Studies of peoples' perceptions of street life reveal that residents clearly perceive the difference between heavy traffic face blocks continuously filled with strangers which are used solely as thoroughfares and corridors between the local neighborhood and the outside world on the one hand and light traffic face blocks which form the basis of lively, close-knit communities where everyone knows each other and residents consider the boundaries between house and street space to be quite permeable on the other hand (Appleyard and Lintell 1986). A tertiary face block has been specifically designed and maintained by governing authorities to promote local and pedestrian traffic. The tertiary face block is a more or

less "natural" unit of face-to-face neighborly interaction (Suttles 1972). Tertiary face blocks are both oriented towards pedestrian travel and local residents, rather than outsiders who arrive by automobile or mass transit. Thus, tertiary face blocks are the types of face blocks most likely to give rise to social interactions (Rabin 1987). Tertiary face blocks provide a meeting place for neighbors (de Jong 1986). People use them for a host of activities including walking pets, riding bicycles, and chatting with neighbors. Shared tertiary face blocks provide a "permeable boundary" between households' private spaces. Tertiary face-block neighbors are "used for easy sociability and assistance when quick physical accessibility is an important consideration."

Face blocks terminate at intersections. An alternative way of thinking about this, of course, is that intersections connect face blocks with each other. The important question then becomes: Do they also connect neighbors and do neighbor networks terminate at intersections or do they bridge them to form larger structures? Intersections form a different metric than face blocks for measuring functional distance. Just as different types of streets differentially induce or fail to induce neighborly relations, different types of intersections may induce or inhibit neighborly relations from bridging them. Fortunately, the intersections of streets can also be operationally defined in a convenient and meaningful way. When streets of different classifications intersect, planners consider the intersection to be of the higher classification. For example, if a larger street intersects a tertiary street, planners would consider the intersection to be part of the larger street, but not part of the tertiary street. This makes intuitive sense because the inhibiting effects of the larger street will dominate. The nature of intersections either facilitates or impedes pedestrian-based neighborly interaction; therefore, an intersection is a tertiary intersection if all of the face blocks, however many, contiguous with it are tertiary face blocks. A non-tertiary intersection, in contrast, is an intersection such that at least one of the face blocks contiguous with it is a non-tertiary face block.

Studies have shown that tertiary intersections combined with tertiary face blocks can serve as a “pedestrian circulation system.” Sidewalks provide access between residence and parks, churches, and neighborhood shops. Neighborly relations bridging face blocks occur “through the routes people take in meeting an average day’s basic needs and desires. The newsstand where one buys the Sunday paper, the store one runs to for a quart of milk, and the streets one travels on to visit a friend” (Anderson 1992, p. 46). People come to envision their neighborhoods as networks of paths and channels along which they move (Lynch 1971). Tertiary intersections guide this “natural movement” (Hillier 1996) within a city.

Potential Neighbor Networks and Actual Neighbor Networks Individual neighbor networks evolve into neighborhood networks through the process of concatenation. Residents have relations with their neighbors who interact with other neighbors, and so on. These neighborly relations concatenate and consolidate neighbor-to-neighbor-to-neighbor. There are several important corollaries of this fact. First, the resultant network is as far reaching as its most extensive ramification. Relations concatenate to form a network typically larger, both relationally and geographically, than any individual’s relations. Thus, relatively micro-level relations can result in a macro-level structure. Second, the resultant network is as fragile as its weakest link. Anything which can cause a relation to not form, no matter how trivial, breaks the network. In contrast to the first corollary, micro-level fragilities can destroy a macro-level structure. Third, the characteristics of the resultant network are not readily predictable from the characteristics of the local networks which concatenate to form it. Only as sets of individual networks concatenate do the characteristics of this aggregated network emerge. Finally, because a neighboring relation cannot exist unless residents are geographically available to each other, the network of potential neighbors cannot transcend the network of geographic availability; it is logically impossible. While individuals’ lifestyles and habits may prevent them from having

contacts and interactions with those who are geographically available to them, they cannot cause them to have contacts with those who are unavailable.

To study efficacious neighborhood communities emerging from neighbor networks, therefore, we need a definition of a neighborhood community whose importance is derived from the potential for neighbor networks to concatenate within it. Grannis (2009) defined this geographic availability in terms of shared walking arenas which mediate, guide, and constrain potential neighborly encounters. Building on this, the concatenated network of overlapping neighborly contacts can be no larger than the concatenated network of walking arenas; conversely, the network of potential neighborly relations, based on concatenated interactions, is a subset of the concatenation of these walking arenas. Grannis argued that tertiary block faces effectively proxy walking arenas in urban areas and thus the maximal concatenation of contiguous tertiary block faces, of walking arenas, represents the maximal consolidation of individual residents’ potential to access each other.

Such a neighborhood equivalent would signify internal access. All residents within it would have a potential for neighborly relations using walking arenas. While it is unlikely that all, or even any, residents would traverse the entirety of this neighborhood equivalent, this internal contiguity would allow residents to interact with their neighbors down the street who interact with other neighbors further down the street, and so on throughout the network. Such a neighborhood equivalent would also signify constraint. To the extent to which the potential for neighboring relations depends upon walking arenas, it would constrain their concatenation. Based on these criteria, Grannis (2009) defined two types of neighborhood equivalents, t-communities and islands, one connecting tertiary block faces using only tertiary intersections and the other connecting tertiary block faces using all intersections. A t-community is a maximal contiguous network of tertiary block faces and tertiary intersections, while an island is a maximal contiguous network of tertiary face blocks and any

intersections. The label “t-community” is short for “tertiary street communities” indicating that these neighborhood equivalents are connected by tertiary streets (Grannis 2009). A pair of neighborhoods would be geographically unreachable if they could only reach each other by crossing water; Grannis (2009) analogously label these discontinuous networks of tertiary block faces as “islands” since it is impossible for households in two different islands to access each other using tertiary block faces, even if they use non-tertiary intersections. Since islands are maximal networks of tertiary block faces and any intersections while t-communities are maximal networks of tertiary block faces and tertiary intersections, t-communities are necessarily subsets (although not necessarily proper subsets) of islands.

Selection vs. Influence These t-communities and islands, these walking arenas, interact with two different social forces to create neighborhood communities and their effects. First, neighborhoods are more than just neighbors residing nearby each other. They are vital entities, or at least they have the potential to be. Even when it appears static, neighborhood social life generally is not; it is a stable equilibrium reached amidst the strife of social flows. The vibrant, living part of neighborhoods consists of the flow and exchange, both spoken and silently modeled, of norms, values, identities, symbols, ideas, affect, sentiment, and other social and cultural goods and resources among neighbors along the conduits provided by neighbor networks. This flow pressures neighbors towards conformity.

This certainly happens verbally, through the exchange of personal information, life histories, and stories, as well as through establishing and enforcing rules for neighborhood children; however, it happens even more nonverbally. Community members model and enforce “appropriate behaviors” in their daily interactions with each other, especially with children (Coleman 1990). In addition to modeling appropriate behaviors, neighbors may use rewards to encourage normative behaviors or sanctions to discourage behaviors not compliant with social or personal norms. The degree of these sanctions or rewards varies greatly with the nature of the society; for

example, in communities that emphasize social control and social cohesiveness, sanctions may be enforced by direct social pressure for conformity (Hogan and Kitagawa 1985). As norms, values, ideas, and other social goods and resources traverse and commingle along neighbor networks, they have the potential to engender a sense of community and identity, social capital, mutual trust, social control, collective efficacy, and many other important facets of neighborhood life social researchers interest themselves in. The collectively efficacious community network which emerges, or which fails to emerge, however, is embedded in the network of neighborly interactions which is embedded in the network of potential neighborly interactions which is embedded in the network of geographic availability.

Second, locational choice and homophily may certainly account for some of the effects of neighborhoods but numerous studies have shown that neighborhoods with similar population demographics, in terms of race, socioeconomic status, family structure, and a host of other characteristics, often yield different outcomes for their constituent members. Market explanations, while intuitively appealing, have failed to account for the richness and complexity of effects correlating to neighborhoods.

Locational-based neighborhood effects such as residential differentiation and segregation correspond to the influence-based neighborhood effects such as social capital and collective efficacy because in choosing to move away from dissimilar households, residents are implicitly choosing to segregate their networks of potential neighborly interactions as well. Since contact is a necessary prerequisite for interaction, if households settle in such a way that their immediate neighbors are similar to themselves, then they have settled in such a way as to not have neighborly interactions with those different from themselves. Neighborhood communities result from both the concatenation of homophilous locational choices and the flow and exchange of norms, values, and beliefs among neighbors. Their correspondence is not additive, as in a regression model, but rather sequential. Relocation, which is responsible for residential differentiation and

segregation, determines geographic availability and the potential for neighborly interactions and thus, of necessity, the actualized neighborly interactions which influence works upon to create social capital, collective efficacy, and other important neighborhood effects.

Children and Their Families Neighborhoods are especially important for households with children because children are much less mobile, and thus more geographically dependent, than adults. Children and their playful interactions depend upon proximity much more than adults and their interactions do. Since children cannot drive and have little, if any, voice in relocation decisions, they are forced to share lives with neighboring children even more than are their parents. For children, the street in front of their home is “the mediator between the wider community and the private world of the family” (Appleyard 1981, p. 4). This is where children first learn about the world. They often play games in the middle of these streets, use them to walk pets and to ride bicycles, and the majority of their recreational activity occurs there (Brower 1977). Sidewalks provide access between residence and schools and parks. As a result, the relationships children form will primarily depend upon the opportunities to interact provided by walking arenas immediately surrounding them (Appleyard and Lintell 1986). Especially for young children, neighboring children are the most likely to become their playmates (Hillier 1996). Thus, the networks of relationships they form will be much more dependent upon the network of geographic availability. Unlike children, adults have many venues for social relationships beyond their neighborhood including work and voluntary activities. School-age children may have some of these to the extent their parents allow. Preschool children, however, have few, if any, of these alternative venues for social opportunities. Their lives are tightly bound by geography.

“The micro-ecology of pedestrian streets bears directly on patterns of interaction that involve children and families. Parents are generally concerned with demarcating territory outside of which their children should not wander

unaccompanied by an adult, to ensure that their children stay in areas that are safe for play and conducive to adult monitoring. To the extent that these limited spaces of children’s daily activities usually do not cross major thoroughfares, defining tertiary communities may provide a foundation for constructing neighborhood indicators of child well-being and social processes more generally” (Sampson et al. 2002).

Not only are your neighbors’ children predisposed to become your children’s friends, but they also determine the character of your children’s playmates (Cochran 1994) and the kinds of role models they emulate (Massey and Mullen 1984). The flow of norms and values discussed above acts not only on children but their families as well. “For example, when parents know the parents of their children’s friends, they have the potential to observe the child’s actions in different circumstances, talk to each other about the child, compare notes, and establish norms. Such intergenerational closure of local networks provides the child with social capital of a collective nature” (Sampson 2001: 9). As a result, households with children are far more influenced by the norms and values of surrounding households with children than households in general are influenced by the norms and values of their surrounding neighbors.

Neighboring parents may become intimately involved in the socialization of each other’s children. Neighbors rear children side by side and together have the potential to co-create a safe and value-laden environment. Parents monitor their own children as well as those of their neighbors (Sampson et al. 1999). Some neighborhoods expect that residents share values and are willing and able to intervene on the behalf of children. They expect that residents will actively engage themselves in the support and social control of children (Bandura 1997) and that the community will work together to successfully support and control children. Parents get to know the parents and families of their children’s friends; they observe children’s actions, both their own and their neighbors’, in a variety of circumstances; they talk with other parents about their children; and they establish norms (Coleman 1990). Such

structural and normative adult-child closure gives children social support, provides parents with information, and facilitates control (Sandefur and Laumann 1998). The choice to live in a neighborhood is to some extent a choice to rear children together with one's neighbors. Ultimately, a community of parents may develop around the community of children, mirroring it. People whose children play together form friendship relations based in part on that fact (Abu-Gazze 1999; Grannis 2009). While it is the children who are immobile and thus confined to neighborhoods that are most immediately impacted by neighborhoods, children's geographic dependence encumbers their parents as well.

The Menu of Neighborhood Equivalents

Altogether, for both children and adults, through influence and selection, neighborhoods cluster outcomes which cannot be accounted for in terms of the characteristics of the individuals or households currently residing in them. It is as if neighborhoods have personalities, enduring characteristic patterns that survive the replacement of their constituent members. Neighborhood communities are geographically constrained because the interactions which produce them are geographically constrained. Neighborhood communities are both geographically identifiable and have effects which persist through the replacement of their residents because the networks of interactions which produce them, which translate neighboring interactions into neighborhood communities, are constrained by predictable urban geographic substrates. To study efficacious neighborhood communities emerging from neighbor networks, therefore, we need a definition of a neighborhood community whose importance is derived from the potential for neighbor networks to concatenate within it.

Grannis (2009) defined this geographic availability in terms of shared walking arenas which mediate, guide, and constrain encounters. What do neighborhoods such as t-communities and islands, defined and measured by their potential for interactions, offer us that traditional neighborhood equivalents do not? They provide us with a lens to focus more closely on

neighborhoods as communities emerging from the interactions of their constituent residents. They use an entirely different metric than census geography, one based precisely on the potential for community generating neighborly relations. In contrast, administrative geography often focuses on neighborhoods as statistical abstractions, perhaps reflecting segregation but agnostic to any potential for community, for interaction, for neighboring. While both types of neighborhood equivalents have their uses, researchers need to use care as to which one they choose and perhaps use both to disentangle different mechanisms that are at work in neighborhood communities, one mechanism provided by the concatenation of neighboring relations into neighbor networks and another provided by service areas, such as those offered by schools or marketplaces or police which unite residents around similar needs and opportunities. T-communities precisely measure the first. Neighborhood equivalents, defined solely by their boundaries, measure the second to the extent their boundaries coincide with the service areas. A careful use of both t-communities and neighborhood equivalents defined by their boundaries could tease apart the different mechanisms at work (Grannis 2009).

Related Entries

- ▶ [Neighborhood Effects and Social Networks](#)
- ▶ [Social Network Analysis of Urban Street Gangs](#)
- ▶ [Spatial Models and Network Analysis](#)

Recommended Reading and References

- Abu-Gazze TM (1999) Housing layout, social interaction, and the place of contact in Abu-Nuseir, Jordan. *J Environ Psychol* 19:41–73
- Alvanides A, Openshaw S, Macgill J (2001) Zone design as a spatial analysis tool. In: Tate NJ, Atkinson PM (eds) *Modeling scale in geographical information science*. Wiley, Chichester, Chapter 8
- Anderson E (1992) *Streetwise: race, class, and change in an urban community*. University of Chicago Press, Chicago

- Appleyard D, Lintell M (1986) The environmental quality of city streets: the residents' viewpoint. In: de Boer E (ed) *Transport sociology: social aspects of transport planning*. Pergamon, Oxford, pp 93–120
- Appleyard D (1981) *Livable streets*. University of California, Berkeley
- Bandura A (1997) *Self-efficacy: the exercise of control*. Freeman, New York
- Brantingham PL, Brantingham PJ (1993) Nodes, paths, and edges: considerations on the complexity of crime and the physical environment. *J Environ Psychol* 13:3–28
- Brooks-Gunn J, Duncan GJ, Aber JL (eds) (1997a) *Neighborhood poverty, vol I, Context and consequences for children*. Russell Sage, New York
- Brooks-Gunn J, Duncan GJ, Aber JL (eds) (1997b) *Neighborhood poverty, vol II, Policy implications in studying neighborhoods*. Russell Sage, New York
- Brower SN (1977) *Streetfronts and backyards: two ways of looking at neighborhood open spaces*. Department of Planning, Baltimore
- Cochran S (1994) Understanding and enhancing neighbourhood sense of community. *J Plann Literat* 9(1):92–99
- Coleman JS (1990) *Foundations of social theory*. Harvard University Press, Cambridge
- de Jong R (1986) The recapture of the street? In: de Boer E (ed) *Transport sociology: social aspects of transport planning*. Pergamon, Oxford, pp 77–91
- Festinger L, Schachter S, Back K (1950) *Social Pressures in informal groups: a study of human factors in housing*. Harper and Row, New York
- Galster GC (2001) On the nature of neighborhood. *Urban Stud* 38(12):2111–2124
- Grannis R (2009) *From the ground up: how the layered stages of neighbor networks translate geography into neighborhood effects*. Princeton University Press, Princeton
- Greenbaum S (1982) Bridging ties at the neighborhood level. *Soc Networks* 4:367–384
- Hillier B (1996) *Space is the machine*. Cambridge University Press, Cambridge
- Hogan DP, Kitagawa EM (1985) The impact of social status, family structure, and neighborhood on the fertility of black adolescents. *Am J Sociol* 90:825–852
- Lynch K (1971). *The image of the city*. MIT Press, Cambridge
- Massey DS, Mullen B (1984) Processes of hispanic and black spatial assimilation. *Am J Sociol* 89:836–873
- Massey DS, Denton NA (1993) *American apartheid: segregation and the making of the underclass*. Harvard University, Cambridge
- Miller H (1999) Potential contributions of spatial analysis to geographic information systems for transportation. *Geogr Anal* 31:373–399
- Morenoff JD, Sampson RJ, Raudenbush SW (2001) Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence. *Criminology* 39(3):517–560
- Openshaw S (1984) *The modifiable areal unit problem*. Geo Books, Norwich
- Rabin Y (1987) The roots of segregation in the eighties: the role of local government actions. In: Tobin GA (ed) *Divided neighborhoods: changing patterns of racial segregation*. Sage, Newbury Park, pp 208–226
- Sampson RJ, Morenoff JD, Earls F (1999) “Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children”. *Am Sociol Rev* 64:633–660
- Sampson RJ (2001) How do communities undergird or undermine human development? Relevant contexts and social mechanisms. In: Booth A, Crouter N (eds) *Does it take a village? Community effects on children, adolescents, and families*. L Erlbaum, Mahwah, pp 3–30
- Sampson RJ, Morenoff JD, Gannon-Rowley T (2002) Assessing “neighborhood effects”: Social processes and new directions in research. *Ann Rev Sociol* 28:443–478
- Sandefur RL, Laumann EO (1998) A paradigm for social capital. *Rational Soc* 10:481–501
- Suttles G (1972) *The social construction of communities*. University of Chicago Press, Chicago

Social Network Analysis of Organized Criminal Groups

Francesco Calderoni

Transcrime, Università Cattolica del Sacro Cuore, Milan, Italy

Overview

This entry discusses the state of the art in the application of network analysis methods to the study of organized criminal groups. It provides an overview of the development of the field from both the academic and law enforcement perspective and discusses the current approaches in the literature. Particular attention is given to methodological issues (data sources, type of network analysis) and to the discussion of the limitations of the application of social network analysis to criminal organizations. In the light of the current state of the art, the entry also discusses the recurrent claim that network analysis may help the law enforcement agencies to more efficiently disrupt criminal networks. Finally, it attempts to identify the future trends in the application of network

methods in the study of organized crime, while suggesting some promising paths from both a research and policy perspective.

Introduction

The use of social network analysis (SNA) in criminology is a relatively recent trend compared to other social sciences. Among the multiple possible applications of SNA in criminology, it may be hard to identify trends and specific fields. Carrington (2011) has identified three main topic areas where network methods have found significant application in criminology: The first is the analysis of the influence of personal networks on delinquency and crime which, according to the author, is “the most common use of social network analysis in criminology” (Carrington 2011, 236). The second area is the influence of neighborhood networks on crime in the neighborhood. The third area is the organization of criminal groups and activities. In turn, the latter may be framed into three fields: The first concerns the application of SNA methods to terrorist networks; the second deals with street gangs, youth gangs, and delinquent groups (Carrington 2011, 244–246); and the last, discussed in this entry, concerns organized criminal groups.

The application of SNA methods to organized crime has a background which is based on the history of both research and law enforcement.

From the research side, the idea that organized crime may be better understood as a network rather than a hierarchical and structured organization is not new in criminological literature. The reaction against the alien conspiracy approach, which suggested to analyze organized crime as a bureaucratic organization, organized along a formal hierarchy and with detailed rules for its functioning, soon led to the alternative hermeneutic perspectives leaning towards more flexible and informal mechanisms. For example, Albin argued that organized crime is “a system of loosely structured relationships,” mainly based on patron-client relations (1971, 288); the works of Ianni suggested that mafia-type organizations should be better understood as social systems

based on shared social cultural and ethnical relations. He explicitly mentioned network analysis, as “an anthropological tool that is used to chart social interactions” (Ianni 1973, 4). Ianni analyzed a number of Puerto Rican and African-American criminal groups as networks, although the application was quite far from current use. Overall, although the concepts of network and network analysis were recurrently evoked to describe the functioning and structure of organized crime, there were very limited empirical applications using network analysis methods (Ianni 1973).

From the law enforcement side, since the mid-1970s there has been a growing attention on the processing and analysis of intelligence data in organized crime. In this context, link analysis (visual representation of the structure of a criminal group performed through manual or computer-assisted drawings) was increasingly applied by law enforcement agencies and a private industry quickly developed offering methodologies and training courses (Lupsha 1983, 63). Link analysis allowed to “establish the relationships that exist among individuals and organizations from bits and pieces of available evidence” (Harper and Harris 1975, 158). It became increasingly popular thanks to the development of intelligence software. The step from link analysis to SNA was relatively short.

Academic interests on organized crime networks gradually met with law enforcement network analysis attempts. Not surprisingly, Ianni and Reuss-Ianni, who had refuted the bureaucratic approach to organized crime for the social system approach, contributed to a volume on criminal intelligence analysis suggesting it be applied to organized crime (Ianni and Reuss-Ianni 1990). Indeed, from the very beginning, this idea was linked with its operational exploitation for law enforcement, at a time when the focus on strategic intelligence and analysis of organized crime was very strong.

Indeed, the first contributions in this field, dating back to the first half of the 1980s, already focused on the opportunities offered by network analysis in the strategic analysis and enforcement of criminal organizations (Davis 1981; Lupsha

1980, 1983). In his 1980 contribution Peter Lupsha suggested that effective enforcement of organized crime required a move towards strategic analysis of groups and operations. He suggested some “first steps that departments might engage in to move from a tactical to a strategic analysis perspective” (1980, 37). Among such steps, he included network analysis as an “essential and necessary step” (1980, 38). One year later, FBI Special Agent Roger Davis provided a first hypothetical example of the application of network methods to a criminal organization, using some basic network concepts, such as density and centrality, to a fictional criminal organization (Davis 1981, 18). One of the first empirical network analyses of a criminal organization was carried out by Lupsha (1983). The study focused on “The New Purple Gang,” a group trafficking in heroin and cocaine active in the New York City metropolitan area in the second half of the 1970s and composed of Italian-Americans with connections to La Cosa Nostra. The author maintained that his study was an application of network analysis and actually developed matrixes and graphs. Rather, the study merged broad concepts familiar to network analysis, such as reciprocity and ego network structure, with other approaches and data. Interestingly, some of the most common network analysis concepts, such as centrality, were barely mentioned. As a result, Lupsha’s study provided a first exploration of network analysis of organized crime, demonstrating how multiple methods could help to extract relevant information from police data.

In the following years, some scholarly contributions advocated the application of network methods to criminal organizations (among the first contributions Ianni and Reuss-Ianni 1990; Sparrow 1991). However, most of these nowadays “classic” contributions, albeit with some exceptions (e.g., Lupsha 1983), did not engage in empirical analysis of criminal groups, possibly due to the limited availability of datasets and software. Interestingly, Mastrobuoni and Patacchini (2010) and Papachristos and Smith (2011) have recently demonstrated that old law enforcement databases and archives may provide

valuable opportunities for the application of SNA to organized crime. Not surprisingly, only 10 years ago, Coles still complained about the “failure by criminologist to adopt Social Network Analysis techniques and concepts in the investigation of criminal networks, particularly of organized crime”.

Only in the last decade the use of network analysis in the study of organized crime has seen significant developments. Since the beginning of the 2000s, interest in this specific field has significantly increased and contributed to opening new research directions in the study of criminal organizations. Indeed, the number of articles and papers on criminal, illicit, or, more generally, “dark” networks has grown only since the 2000s. Significantly, a special issue of a specialized academic journal was dedicated to SNA and organized crime (*Trends in Organized Crime*, Volume 12, Number 2/June 2009). Studies on criminal organizations are now an important sector in this increasing trend. The following sections of this entry attempt to review the current state of the art of the application of SNA to organized crime groups and to discuss present issues and future trends.

Criminal Organizations and Markets

In a very rough categorization, most current applications of SNA to organized crime have adopted either a “micro” or a “macro” approach.

Studies within the “micro” approach have normally focused on one organization or network with a relatively low number of nodes. One of the first examples was Natarajan’s study of a cocaine trafficking organization (2000). The study provided a detailed analysis of the structure and functioning of a criminal organization, through a mix of multiple methods (contact analysis, task analysis, status analysis, and network analysis). As to SNA, Natarajan applied some network concepts (centrality and density) and demonstrated that a systematic quantitative analysis of law enforcement and judicial sources could complement more traditional research methods in the study of higher-level drug trafficking.

In the following years, further empirical works have contributed to analyzing criminal organizations using network analysis. Most of them focused on drug trafficking (Morselli 2005, 2009; Calderoni 2012; Malm et al. 2011; Bright et al. 2012), but some studies analyzed carrying organizations (Morselli and Roy 2008), extortion/money-laundering activities (Varese 2006), or the structure or evolution of a prominent criminal within a New York mafia family (Morselli 2005).

As to drug trafficking, the application of SNA has provided unprecedented insight in the structure of trafficking organizations. In most drug-trafficking networks, a subset of nodes concentrated the majority of criminal contacts and communications. This emphasized that drug trafficking was mostly conducted by a subset of highly connected and active individuals who were able to bring different people and resources together. In general, these subjects were skilled criminal brokers tasked with the management of drug-smuggling operations. Conversely, the majority of nodes involved in drug-trafficking groups appeared to be marginal players, involved only in one specific operation or task. On the one side, the results revealed a specific organizational structure of drug trafficking. This was a rather horizontal structure, ideal for a flexible and fast-changing environment, where criminal activities and illicit partnerships depended more on criminal opportunities than on any internal organizational arrangement. On the other side, the mentioned structure did not correspond to an extremely elaborated hierarchy where different organizational layers/ranks could be observed and where bosses delegated to underbosses or capos who in turn command soldiers. Signals of structure and hierarchy did not emerge from the network analysis, but rather from complementary methods highlighting the informal division of tasks and differences in status among the participants. Overall, leaders/brokers were actually in contact with most of the other individuals in the criminal networks and maintained direct control over the criminal activities. Hardly any traces of delegation and rigid organizational layers were found, and this is probably due to the highest risks

that may be faced due to information dispersion, lack of trust, and difficulties in controlling subordinates.

This picture was substantiated by evidence coming from different studies on drug networks from a variety of countries and typologies of criminal organizations, such as the United States, Canada, Spain, Australia, the UK, and Italy. These findings were consistent with, and provided further empirical evidence to, previous drug-trafficking research.

Studies within the “macro” approach have developed in very recent years, possibly because they were based on larger and more complex databases. They have generally analyzed larger networks, focusing on specific national or regional criminal markets or offender categories. This research has aimed at analyzing specific aspects of criminal networks and population. Malm and Bichler studied a network of 2,197 individuals derived from law enforcement intelligence data in Canadian Pacific Region (Malm et al. 2009; Malm and Bichler 2011; Malm et al. 2011). They used multiple networks (co-offending, kinship, formal organization, and legitimate association) and this allowed to focus on the interplay of different types of connections (e.g., family relation or partnership in legal business) within a large network of individuals involved in criminal organizations (Malm et al. 2009). In another study, they focused on a subset of 1,696 individuals involved in drug trafficking and analyzed the network characteristics of specific market niches (i.e., specific tasks within the drug-trafficking chain) (Malm and Bichler 2011). They have also explored the extent of co-offending patterns among criminal groups of different (ethnic or nonethnic) origin (Malm et al. 2011). Heber (2009) studied a sample of 127 serious drug offenders (principal offenders) in Sweden and further enlarged her network including the principal offenders’ co-offenders and eventually the co-offenders of the latter. As a result, her network included several thousands of offenders allowing an analysis of their distribution across the Swedish territory and of the structure of drug-trafficking activities. Some further contributions focused on American

organized crime, namely, in Al Capone's Chicago (Papachristos and Smith 2011) and La Cosa Nostra in 1960 (Mastrobuoni and Patacchini 2010).

Overall, network studies have pointed out that the majority of drug-trafficking (and other criminal) groups are flexible and rapidly changing organizations, hardly fitting the typical pyramidal picture which is frequently suggested by the media and sometimes by law enforcement. Network analysis has contributed critically to the production of empirical evidence in this field.

Methodologies

Compared to the most advanced application of SNA in other social sciences, the methods currently applied in most criminological contributions may appear somewhat simpler. This may be due to the relatively recent development of network approaches in this specific field. In turn, this may imply a limited experience of most criminologists in the application of SNA but also, and more critically, a limited scientific acceptance of such methods within the discipline. Not surprisingly, in the past years most criminologists applying SNA have experienced difficulties in publishing in mainstream peer-reviewed journals. Another reason for the relative lag in the application of more complex network methods to organized crime may inherently relate to the this specific field. Indeed, available data on criminal organizations and criminal markets rarely allow to apply more complex analyses such as dynamic network analysis (DNA), multiplexity (multiple types of relations, e.g., co-offending, phone communications, and kinship), and exponential random graph models which would allow to study the evolution of organized crime networks in time and structure. Notwithstanding these difficulties, some studies have applied ERGMs and multiplex relations (Malm et al. 2009) and one recent paper announced further developments in the field of organized crime (Papachristos and Smith 2011).

Concerning data sources, the majority of studies relied on secondary data. Most micro studies

used judicial documents, intelligence reports, or investigation files. Macro studies have used intelligence databases collected by law enforcement agencies (Malm and Bichler 2011; Malm et al. 2009; Heber 2009) or databases created through archival analyses (Mastrobuoni and Patacchini 2010; Papachristos and Smith 2011).

As for the types of SNA, most of the literature on organized crime groups has worked with complete (or whole or full) networks rather than the ego networks, differently from other main areas identified by Carrington (2011). Indeed, the complete network approach is naturally suited to the analysis of criminal organizations or criminal markets. Nevertheless, some studies used ego network approaches for studying the career of a prominent La Cosa Nostra member (Morselli 2005) or the impact of drug dealers' personal networks on their survival and success (Bouchard and Ouellet 2011) and the co-offending patterns among different typologies of criminal enterprises (Malm et al. 2011).

Within the full network perspective, the overwhelming majority of the reviewed contributions have focused on one-mode networks with individuals as nodes. Exceptionally, some contributions used two-mode networks with individuals and specific positions in the drug-trafficking chain (Bouchard et al. 2010), while another study, focusing on the world drug-trafficking market, elaborated a full network with world countries as nodes (Boivin 2011).

As already pointed out by Carrington (2011, 244), most network analyses of organized criminal groups are not particularly concerned with the testing of specific criminological theories. Nevertheless, most studies are concerned with the verification of hypothesis and concepts suggested by previous literature on organized crime (Natarajan 2000, 2006; Varese 2006; Morselli 2009; Malm et al. 2011). The existing literature shows a wide variety of approaches and different methods as to the network measurements and concepts. The most popular concepts are density and centralization, the analysis of subgroups (e.g., factions, cut-points, core-periphery), and the centrality of individuals (e.g., degree, betweenness, flow betweenness, eigenvector,

closeness, clustering coefficient). Some contributions have merged network analysis with other methodologies such as script analysis (Morselli and Roy 2008), the content analysis of transcripts (Natarajan 2000; Varese 2006; Natarajan 2006; Calderoni 2012; Campana 2011), or spatial analysis (Malm et al. 2008).

Finally, a number of contributions discussed the theoretical, methodological, and practical repercussions of the application of SNA to organized crime groups (e.g., Morselli 2009). These works signal the ongoing meditation about the implications of network analysis in the study of organized crime.

Problems and Issues

One of the main problems encountered in studies applying social network analysis to organized crime is data validity and reliability. The different data sources used in the literature are all exposed to possible biases. Indeed, most of them are secondary sources originating from judicial cases, law enforcement, or intelligence databases. Inevitably, the analysis based on such sources will reflect the perception of law enforcement. Interestingly, however, most of the abovementioned problems seem to be inherent to organized crime research in general. The biases affecting network analysis are often likely to occur also when using alternative methodologies. Indeed, scholars using network methods frequently reported difficulties in controlling for law enforcement biases, and they discussed the possible limitations of their analyses in detail (Morselli 2009, 44–50; Bouchard and Ouellet 2011, 83–85; Varese 2006, 45–47; Calderoni 2012; Malm et al. 2009, 70–71). A recent contribution further explored how the selection of different judicial sources affects the extracted networks (Berlusconi 2013). In general, authors appeared aware of the limitations of their studies. This has not spared criticism and skepticism towards the uncritical use of network methods in the study of organized crime. Undeniably, some contributions have exploited the growing interest in network analysis and have inserted it in

the titles, abstract, or keywords, but without actually using any network analysis concept. This may suggest that the increasing use of network analysis may have generated some side effects, not excluding the unnecessary abuse of network-related terminology and techniques, just because this has become “fashionable.” Besides such collateral effects, critics have rarely pointed out what alternative methods could replace the advantages provided by SNA. In fact, an increasing number of scholars are using network methods and concepts and applying them to the study of criminal organizations in different countries. This general trend suggests that SNA may actually bring significant added value to this field.

Concerning the risk of biases related to the law enforcement perception of crime, this appears to be a major issue particularly for studies adopting a macro approach. When trying to analyze regional or national markets, these studies most often reflect law enforcement knowledge of markets and criminal networks. The validity and reliability of the sources relate to the effectiveness of law enforcement and to the quality of the databases. Inevitably, however, any law enforcement source is inherently incomplete. Individuals and activities not (yet) uncovered or investigated by law enforcement will be missing (Sparrow 1991, 268; Bouchard and Ouellet 2011, 84; Malm et al. 2009, 70–71).

Also studies adopting a micro approach, focusing on specific groups or networks, may be affected by the mentioned problems of validity and reliability. However, these problems may have a lower impact than in the case of macro studies. Focusing on particularly long and detailed investigations is likely to reduce the risk of missing data in a criminal network. When law enforcement has been intercepting and monitoring a criminal group for months, if not years, the chances for some skilled criminal to avoid detection are likely to be low (Morselli 2009; Calderoni 2012). Nevertheless, the risks that law enforcement may have misunderstood or misjudged the relevance of specific nodes remain and so also the possibility of a willing omission by the police (e.g., this may be the case of informants or undercover agents which have not yet been disclosed).

Interestingly, recent studies have demonstrated that network properties and measures are strong even if tested for missing data (Morselli 2009, 48; Malm et al. 2009, 70–71). Although these findings are encouraging as to the validity and reliability of network methods applied to organized crime, there is a need of further research in this direction.

Social Network Analysis as a Tool for the Enforcement of Criminal Organizations

As already discussed, one of the main drivers of the development of network studies on criminal organizations has always been its possible application in law enforcement. Scholars have claimed the operational benefits of network analysis for intelligence and investigations for years. Given the increasing interest in network analysis of organized crime, a number of law enforcement agencies have started to apply SNA in their criminal investigations, and some contributions from law enforcement analysts acknowledged the use of network methods in the enforcement of criminal organizations. However, operational results so far appeared quite weak and many law enforcement agencies showed some skepticism towards the application of network methods in organized crime cases. Although this may be related to some difficulties in cooperation between law enforcement and academics in general (an issue which is not limited to the application of network analysis and which goes largely beyond the scope of this entry), discussion with law enforcement analysts (from Germany, the United Kingdom, the Netherlands) who have actually experimented network analysis in their job provides more detailed information. According to their experience, the current applications of network analysis are not likely to provide critical advantages to law enforcement agencies in their everyday investigations. Indeed, in long-lasting investigations, the police normally acquires a detailed knowledge of the case. Each individual is monitored, her/his background scanned, and phone and e-mails are frequently intercepted. The application of

SNA can hardly provide any additional knowledge to the officers and prosecutors working daily on the case. The idea that SNA can effectively identify the most important actors and show the police the most efficient way to disrupt a criminal network does not take into account that the police normally has a far better insight on the case than any researcher. More often than not, scholars identify the most important individuals, the best targets for network disruption, and the “usual suspects” that law enforcement had already identified.

This may explain the skepticism so far shown by some law enforcement professionals (some of which are increasingly being trained in SNA methods) when claims are made that SNA can help the police to identify the best targets. Network analysis can assist a researcher in reconstructing a network and in making sense of the structure and functioning of a criminal organization. While these analyses have provided unprecedented academic insight on the structure and functioning of criminal organizations, the risk is that current applications may have limited interest for law enforcers.

The above considerations should not lead to hasty conclusions about the futility of SNA and the enforcement of organized crime. Rather, they suggest that the exchange of experiences between academics and law enforcement should be improved. Notwithstanding the operational limits of current analyses, SNA concepts and methods have gained the interest of many law enforcement agencies around the world. As with most new techniques, it will take time to have it accepted and applied. The future trends discussed in the next section suggest that the use of SNA in the enforcement of criminal organizations may become a reality.

Future Issues and Trends

As already mentioned, the application of SNA to organized criminal group is a very recent phenomenon. Therefore, it is likely that further developments will come in the next years. Some trends may be reasonably anticipated.

The number of publications will increase. The rate of publications applying SNA to criminal organizations has been significantly growing, and this trend should continue, if not increase, in the coming years. Hopefully, this may improve the overall quality and standards in this specific field. It will also prevent and eventually remove the inappropriate exploitation of the word network analysis in contributions which do not actually make any use of SNA methods and concepts. Eventually, this may provide SNA applications wider academic acceptance in this specific field and reduce skepticism about the network methods.

Further, the increase in the number of published works will likely provide more empirical studies of specific organized crime groups (micro approach) and/or markets (macro approach). At the moment, the promising results produced by the first papers need further support or verification with other groups and countries. The replication of methodologies and analyses will play an important part until a critical mass of criminal networks and datasets will be analyzed and available.

The methodology will improve. As already mentioned above, so far most SNA studies on organized crime have applied relatively basic measures and concepts. Although this is not necessarily a problem, since also very basic analyses have provided valuable results, the forthcoming years will likely bring new, more complex network methods to be applied also in the study of organized crime groups. Inevitably, the main challenge in this direction will be data availability and, ultimately, the improvement of data collection procedures by law enforcement. The availability of new and better data sources will likely allow the application of advanced SNA techniques such as dynamic network analysis, multiplexity, and ERGMs. These will provide new avenues for analyzing the structure of complex networks and their evolution through time.

In addition to more complex network techniques, future studies will likely mix network methods with other methods, both quantitative and qualitative. Network analysis can be complemented with other quantitative techniques

such as spatial analysis and hotspot analysis. Furthermore, the inherently quantitative approach implied in SNA may provide the best results when balanced by qualitative analyses. Studies within the micro approach may take advantage of a detailed analysis of the case studies, including interviews with law enforcement to verify possible biases, interviews with network participants, or content analysis of judicial sources and transcripts (Campana 2011). As anticipated by Morselli, “content analysis is indeed the next step toward enriching the various analyses conducted throughout this book” (Morselli 2009, 164).

SNA and the enforcement of organized crime groups. Notwithstanding the issues highlighted in the previous section, future improvements of network applications to organized crime groups may provide interesting developments for the operational/enforcement point of view. In particular, SNA may be applied not only within investigations, but across different investigations. This contribution may have a geographical and a chronological dimension.

From the geographical perspective, criminal organizations are frequently operating in different areas which may not always be within the jurisdiction of the same law enforcement agency. This is particularly true for large criminal organizations such as mafia-type associations which may commit crimes in different countries. Inevitably such extended criminal activity is likely to involve a number of different police forces and prosecutors’ offices, with increasing problems in coordination and information sharing. Different cases and files are going to be opened. It may happen that a marginal subject in the investigation led by District A shows up to be a high-ranking criminal in the investigation conducted by District B. The application of SNA methods across investigations may provide critical information which would not otherwise be available to the single investigators/prosecutors. Through network analysis it may be possible to merge investigative data from a number of cases and provide an overall picture which may be otherwise too complex or difficult to achieve. The literature has already provided some examples

of SNA studies merging data from different investigations (Morselli 2009, Chaps. 8 and 9). Results demonstrated that network structure and positioning may significantly change when different cases are merged.

From the chronological perspective, some criminal investigations are capable to survive prosecution and convictions. Frequently, family ties and kinship provide important connections among the members, and this allows such criminal groups to last over years. Mafia-type organizations (e.g., the five families in New York) may last for years, notwithstanding constant law enforcement monitoring and repeated prosecutions. In these cases, it may be important to achieve a very good knowledge of the history of a criminal group. Surely, prosecutors and policemen achieve excellent knowledge of the most dangerous groups in their jurisdictions, but in some cases this important knowledge asset may be dispersed due to reorganization, retirement, transfer, or resignation. The application of network analysis may provide important strategic information about the development and structure of a single criminal organization through time. Furthermore, SNA may allow to assess the changes and reactions of criminal organizations as a reaction to law enforcement. In this perspective, Morselli and Petit showed that systematic drug seizures by law enforcement had a significant impact on a criminal organization, although the group was to continue its operations for months (Morselli and Petit 2007). Further studies may analyze the impact of arrests (either random or targeted) on criminal organizations, applying dynamic network analysis techniques to organized crime groups. Ultimately, this may allow law enforcement agencies to estimate the impact of different strategies and investigations on the evolution of the structure of large and continuing criminal organizations. SNA methods may represent a valuable quantitative benchmark in this evaluation.

In conclusion, organized crime is one of the last criminological areas where SNA methods have been applied. If Carrington has argued that “the use of social network analysis in criminology is in its infancy,” this is even more relevant

for the use in the study of organized criminal groups. In the last 10 years, important steps have been made. The progress are encouraging and it is likely that, through the trial and error which is one of the basic methods of science, SNA will play an important part in the study of organized crime in the next years.

Related Entries

- ▶ [Gangs and Social Networks](#)
- ▶ [Neighborhood Effects and Social Networks](#)
- ▶ [Network Analysis in Criminology](#)
- ▶ [Social Network Analysis and the Measurement of Neighborhoods](#)
- ▶ [Social Network Analysis of Urban Street Gangs](#)
- ▶ [Spatial Models and Network Analysis](#)

Recommended Reading and References

- Albini JL (1971) *The American mafia; Genesis of a legend*. Sociology series. Appleton, New York
- Berlusconi G (2013) Do all the pieces matter? Assessing the reliability of law enforcement data sources for the network analysis of wire taps. *Global Crime* 14(1): 61–81.
- Boivin R (2011) Drug trafficking networks in the world economy. In: Montréal. <http://www.erdr.org/textes/boivin.pdf>
- Bouchard M, Ouellet F (2011) Is small beautiful? The link between risks and size in illegal drug markets. *Glob Crime* 12(1):70–86
- Bouchard M, Descormiers K, Gallupe O, Wang W (2010) Panel: analyzing illegal drug supply chains as 2-mode social networks. Paper presented at the crime and social institutions, 62nd annual meeting of the American society of criminology, San Francisco, 17 November
- Bright DA, Hughes CE, Chalmers J (2012) Illuminating dark networks: a social network analysis of an Australian drug trafficking syndicate. *Crime Law Soc Change* 57(2):151–176. doi:10.1007/s10611-011-9336-z
- Calderoni F (2012) The structure of drug trafficking mafias: The Ndrangheta and Cocaine. *Crime Law Soc Change* 58(3):321–349.
- Campana P (2011) Eavesdropping on the mob: the functional diversification of mafia activities across territories. *Eur J Criminol* 8(3):213–228
- Carrington PJ (2011) Crime and social network analysis. In: Scott JP, Carrington PJ (eds) *The SAGE handbook of social network analysis*. SAGE Publications, London, pp 236–255

- Davis RH (1981) Social network analysis: an aid in conspiracy investigations. *FBI Law Enforce Bull* 50(12):11–19
- Harper WR, Harris DH (1975) The application of link analysis to police intelligence. *Hum Factors* 17(2):157–164
- Heber A (2009) The networks of drug offenders. *Trends Organ Crime* 12(1):1–20
- Ianni FAJ (1973) Ethnic succession in organized crime: summary report. U.S. Government Printing Office, Washington
- Ianni FAJ, Reuss-Ianni E (1990) Network analysis. In: Andrews PP, Peterson MB (eds) *Criminal intelligence analysis*. Palmer Enterprises, Loomis, pp 67–84
- Lupsha P (1980) Steps toward a strategic analysis of organized crime. *The Police Chief* 47:36–38
- Lupsha P (1983) Networks Vs. networking: analysis of an organized crime group. In: Waldo GP (ed) *Career criminals*. Sage, Beverly Hills
- Malm AE, Bichler G (2011) Networks of collaborating criminals: assessing the structural vulnerability of drug markets. *J Res Crime Delinq* 48:271–297
- Malm AE, Kinney JB, Pollard NR (2008) Social network and distance correlates of criminal associates involved in illicit drug production. *Secur J* 21(1–2):77–94
- Malm AE, Bichler G, Van De Walle S (2009) Comparing the ties that bind criminal networks: is blood thicker than water? *Secur J* 23(1):52–74
- Malm AE, Bichler G, Nash R (2011) Co-offending between criminal enterprise groups. *Glob Crime* 12(2):112–128. doi:10.1080/17440572.2011.567832
- Mastrobuoni G, Patacchini E (2010) Understanding organized crime networks: evidence based on federal bureau of narcotics secret files on American mafia. Carlo Alberto notebooks 152. Collegio Carlo Alberto, Moncalieri
- Morselli C (2005) *Contacts, opportunities, and criminal enterprise*. University of Toronto Press, Toronto
- Morselli C (2009) *Inside criminal networks*. Studies of organized crime. Springer, New York
- Morselli C, Petit K (2007) Law-enforcement disruption of a drug importation network. *Glob Crime* 8(2):109–130
- Morselli C, Roy J (2008) Brokerage qualifications in ring-ing operations. *Criminology* 46(1):71–98
- Natarajan M (2000) Understanding the structure of a drug trafficking organization: a conversational analysis. In: Mangai N, Mike H (eds) *Illegal drug markets: from research to prevention policy*, vol 11, Crime prevention studies. Criminal Justice Press/Willow Tree Press, Monsey, pp 273–298
- Natarajan M (2006) Understanding the structure of a large heroin distribution network: a quantitative analysis of qualitative data. *J Quant Criminol* 22(2):171–192
- Papachristos AV, Smith CM (2011) The small world of Al Capone: the embedded nature of criminal and legitimate social networks. In: Third annual illicit networks workshop, Montréal. http://www.erdr.org/textes/papachristos_smith.pdf
- Sparrow MK (1991) Network vulnerabilities and strategic intelligence in law enforcement. *Int J Intell Counter Intel* 5(3):255–274
- Varese F (2006) The structure of a criminal network examined: the Russian mafia in Rome. Oxford Legal Studies Research Paper No. 21/2006

Social Network Analysis of Urban Street Gangs

Jenny C. Piquette¹, Chris M. Smith² and Andrew V. Papachristos³

¹Department of Sociology, The Ohio State University, Columbus, USA

²Department of Sociology, University of Massachusetts Amherst, Amherst, USA

³Department of Sociology, Yale University, New Haven, CT, USA

Overview

This entry synthesizes the current state of network-related research on street gangs around the themes of network structure, network action, and network location. At their core, street gangs are social networks created by the coming together, socializing, and interacting of individuals in particular times and in particular places. The employment of social network analysis has the potential to examine patterns of interaction among gang members and gangs, illuminate structural variation across gangs, and measure the influence of gang networks on individual action. This entry includes an overview of social network analysis, suggestions and directions for future gang network research, a discussion of limitations, and a vision for how a social network approach to gangs might inform theory, research, and practice.

Introduction

Over the past decade, the field of criminology has increasingly employed both the theoretical and

methodological tools of social network analysis (McGloin and Kirk 2010; Papachristos 2011). Although criminology lags behind other disciplines in this network turn, criminologists have applied social network analysis to investigate organized crime syndicates, narcotics trafficking, terrorist organizations, white-collar conspiracies, and adolescent delinquency. But perhaps one of the areas of criminological inquiry most in need of a network-related approach is the study of street gangs.

Street gangs are, if anything, social networks created by the coming together, socializing, and interacting of individuals in particular times and in particular places. Though scholars still heavily debate the definition of what constitutes a gang, virtually all gangs have some element of a group and it is this groupness that is at the heart of this entry. Street gangs are groups of individuals engaging in activities that are bolstered by group processes, culture, and structures. Social network analysis studies street gangs *empirically* to examine precise patterns of interaction among individuals and collectivities (see Papachristos 2006). *Theoretically* gangs operate as unique sociological entities with various structural and interactional manifestations. Research gathers, compares, and analyzes dozens of network properties at the individual, group, and system level, which brings fresh perspectives to old debates. Contemporary social network approaches have the potential to push past criminology's definitional bulwarks and typologies and reveal how real-life behavior of individuals and groups might inform theory, research, and practice.

This entry synthesizes the current state of network-related research on street gangs. Rather than simply plod forward in the standard literature review format of "this article says this" and "that article says that," the aim will be to review research in a way that provides a road map for future lines of inquiry. Specifically, the focus will be on measuring and theorizing street gangs from multiple levels of analysis in order to shed light on a broader range of empirical gang behaviors, structures, and processes. To wit, this entry is structured around three themes: (1) the structure of gang networks, (2) the actions of gang

networks, and (3) the location of gang networks. In the following sections particular attention will be paid to the group nature of gangs as well as the locations and actions of individuals in gangs. This entry concludes with a brief discussion of the challenges to and limitations of social network analysis and gang research. To begin, the next section provides a brief discussion of social network analysis more generally and proposes its application to street gangs.

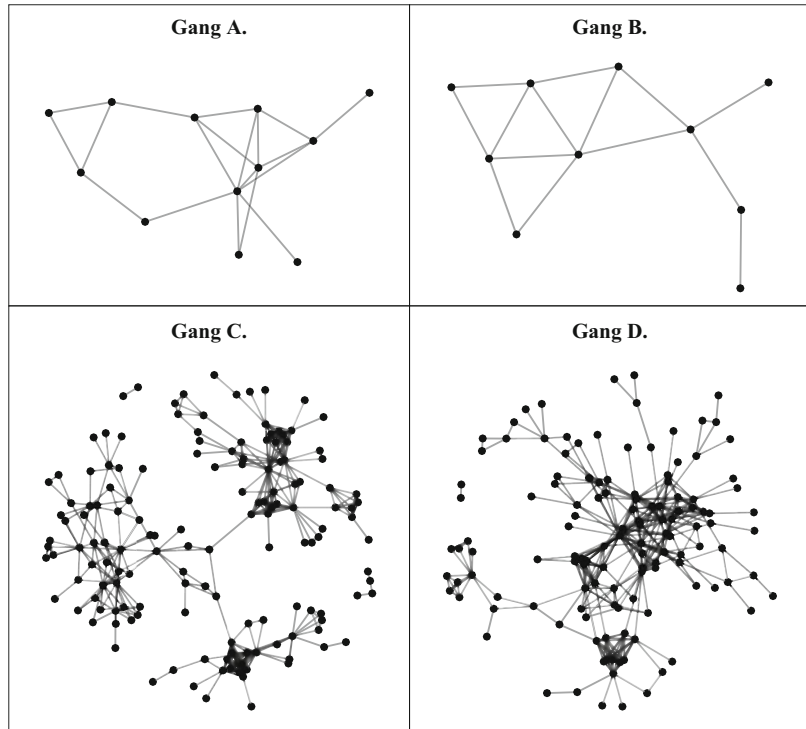
Social Network Analysis and Applications to Street Gangs

Social network analysis (SNA) refers to the study of social relationships among sets of actors and, more importantly, the analysis of how the patterns of relationships affect actors' behaviors (Wasserman and Faust 1994). Unlike traditional regression analyses in which effects are net of other variables (Emirbayer 1997), SNA assumes *interdependence* rather than *independence*. Indeed, it is such interdependence that is of analytic interest. Most network studies attempt to (a) analyze and describe the observed structure of relationships among a set of actors and/or (b) investigate how sets of relationships affect actors' behaviors, actions, opinions, and attitudes. Empirical research has demonstrated the importance of networks on a variety of phenomena such as getting a job, income attainment, political decision making, the diffusion of ideas, and even political revolutions.

Formally, social network analysis relies on graph theory: mathematical models that capture the pairwise relationships among a specific class of objects (Wasserman and Faust 1994). In graph theoretical terms, a network consists of a set of vertices (V) that represent a bounded set of actors or units and a set of edges (E) that define the relationship among the vertices. Relationships are commonly represented in a square matrix $X = x_{ij}$ where x_{ij} is the strength of relationship between unit i and unit j (Wasserman and Faust 1994). Edges can be binary indicating the presence of a relationship. Alternatively, edges can be valued measuring the intensity of a relationship, the

Social Network Analysis of Urban Street Gangs,

Fig. 1 Four African American gang networks from Chicago’s west side



level of interaction, or any other theoretically sound quantity of a relationship. Ties can also have direction, such that one node sends a tie and another receives it. Sociograms or graphs display networks visually in which nodes depict the vertices and lines depict the edges (see, e.g., Fig. 1).

Conceptually, gangs are clearly social networks. Gangs consist of actors who are tied to each other through group participation and identity, as well as neighborhood, familial, and peer relationships. Furthermore, gangs collectively act in particular moments and places, whether simply hanging out or engaging in gang warfare. Of particular interest for researchers is a gang’s ability to activate ties in order to accomplish certain tasks and the ensuing consequences of those tasks. Figure 1 illustrates this point by mapping the network structures of four African American gangs in Chicago using police observational data. Each node represents a unique individual and each tie represents a non-arrest observation made by the police – sometimes called “field observations” or “field contacts” (see

Papachristos et al. 2012a). Each tie is a police observation of two individuals hanging out on a street corner, which is perhaps one of the most fundamental gang activities. By aggregating such information over time and within multiple locations, a network structure emerges representing the patterns of direct and indirect associations among gang members as well as their non-gang companions.

All four of the gangs in Fig. 1 are from the west side of Chicago and are part of larger gang nations that date back to the 1960s. By virtually any survey or qualitative account, all four of these gangs should have the following: a hierarchical structure, formalized sets of rules and regulations, a division of labor, and intricate drug-dealing portfolios. Yet, the patterns of association shown in Fig. 1 reveal that some of the networks deviate from such accounts. Gangs A and B, for example, are smaller in size and less dense than gangs C and D. Moreover, the network structures of gangs A and B look more like ordinary friendship networks and are decidedly *not* hierarchical. In gang B, triangles of individuals overlap to make up the

larger networks; this pattern suggests that the network is comprised of small cliques with overlapping membership as opposed to some formalized division of labor. In contrast, gangs C and D have a starburst type of structure with dense pockets of individuals linked together by one or two ties. These gangs are more centralized, a trait that is especially evident in gang C with a single individual joining three different parts of the graph.

The analysis of gang structures such as those in Fig. 1 can augment more traditional analyses of gangs by comparing structural traits of networks across time and space. A whole range of network measures and statistics – such as density, clustering, and degree distribution (Wasserman and Faust 1994) – can be compared and analyzed across samples of gangs to detect points of variation and similarity. Perhaps more importantly, SNA can examine the structural properties that distinguish a gang from other types of social groups.

A networked approach to gangs requires two related and nested units of analysis: (1) gangs as collectives and (2) gangs as a collection of individual gang members. The questions gang researchers ask rely on the unit of analysis. At the group level, researchers examine organizational structures, gang interactions within larger social spaces, and intragroup relations affecting larger social processes. Individual-level hypotheses center on relations between gang members and gang influence on individual behavior. Research agendas might include (1) how friendship networks within a gang mediate or facilitate violence, (2) the types of relationships from which gang leadership or influence derive, (3) patterns of social status and power among gang members, or (4) the ways in which gang members integrate into larger noncriminal networks (see Papachristos 2006). The remainder of this entry grapples with some of these hypotheses, which are organized around the themes of network structure, network action, and network location.

Network Structures of Street Gangs

Though not specified in formal network language, researchers' conceptualization of street

gangs has relied on network concepts and ideas for decades. At the most basic level, gangs can be conceptualized as groups of delinquent individuals who socialize with each other. Rather than a random collection of strangers who happen to share social and geographic space, gangs have a collective identity, a culture, and sense of togetherness. As such, criminologists strive to understand group formation, cultural precepts, or collective processes. Prior research has provided tremendous insight into street gangs' culture, experiences, and understandings of their group, but has fared less well in explaining the structural variation and scope of street gangs. SNA, however, provides theoretical and methodological tools to evaluate gangs' organizational structure and participation in those structures.

Organizational Structure

Descriptions of the gangs as organizations range from informal friendships and associations (Fleisher 1999) to formalized corporate bureaucracies (Venkatesh and Levitt 2000). Multiple typologies have emerged attempting to classify gangs according to various organizational or group characteristics. For example, research has debated whether leadership and influence within a gang are vertical (bureaucratic) or horizontal (influence based) (Sanchez-Jankowski 1991). Other typologies differentiate gangs based on certain organizational characteristics such as membership size, meeting regularity, formalized rules and regulations, and gang activities (Bouchard and Spindler 2010; Pyrooz et al. 2012). Gangs are also frequently differentiated by the kinds of criminal behaviors in which they engage – i.e., violent gangs, income generating gangs, and delinquent gangs (Starbuck et al. 2001). Thus far, the majority of the research on gang structure has been largely descriptive and, therefore, limited from making topological comparisons.

SNA can advance this line of research by moving away from descriptive organizational characteristics to formal measures of group structures. This transition allows for testing organizational types, examining structural changes over time, as well as the ability to look at patterns of behavior, influence, and interaction. For

example, network research on formal organizations suggests that informal relationships are often more significant in organizational behavior than formal hierarchical positions. Asking whether or not a gang has regular meetings overlooks the informal associations that may be at the heart of gang life. SNA has challenged conventional understandings of gang membership by showing how ties among individuals extend beyond the in-group and has also revealed more intricate relations within the larger gang structure (McGloin 2005). Frequently gang members are not friends, do not know one another, or have animosity toward one another. From a networked perspective, one builds the organizational structure from the ground up by tracing lines of influence and interaction among members.

A network approach to gangs has the potential to illuminate structural variation, model organizational changes over time, and examine organizational threat or stability. Figure 1, for example, presents structures representing actual lines of interaction among members, as opposed to a theoretical structure that may or may not fit the actual structure of the group. Formal analysis of networks sheds light on current typologies and develops new understandings of gang structures. One potential avenue of inquiry would compare the patterns of interaction within specified groups to group structures as reported by law enforcement or gang members. In many non-gang settings, often the informal networks of friendship have a stronger influence on behavior (see Haynie 2001). Another avenue of inquiry would compare structures at the group level in order to examine if particular network forms are more or less conducive to certain behaviors such as violence or drug dealing.

Structural Position

Not all gang members are alike. For decades, gang research has made important distinctions between the levels at which members participate in gang life and gang culture. Terms such as wannabes, hanger-ons, and associates typically refer to those individuals only causally involved in a gang, whereas terms such as hard core, down, or lifer are reserved for those more committed to gang life (Klein and Maxson 2006). Scholars

often make the distinction between core and periphery gang members where the former refers to those who are active in the gang and the latter refers to members not central to the gang structure. The basic presupposition is that those who are more central to the gang structure are participating more fully in gang activities and those who are on the outskirts participate less.

Scholars also recognize the naïveté of such a core-periphery distinction. Gang participation is more amorphous than most typologies allow. Although popular media often depict gangs as secret societies with blood rituals and initiations, the reality is that gang membership is fluid (Fleisher 1999). Furthermore, membership is often unofficial; it is just as likely that membership is based on familial relations than a formal initiation. Some fully indoctrinated members may actually exert very little influence, while others who are not formal members may exert tremendous influence. Small cliques form within the group, which can at times divide the group into subdivisions, each with its own level of participation (Fleisher 1999; McGloin 2005; Papachristos 2006).

SNA can move research on gang participation forward by considering how positions within networks may or may not relate to levels of participation or vice versa. By empirically returning to basic gang networks (such as those in Fig. 1), analysts can use various network measures to find structural similarities or differences within and across groups. For example, one promising idea might be to analyze a sample of gang structures and test for structural equivalence in order to identify unique positions or roles in gangs. Another potential line of inquiry might locate sources of influence within gang structures – i.e., locating individuals, such as brokers or gatekeepers, who might be significant not because of formal titles or positions but because of their influence on member behavior.

Network Action of Street Gangs

Like other types of social groups, gangs can pursue collaborative goals, form alliances, and

engage in collective action. Perhaps one of the most important areas of research on SNA and gangs investigates how gang network structures affect collective and individual action.

Gang Landscapes

SNA has the capacity to survey *gang landscapes*: the general patterns of relationships among a population or system of gangs (Kennedy et al. 1997). Social mapping of the collective relationships within a gang or across gangs offers insight into the organizational and territorial niches in which gangs exist. A gang landscape can map the alliances and rivalries between groups and individuals across a geographic space. Gang landscapes have the potential to situate specific gang activities and exchanges (such as violence and drug dealing) in the context of a total networked system. In other words a structural analysis illuminates gang actions and behaviors. For example, Descormiers and Morselli (2011) analyze how gang-level attributes (such as gang size, age of members, and ethnicity) and ecological factors (such as gang turf and underground economies) predict maps of gang alliances and conflicts in Montreal. In a slightly different approach, Papachristos (2009) maps the networks of violent exchanges between Chicago gangs over time and finds that prior relationships and the position of a gang in conflict networks are strong predictors of retaliation. This latter example demonstrates how patterns of the gang landscape provide insight into a social process of retaliation.

Group Processes

Gangs' group-level actions include protection, socializing members, and supplying resources. These behaviors imply a number of social processes that are negotiated within the gang. Prior research has revealed the potential mechanisms within gangs that contribute to their crime and delinquency – e.g., loyalty, status, and protection (e.g., Short and Strodtbeck 1963). Protection is of particular importance because survey research has consistently found that protection is one of the most commonly cited reasons for joining youth gangs (Klein and Maxson 2006). In the

quest for mutual protection, gangs face a basic collective action problem: they must motivate members to actually come to each other's mutual aid in the face of a threat.

Network structures can facilitate or impede acts of mutual protection. In particular, the idea of *cohesion*, often measured as the density of ties within a network (Wasserman and Faust 1994), offers a fruitful direction for empirical analysis. A second approach to examining protection is by calculating the network density of gangs and correlating this with acts of violence – especially retaliation. For example, Papachristos (2009) found that gangs are more likely to reciprocate an act of violence when their actions are visible to other gangs. In other words, the collective processes involved in retaliation are more likely to occur when retaliation becomes a signal to other groups that a gang is capable of vengeance and protection. In addition to reinforcing the *esprit de corps* of the group, Papachristos (2009) finds such acts of retaliation can actually flame the spread of violence and create enduring networks of gang conflict that persist over time.

Research on gangs' group processes and the ensuing consequences has been virtually untouched. The launch pad for such inquiry should focus on how specific hypothesized processes might manifest themselves structurally or on how structures affect the hypothesized processes. Just as understanding the sociodemographic characteristics of neighborhoods allow one to understand how social processes ecologically unfold, understanding the networks of street gangs better unpacks the social context in which group processes unfold.

Group Influence on Individual Action

One of the fundamental questions in gang research is the extent to which gangs (as groups) actually facilitate individuals' delinquent or criminal behavior. This is in contrast to the selection effect perspective, which assumes like-minded and previously delinquent individuals come together to form a group (Bendixen et al. 2006; Thornberry et al. 2003). Mounting evidence from longitudinal studies supports a facilitation effect: individual delinquency is

amplified in the gang context even when controlling for factors such as individual self-control, prior delinquency, and sociodemographic characteristics (Krohn et al. 2011; Thornberry et al. 2003). Individual delinquency, even among those with higher baseline delinquency, increases during periods of gang membership. However, what is still unclear is exactly how the gang facilitates crime and delinquency for the individual.

A growing body of network research shows that the structure of adolescent friendship networks is associated with self-reported delinquency, which can range from statutory crimes like underage drinking to more serious crimes like theft (Haynie 2001, 2002; McGloin 2009). In particular, this research finds considerable evidence for peer influence on delinquency: one's own delinquency is related to the attitudes, behaviors, and delinquency of one's peers. At the same time, only a handful of studies look at co-offending networks involved in serious crimes. Given the observed facilitation effect of gangs and the simple fact that gang associations appear more group-based than friendship networks, it is reasonable to hypothesize that the network effect of gangs on individual action should be larger than effects of friendship networks or co-offending networks. Furthermore, certain network configurations are more apt than others to facilitate imitation of behavior, to provide opportunities for deviance, and to exert stronger influence on its members. As such, we hypothesize that (a) the gang context amplifies group-level processes that foster delinquency and (b) gangs exhibit additional processes not found in delinquent peer groups.

Two recent studies by Papachristos and colleagues provide evidence for the group facilitation effect of gangs. In a study of gang members and their associates in Boston's Cape Verdean community, Papachristos et al. (2012a) examine a social network of more than 700 individuals, approximately 30 % of whom are gang members. Their study finds that individuals in the network are more likely to be gunshot victims when (a) they are gang members, (b) their immediate social network is saturated with gang members, and (c) their immediate social networks contain other gunshot victims. In other words, network

exposure to gang members and gunshot victims increases one's own risk of being a victim. The second study surveys 150 active gun offenders in Chicago. Papachristos et al. (2012b) find that the network structure of individuals affects both their perceptions of the law and their subsequent law-violating behavior. In particular, individuals are more cynical of the law and subsequently more likely to engage in deviant behavior when their social networks are saturated with gang members. These two studies offer only a small glimpse into the ways network structure might influence individual actions and/or individual-level consequences. Future research should consider a wider range of group types (such as comparing gang members vs. non-gang members or members of other types of delinquent groups), different types of group processes (imitation, peer influence, and homophily), as well as various types of individual-level outcomes (beliefs, decisions, and behaviors).

Violence Prevention

From a social problem perspective, one of the most critical gang activities requiring attention is the incredibly high rate of lethal and nonlethal violence. A small but growing body of research has taken gang landscape mapping quite seriously as both a means of gang violence prevention and a research area. As part of a crime reduction initiative, Kennedy et al. (1997) mapped out gang rivalries and alliances in Boston. Violence intervention and prevention efforts relied on these network maps to direct attention toward gangs that were actively involved in shootings and target resources scarce to those gangs. This targeted group-based intervention strategy (of which gang network mapping was a key part) produced significant reductions in youth homicide and nonfatal shootings (Braga et al. 2001). Similar approaches evaluate the efficacy of group-based and outreach worker strategies in other US cities (Skogan et al. 2009; Tita et al. 2003).

Network Locations of Street Gangs

Gangs are decidedly local phenomena with strong connections to the character and culture

of urban neighborhoods. In fact, part of sociology's continued fascination with gangs stems from their relationship to neighborhood social organization. Gangs are seen as de facto social institutions or collectivities that meet the basic needs of community and youth in otherwise disadvantaged communities. More than that, however, the neighborhood becomes a defining element of gang identity and group membership (Garot 2007; Grannis 2009). This association is so strong that many gang names incorporate their neighborhood of origins: the 35th Street Kings, Columbia Point Dogs, Westville Saints, and so on. Indeed, often the simple question "Where you from?" is a shorthand way for identifying gang members (Garot 2007). In a way, the gang itself becomes a reflection of many of the conflicts, cultures, and idiosyncrasies of the larger neighborhood.

Prior research has shown that space matters in shaping gang behaviors in at least two ways. First, and foremost, gang violence is commonly associated with the defense of gang turf or territory. Given the aforementioned attachment of gangs to urban space, threats to turf are considered threats to home, members' safety, and the identity of the group. Violence stemming from the protection of turf becomes as much a symbolic measure signaling the solidarity and reputation of the gang and community as well as an instrumental gesture in protecting a piece of land over which a gang claims ownership. Secondly and perhaps related to such matters, previous research also suggests that gang violence is spatially contagious; gang violence in a given neighborhood affects levels of violence in adjacent neighborhoods. In a sense, the group and dynamic nature of gang violence spills over into adjacent communities. Though such spatial effects are associated with crime and violence more generally, a recent study of gangs in Los Angeles suggests that the spatial effects of gang violence extend beyond immediately adjacent geographic areas to affect more geographically distant neighborhoods (Tita and Radil 2011; Tita et al. 2003).

To date, with few exceptions, such spatial analyses of gangs rely mainly on regression

models of aggregate crime measures – i.e., the extent to which crime rates in a focal neighborhood are associated with rates in surrounding communities. Yet, crime rates themselves do not move across gang turf or other geopolitical boundaries: gangs and gang members do. Nearly all empirical studies on the spatial nature of crime theoretically acknowledge that the social networks of criminal activity – such as drug dealing, prostitution, and gangs – drive the observed contagion of crime across neighborhood boundaries (e.g., Cohen and Tita 1999; Morenoff et al. 2001). Yet only a handful of studies actually incorporate social networks and geographic space into the same analysis, and even fewer studies have applied both to the study of gangs.

The inherent networked nature of gangs and their strong links to urban space propose an excellent area of study for understanding the convergence of social and geographic space. This line of inquiry would be relevant not only to gang research but also to social science more broadly. In fact a recent issue of *Social Networks* (Adams et al. 2012) was dedicated entirely to the convergence between spatial and network thinking. Statistical models of geographic and network autocorrelation share many similarities and, in their most basic principles, have a similar functional form. The gang, both theoretically and empirically, affords an important instance in which to understand the convergence of both geographic and social space. For example, combining social networks of gang relationships with geographic models of diffusion could shed light on whether the clustering of gang violence in particular neighborhoods is related to ecological conditions, group processes (such as reciprocity), or some combination therein. Likewise, analyses such as these might advance our understanding of whether institutional memory captured in social networks might trump geographic barriers. Alternatively, spatial network research could also provide valuable insight into the ways that group processes and geography may impede or promote behavior – the relevance of such findings extends well beyond the world of the street gang.

Concluding Remarks

To review, SNA affords many unique opportunities to consider the group nature of contemporary street gangs. By considering gangs as networks of individuals or by mapping systems of gangs, researchers can more accurately measure structures and processes that make the gang a unique sociological entity. This entry has outlined some promising avenues for inquiry and, no doubt, many more exist. In advancing this network agenda for gang research, researchers should keep two things in mind. First, at the core of gang research is the inherent “groupness” of the gang. The application of SNA to gang research should proceed in a way that exploits new methodologies and data to understand this aspect of the gang. Second, the potential influence between SNA and gang research is bidirectional. This means that the application of SNA to gangs should not simply be a “cut-and-paste” of existing methodologies and ideas. Rather, the diligent application of SNA to gang research should provide insight into a range of small group behavior relevant to social science more generally.

Though enthusiastic about the direction of SNA and gang research, this entry recognizes that the methodology is not without challenges or limitations, two of which are mentioned here. First and foremost, boundary specification issues haunt all SNA research (Laumann et al. 1989) and pose even greater limitations for theoretical and empirical considerations of gangs. Boundary specification refers to the problem in SNA of identifying where the network begins and where it ends. Boundary specification can be driven by data considerations (who is in a dataset) or by theoretical considerations (who is part of the group). As discussed above, SNA research on gangs has already demonstrated that boundaries of gangs – defined by gang membership – are rather amorphous (Fleisher 1999; McGloin 2005). This suggests that finding the “true” group may prove elusive. More precisely, the shape of the network will be contingent upon where analysts draw the boundary of the group – whether, for instance, to include only “gang members” or to also include their

associates. Figure 1, for instance, stopped at 2° of separation from known gang members (i.e., those individuals who were friends of friends of known gang members). In short, SNA provides a new way to discuss group membership and structure, but ultimately boundary specification becomes immensely important for theory and empirical research.

Second, future research should explore the limits of existing data as well as consider how to improve future data collection efforts. SNA requires relational data: a data source that contains some reference to the relationship among units. In some instances, this means revisiting old data sources with a new theoretical lens. For example, the study of co-offending networks (Schaefer 2012) and the study of gang homicide networks (Papachristos 2009) relied on official police data common in criminological research. But it required these authors to understand a relational aspect of the data not previously considered: how events link individuals and groups. The network literature refers to this as two-mode data (see Wasserman and Faust 1994). Similarly, ethnographic data often provides extremely rich information on the social relationships of its subjects. Fleisher (1999), for example, coded his field notes for specific instances of social ties to help uncover the patterns of the group – a time-tested anthropological technique for understanding group membership. Papachristos (2006) has even gone back and coded published ethnographies to a similar end. Moving forward, gang researchers should include network-related questions in future survey collection efforts. Indeed, much of the growth of SNA in criminology can be attributed to the network focus of a handful of surveys of school youth. Subsequent surveys of gang members or law enforcement could advance this area of research considerably by adding even a single network question to their survey instruments.

Related Entries

- ▶ [Co-Offending](#)
- ▶ [Gangs and Social Networks](#)

- ▶ [Neighborhood Effects and Social Networks](#)
- ▶ [Network Analysis in Criminology](#)
- ▶ [Race, Ethnicity, and Youth Gangs](#)
- ▶ [Risk Factors for Gang Membership](#)
- ▶ [Social Network Analysis of Organized Criminal Groups](#)

Recommended Reading and References

- Adams J, Faust K, Lovasi GS (eds) (2012) Capturing context: integrating spatial and social network analysis [Special issue]. *Social Networks* 34(1):1–158
- Bendixen M, Endresen IM, Olweus D (2006) Joining and leaving gangs: selection and facilitation effects on self-reported antisocial behavior in early adolescence. *Eur J Criminol* 3(1):85–114
- Bouchard M, Spindler A (2010) Groups, gangs, and delinquency: does organization matter? *J Crim Justice* 38(5):921–933
- Braga AA, Kennedy DM, Piehl EJ, Morrison A (2001) Problem-oriented policing: an evaluation of Boston's operation ceasefire. *J Res Crime Delinq* 38(3):195–225
- Cohen J, Tita G (1999) Diffusion in homicide: exploring a general method for detecting spatial diffusion. *J Quant Criminol* 15(4):451–493
- Descormiers K, Morselli C (2011) Alliances, conflicts, and contradictions in Montreal's street gang landscape. *Int Crim Justice Rev* 21(3):297–314
- Emirbayer M (1997) Manifesto for a relational sociology. *Am J Sociol* 103(2):281–317
- Fleisher M (1999) Dead end kids: gang girls and the boys they know. University of Wisconsin Press, Madison
- Garot R (2007) 'Where you from!': gang identity as performance. *J Contemp Ethnogr* 36(1):50–84
- Grannis R (2009) From the ground up: translating geography into community through neighbor networks. Princeton University Press, Princeton
- Haynie D (2001) Delinquent peers revisited: does network structure matter? *Am J Sociol* 106(4):1013–1057
- Haynie D (2002) Friendship networks and delinquency: the relative nature of peer delinquency. *J Quant Criminol* 18(2):99–134
- Kennedy DM, Braga A, Piehl A (1997) The (un)known universe: mapping gangs and gang violence in Boston. In: Weisburd D, McEwen T (eds) *Crime mapping and crime prevention*. Criminal Justice Press, New York, pp 219–262
- Klein MW, Maxson CL (2006) *Street gang patterns and policies*. Oxford University Press, New York
- Krohn M, Ward J, Thornberry TP, Lizotte A, Chu R (2011) The cascading effects of adolescent gang involvement across the life course. *Criminology* 49(4):991–1028
- Laumann EO, Marsden PV, Prensky D (1989) The boundary specification problem. In: Freeman L, White D, Romney A (eds) *Research methods in social network analysis*. George Mason University Press, Fairfax, pp 61–87
- McGloin JM (2005) Policy and intervention considerations of network analysis of street gangs. *Criminol Pub Policy* 4(3):607–635
- McGloin JM (2009) Delinquency balance: revisiting peer influence. *Criminology* 47(2):437–477
- McGloin JM, Kirk D (2010) Social network analysis. In: Piquero A, Weisburd D (eds) *Handbook of quantitative criminology*. Springer, New York, pp 209–224
- Morenoff JD, Sampson RJ, Raudenbush SW (2001) Neighborhood inequality, collective efficacy, and the spatial dynamics of homicide. *Criminology* 39(3):517–560
- Papachristos AV (2006) Social network analysis and gang research: theory and method. In: Short JF, Hughes LA (eds) *Studying youth gangs*. AltaMira Press, Lanham, pp 99–116
- Papachristos AV (2009) Murder by structure: dominance relations and the social structure of gang homicide. *Am J Sociol* 115(1):74–128
- Papachristos AV (2011) The coming of a networked criminology? Using social network analysis in the study of crime and deviance. In: MacDonald J (ed) *Measuring crime and criminality: advances in criminological theory*, vol 17. Transaction Publishers, New Brunswick, pp 101–140
- Papachristos AV, Braga AA, Hureau DM (2012a) Social networks and the risk of gunshot injury. *J Urban Health* 89(6):992–1003
- Papachristos AV, Meares TL, Fagan J (2012b) Why do criminals obey the law? The influence of legitimacy and social networks on active gun offenders. *J Crim Law Criminol* 102(2):101–144
- Pyrooz DC, Fox AM, Katz CM, Decker SH (2012) Gang organization, offending, and victimization: a cross-national analysis. In: Esbensen FA, Maxson CL (eds) *Youth gangs in international perspective: tales from the Eurogang program of research*. Springer, New York, pp 85–105
- Sanchez-Jankowski MS (1991) *Islands in the street: gangs and American Urban Society*. University of California Press, Berkeley
- Schaefer DR (2012) Youth co-offending networks: an investigation of social and spatial effects. *Social Networks* 34(1):141–149
- Short JF, Strodtbeck FL (1963) The responses of gang leaders to status threats: an observation on group process and delinquent behavior. *Am J Sociol* 68(5):571–579
- Skogan WG, Hartnett SM, Bump N, Dubois J, Hollon R, Morris D (2009) Evaluation of ceasefire-Chicago. http://www.ipr.northwestern.edu/publications/ceasefire_papers/mainreport.pdf. Retrieved 8 Mar 2012
- Starbuck D, Howell J, Linquist D (2001) *Hybrid and other modern gangs*. U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention, Washington, DC

- Thornberry TP, Lizotte AJ, Krohn MD, Smith CA, Porter PK (2003) Causes and consequences of delinquency: findings from the rochester youth development study. In: Thornberry TP, Krohn MD (eds) *Taking stock of delinquency: an overview of findings from contemporary longitudinal studies*. Kluwer/Plenum, New York, pp 11–46
- Tita G, Radil S (2011) Spatializing the social networks of gangs to explore patterns of violence. *J Quant Criminol* 27(4):521–545
- Tita G, Jack Riley K, Ridgeway G, Grammich C, Abrahamse AF, Greenwood PW (2003) *Reducing gun violence: results from an intervention in East Los Angeles*. RAND Corporation, Santa Monica
- Venkatesh S, Levitt SD (2000) An economic analysis of a drug-selling gang's finances. *Q J Econ* 115(3):755–789
- Wasserman S, Faust K (1994) *Social network analysis: methods and applications*. Cambridge University Press, New York

Social Networks

- ▶ [Network Analysis in Criminology](#)
- ▶ [Social Network Analysis and the Measurement of Neighborhoods](#)

Societies

- ▶ [History of Police Unions](#)

Socio-spatial Criminology

- ▶ [Geography of Crime and Disorder](#)

Spatial Approaches to Crime

- ▶ [Inferential Crime Mapping](#)

Spatial Influence

- ▶ [Forecasting of Shootings Using Risk Terrain Modeling](#)

Spatial Models and Network Analysis

Steven M. Radil

Department of Geography, Ball State University, Muncie, IN, USA

Overview

Several issues in criminology have led to the widespread use of spatial econometric statistical models. This entry outlines the principles and concepts behind the use of spatial models in criminology, including a description of the most commonly used models in spatial criminology. Next is a discussion of a critical but often overlooked element of spatial model specification, the construction of the spatial weights matrix, which should be theory dependent. The entry concludes with a discussion of the limitations of distance- or proximity-based approaches to constructing the spatial weights matrix, including the use of social network analysis to overcome these limitations.

Foundations of Spatial Modeling in Criminology

The recognition of geography as a factor in the explanation of a multitude of social phenomena has been an increasingly notable component of quantitative criminology. Research concerned with basic and applied questions pertaining to crime, criminal justice, and policy now typically incorporate geographic or spatial elements into analyses that utilize quantitative methodologies. An important reason for the adoption of spatial perspectives for quantitative criminology has been a growing recognition of the importance of context to human action. Incorporating context into quantitative models of human behavior is an ongoing focus of the subfield of spatial criminology.

The analysis of spatial phenomena in criminology has been made possible in recent years by the ongoing development of statistical techniques

that attempt to deal with some of the unique problems of spatial data, especially spatial dependence. Dependence, or the tendency of characteristics of a given location to correlate with those of nearby locations, is a foundational issue in quantitative geographical analysis (e.g., Anselin 1988). When present in spatially organized data, spatial dependence can result in the spatial autocorrelation of regression residuals, which simply means that the values of variables sampled at nearby locations are not independent from each other (such values may be either positively or negatively correlated). The presence of spatial autocorrelation may lead to biased and inconsistent regression model parameter estimates and increase the risk of a type I error (falsely rejecting the null hypothesis). Accordingly, a critical step in statistical model specification when using spatial data is to assess the presence of spatial autocorrelation.

The statistical models developed to deal with spatial dependence, commonly called spatial econometric models, are now increasingly of interest to researchers from diverse disciplines investigating a wide range of issues, including crime and criminal justice. Many criminologists utilizing quantitative techniques now view spatial econometric regression models as an improvement non-spatial techniques as a result of the growing recognition that dependencies in the spatial structure of research data limit the reliability of inferences from conventional regression models (Anselin 1988). While there are now a variety of methods to address spatially auto correlated data in regression models, the standard workhorse in spatial regression is what are commonly referred to as simultaneous autoregressive (SAR) models.

SAR models can take three different basic forms (see Anselin 1988, 2002). The first SAR model assumes that the autoregressive process occurs only in the dependent, or response, variable. This is called the “spatial lag” model, and it introduces an additional covariate to the standard terms for the independent or predictor variables and the errors used in an ordinary least squares (OLS) regression (the additional variable is referred to as a “spatial lag” variable which is

a weighted average of values for the dependent variable in areas defined as “neighbors”). Drawing on the form of the familiar OLS regression model and following Anselin (1988), the spatial lag model may be presented as

$$Y = \rho W y + X \beta + \varepsilon,$$

where Y is the dependent variable of interest, ρ is the auto regression parameter, W is the spatial weights matrix, X is the independent variable, and ε is the error term.

The second SAR model assumes that the autoregressive process occurs only in the error term. In this case, the usual OLS regression model is complemented by representing the spatial structure in the spatially dependent error term. The error model may be presented as

$$Y = X \beta + \varepsilon, \quad \varepsilon = \lambda W \varepsilon + \mu,$$

where λ is the auto regression parameter and ε is the error term composed of a spatially auto correlated component ($W \varepsilon$) and a stochastic component (μ) with the rest as in the spatial lag model.

The third SAR model can contain both a spatial lag term for the response variable and a spatial error term, but is not commonly used. Other SAR model possibilities include lagging predictor variables instead of response variables. In this case, another term must also appear in the model for the auto regression parameters (γ) of the spatially lagged predictors (WX). This model takes the form

$$Y = X \beta + WX \lambda + \varepsilon.$$

Combining the response lag and predictor lag terms in a single model is also possible (sometimes referred to as a “mixed” model).

As Anselin (1988) observes, spatial dependence has much to do with notions of relative location between units in potentially different kinds of space, such as a conceptual social space. Accordingly, SAR models share a number of common features with network autocorrelation models. Substantively, spatial and network approaches have been used to explore similar

questions pertaining to influence and contagion effects, albeit among different units of observations (see Marsden and Friedkin 1993 for examples). In such cases proximity or connectedness is assumed to facilitate the direct flow of information or influence across units. Individuals or organizations are also more likely to be influenced by the actions, behaviors, or beliefs of others that are proximate on different dimensions, including geographical and social space.

The Spatial Weights Matrix (W)

Spatial econometric regression models demand a careful consideration of both the theoretical and empirical spatial structure of the data in question (e.g., Florax and Rey 1995). However, the advent of new spatial econometric software packages has lowered the traditional technical barriers to the use of these techniques while simultaneously making it easy for users to choose among a few predefined spatial structures. New users may therefore be unfamiliar with the importance of the formal spatial structure to analytic outcomes and less likely to carefully consider their choice (Leenders 2002). The small literature new users may draw upon to understand this issue remains quite technical in nature, even when proclaiming the opposite intent (e.g., Griffith 1996). Taken as a whole, this state of affairs is less than ideal and unlikely to encourage careful thinking about space. However, the need to formalize the empirical spatial structure of the data for modeling is also an opportunity to reflect on the theoretical interaction of the geography and social processes being studied. In this manner, spatial analyses of human behavior and outcomes are at the core of an emerging “spatially integrated social science” identified by Goodchild et al. (2000), and an opening for the investigation of how human behavior and space is mutually constituted.

As typically argued in many geographic literatures and increasingly in quantitative criminology, spatial perspectives are important for both theoretical and practical reasons. Theoretically, spatial perspectives are of interest because of the long-standing interest in the social production of space in geography. Following Lefebvre’s (1991)

understanding of space as both social process and social product, the spatial structures of a given phenomenon are commonly investigated as underlying cause, constructed outcome, or both. For example, the perception of rivalries over territorial control between a set of gangs and the actual construction of bounded turf are simultaneously a social and a geographic phenomenon. The competitions are inseparable from the geography and vice versa. Hence modeling a social process such as gang rivalry requires a consideration of the modeling the social construction of space (Radil et al. 2010).

On the other hand, consideration of geography is a practical methodological matter. Data with a geographical component has important implications for statistical analyses; if processes that are affected by the underlying spatial structure in a study area are not accounted for, inferences will be inaccurate and estimates of the effects of independent variables may be biased (Anselin 1988). Perhaps the most important reason for the interest in quantitative spatial methods is the most straightforward: nearly all social science data is spatially organized, and ignoring this structural element is increasingly seen as untenable. To accommodate these issues, statistical models have been developed that attempt to deal with issues of spatial dependence. Conventionally, this is done through either introducing an additional covariate (the “spatial lag” variable discussed in the previous section) or by specifying a spatial stochastic process for the error term.

These models, now seen as both viable and important for social science research, have been discussed and exemplified in depth (see Anselin 2002 for an overview), but an essential element of these models remains largely ignored in the literature, despite the major theoretical and methodological implications. Both the lag and error models attempt to estimate regression parameters in the presence of presumably interdependent variables (Anselin 1988; Leenders 2002). This estimation process requires the analyst to define the form and limits of the interdependence and formalize the influence one location has on another. In practice, this is accomplished by identifying the connectivity between the units of the

study area through a $n \times n$ matrix. The matrix is usually described in the literature as a “spatial weight” or “spatial connectivity” matrix and referred to in the preceding SAR models as “ W .”

This W , or matrix of locations, formalizes a priori assumptions about potential interactions between different locations, defining some locations as influential upon a given location and ruling others out. A simpler way of describing this is that W identifies, in some cases, who is a neighbor and who is not or with whom an actor interacts. However, the construction of W is more than just an empirical choice about neighbors. It is a theoretical decision regarding the spatiality of the social processes being discussed and one that has implications for the statistical estimates generated.

As W is supposed to represent a formal model of connections between geographic locations, how one translates theories about influence and its mechanisms across space into a formal mathematical construct is an important step. Put another way, at its core, W is really a theoretical geography of interaction. However, as a practical matter, the spatial analytic geography literature focuses on modeling interaction through a distance-based logic that typically takes one of two forms: contiguity or distance (Cliff and Ord 1981; Griffith 1996). Both of these spatial themes have been mobilized for constructing W s for various kinds of measures of spatial autocorrelation dating back to Cliff and Ord’s seminal work (1981). Contiguity, or the physical connections between locations, is emphasized in issues that focus on areal spaces, especially ecological studies that use aggregated data and attempt to evaluate the relationships between neighborhood composition and crime outcomes. Distance between locations of interest also remains an important concept for criminology, in particular for journey-to-crime research.

Theorizing W with Social Network Analysis

Whether contiguity- or distance-based, the geography of influence is typically imagined and implemented in SAR models as a kind of gradient that uniformly diminishes with increasing distance, also known as the concept of distance decay (see Stephenson 1980). Classic examples

of distance-decay thinking in social science are Boulding’s (1962) “loss of strength gradient” which argued that military power has a direct inverse relationship with distance and Tobler’s (1970) so-called “first law” of geography which states that everything is related to everything else, but near things are more related than are far things.

At a basic level, distance decay is the foundational concept behind the implementation of W in most spatial economic models, and the development and dissemination of spatial analytic software allow users to easily create W s from spatially organized data using the classic spatial forms of areal contiguity or point-to-point distance. While such software often guides researchers through the practical steps needed to create a matrix of spatial interaction or influence grounded in distance decay, there is no drop-down menu to offer guidance as to how best to capture the specific geography, or spatiality, of the social processes being analyzed. For any given research topic, are immediately contiguous areal neighbors enough, or should more distant neighbors also be included? If distance matters, at what distance does influence begin to diminish? Does influence diminish over distance in the same way for all social actors or behaviors? More to the point, why and in what way does distance “matter” in the operation of the social process under investigation? Addressing these sorts of questions remains the key challenges for a theoretically informed spatial analysis in criminology.

Somewhat surprisingly, theoretical discussions about the nature of W and practical discussions of how different specification choices may affect regression results have also been underemphasized in most spatial analytic literature. Despite these noteworthy efforts, Leenders (2002: 44) is correct in his assessment that “the effort devoted by researchers to the appropriate choice of W pales in comparison to the efforts devoted to the development of statistical and mathematical procedures.” The net effect of this lack of attention is that theoretical conceptions about the role space plays in producing empirical patterns in a given dataset are often afterthoughts and little effort is given to understanding how

sensitive model results may be to different specifications of W . Hence, the vision of a “spatially integrated social science” (Goodchild et al. 2000) remains unfulfilled, because when space is included in the analysis of social processes, it is often added in a default form without consideration of the geographic expression of the processes in question.

Criminologists have begun to explore such questions, often turning explicitly to the concept of social networks and the associated field of social network analysis. For example, social networks within neighborhoods are often implicated as being important mechanisms in the production and maintenance of safe, low-crime neighborhoods (e.g., Sampson et al. 1997). Such approaches argue that individual-level social bonds among local residents facilitate the formation of informal social control and the creation of shared goals and trust that regulate and censure local activities. A general acceptance of the importance of networks of local social relationships for understanding rates and patterns of neighborhood crime has prompted researchers to look into the kinds of social networks operating in communities in an effort to gather information regarding social ties among local residents as well as peer associations with delinquent others.

Beyond examining the degree to which individual-level ties influence crime patterns within a community, researchers are also beginning to explore the importance of institutional or organizational ties that can bridge communities. Sampson (2004: 158) argues for reconceptualizing neighborhoods “as nodes in a larger network of spatial relations” in order to account for the various ties that can link residents across space. Although Sampson (2004) does not specifically suggest what kinds of local institutions or organizations one should consider as important to explaining crime, in general terms he refutes the notion that neighborhoods are analytically independent and argues that ecological models of crime need to consider the different ways in which the observable outcomes in one neighborhood are partly the product of social actions and activities that can stretch beyond local communities (Morenoff et al. 2001; Sampson 2004).

Drawing on such arguments about the importance of both localized and no localized social networks, an emerging framework in quantitative criminology argues for a “deductive approach” to incorporating space built on the understanding that social networks are inherently geographic, existing both within and across localities and connecting communities separated by distance (e.g., Radil et al. 2010; Tita and Radil 2010). This analytic framework allows influence to take place not just between geographically proximate neighbors (as with conventional distance-decay-based spatial regression models) but also between locations that are connected directly by social networks. By carefully considering and allowing for processes that extend beyond (or perhaps preclude) spatially adjacent areas, one can ensure that the spatial weights matrix adequately captures the realities of the mechanisms of influence. Just as Morenoff (2003: 997) argues that spatial analysis “expands the neighborhood-effects paradigm by considering not only the local neighborhood but also the wider spatial context within which that neighborhood is embedded,” a careful consideration of the spatial dimensions of social influence is under way as part of an attempt to facilitate the inclusion of the “wider *social* context of neighborhoods” into quantitative models of crime.

An early example of this emerging framework was found in the work of Mears and Bhati (2006), which links adjacent as well as nonadjacent areas to one another if the residents are economically and demographically similar by constructing W based on “similarity” that links socially similar areas together regardless of spatial proximity. Most recently, this framework has been advanced by research into the importance of both social and spatial linkages among gangs embedded within various types of local networks (Tita and Greenbaum 2009; Radil et al. 2010; Tita and Radil 2010). Rather than inferring connections through social similarity, the framework first demonstrated by Tita and Greenbaum (2009) accounts for influence by simultaneously considering linkages through geographic proximity between neighborhoods as well as through specific social ties that connect places. This approach carefully

identifies direct social connections between neighborhoods based on rivalries between urban street gangs that are not geographically proximate, while also preserving the underlying spatial structure of the entire study area. Tita and Radil (2010) examined the utility of such an approach by comparing spatial regression model outcomes of gang violence using three differently constructed W s based on geographic adjacency, social connectivity, and a blend of both. Tita and Radil (2010) argued that violence committed by, and against, gang members in a socially and geographically distinct area of Los Angeles is largely a function of a *social* process that spans the local geography in such a way that violence in noncontiguous areas impacts levels of violence in a focal neighborhood and that blended W that incorporated both distance-based spatial effects and connections formed by specific social relations across neighborhoods offered the most meaningful results to the explanation of gang violence, both theoretically and statistically.

These studies reflect a growing interest and recognition in quantitative criminology of the need to not simply incorporate space and spatial effects in the study of crime but to do so in a way that moves beyond the uncritical notions of uniform distance decay that underpin SAR models. Specifications of the spatial weights matrix, or W , that rely solely on spatial contiguity or distance to define the spatial reach of the various social processes posited to be responsible for clustering force researchers to assume that all such processes decay rapidly and uniformly over geographic distance and therefore matter only among spatially adjacent neighbors. Furthermore, even when multiple social processes are considered, the conventional modeling approach has been to specify a single W rather than specify different kinds of connections between places for different social processes. In addition to making it impossible to parse the impact of one process from that of another, this is an a theoretical approach to understanding why and how space matters (Leenders 2002). By carefully considering the socio-spatial dimensions of the criminal phenomena of interest and of the actors behind them, the work of Mears and Bhati (2006) and

(Tita and Radil 2010) demonstrate new possibilities by creating W s that explicitly capture the social and geographic dimensions of spatial influence through incorporating social networks. Further, such approaches allow theories of spatial influence to guide the construction of W , a meaningful advance given Leenders' (2002) concern with the lack of careful consideration of the underlying social processes of influence exhibited by researchers in their construction of weights matrices.

This is not to say that all the challenges of incorporating space into studies of crime have been met. For example, the developing arguments about the importance of microscale units of analysis in criminology (Hipp 2007) complicate the need to model interaction as the likelihood for spillover or other unmeasurable spatial effects increases as the areal scale of the unit of analysis decreases. Additional challenges include the need to incorporate dynamism and change over time in social networks and space and the need to consider overlapping networks (Ettliger 2003) or multiple spatialities (Leitner et al. 2008) when theorizing the interactions between people and space. Nonetheless, the value of theoretically grounded notions of the geographic complexities of the mechanisms of influence in terms of positing various theories and mechanisms responsible for observable patterns of crime cannot be overstated. The current efforts underway in quantitative criminology have demonstrated the value of employing such an approach which should encourage future research into how social networks are involved in the social production of space and how the methods of social network analysis may be further utilized to understand it and how the structure of social networks have implications for material geographies of crime.

Recommended Reading and References

- Anselin L (1988) *Spatial econometrics: methods and models*. Kluwer, Boston
- Anselin L (1998) Exploratory spatial data analysis in a geocomputational environment. In: Longley P, Brooks S, McDonnell R, Macmillan B (eds) *Geocomputation, a primer*. Wiley, New York

- Anselin L (2002) Under the hood: issues in the specification and interpretation of spatial regression models. *Agric Econ* 27(3):247–267
- Anselin L, Florax R, Rey S (2004) *Advances in spatial econometrics: methodology, tools and applications*. Springer, Berlin
- Boulding K (1962) *Conflict and defense*. Harper, New York
- Cliff A, Ord J (1981) *Spatial processes, models, and applications*. Pion, London
- Ettlinger N (2003) Cultural economic geography and a relational and microspace approach to trusts, rationalities, networks, and change in collaborative workplaces. *J Econ Geogr* 3(2):145–171
- Florax R, Rey S (1995) The impacts of misspecified spatial interaction in linear regression models. In: Anselin L, Florax R (eds) *New directions in spatial econometrics*. Springer, Berlin, pp 111–135
- Goodchild M, Anselin L, Appelbaum R, Harthorn B (2000) Toward spatially integrated social science. *Int Reg Sci Rev* 23:139–159
- Griffith D (1996) Some guidelines for specifying the geographic weights matrix contained in the spatial statistical models. In: Arlinghaus S (ed) *Practical handbook of spatial statistics*. CRC Press, Boca Raton, pp 65–82
- Hipp J (2007) Block, tract, and levels of aggregation: neighborhood structure and crime and disorder as a case in point. *Am Sociol Rev* 72:659–680
- Leenders R (2002) Modeling social influence through network autocorrelation: constructing the weight matrix. *Soc Netw* 24:21–47
- Lefebvre H (1991) *The production of space* (trans: Nicholson-Smith D). Blackwell, Oxford/Basil
- Leitner H, Sheppard E, Sziarto K (2008) The spatialities of contentious politics. *Trans Inst Br Geogr* 33(2):157–172
- Marsden P, Friedkin N (1993) Network studies of social influence. *Soc Methods Res* 22:125–149
- Mears D, Bhati A (2006) No community is an island: the effects of resource deprivation on urban violence in spatially and socially proximate communities. *Criminology* 44(3):509–548
- Morenoff J (2003) Neighborhood mechanisms and the spatial dynamics of birth weight. *Am J Sociol* 108:976–1017
- Morenoff J, Sampson R, Raudenbush S (2001) Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence. *Criminology* 39:517–560
- Radil S, Flint C, Tita G (2010) Spatializing social networks: using social network analysis to investigate geographies of gang rivalry, territoriality, and violence in Los Angeles. *Ann Assoc Am Geogr* 100(2):307–326
- Sampson R (2004) Networks and neighborhoods: the implications of connectivity for thinking about crime in the modern city. In: McCarthy H, Miller P, Skidmore P (eds) *Network logic: who governs in an interconnected world?* Demos, London, pp 157–166
- Sampson R, Raudenbush S, Earls F (1997) Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science* 277:918–924
- Stephenson L (1980) Central graphic analysis of crime. In: Georges-Abeyie D, Harries K (eds) *Crime: a spatial perspective*. Columbia University Press, New York, pp 146–155
- Tita G, Greenbaum R (2009) Crime, neighborhoods and units of analysis: putting space in its place. In: Weisburd D, Bernasco W, Bruinsma G (eds) *Putting crime in its place: units of analysis in spatial crime research*. Springer, New York, pp 145–170
- Tita G, Greenbaum R (2011) Spatializing the social networks of gangs to explore patterns of violence. *J Quant Criminol* 27(4):521–545
- Tita G, Radil S (2010) Making space for theory: the challenges of theorizing space and place for spatial analysis in criminology. *J Quant Criminol* 26(4):467–479
- Tobler W (1970) A computer movie simulating urban growth in the Detroit region. *Econ Geogr* 46:234–240
- Weisburd D, Bernasco W, Bruinsma G (eds) (2009) *Putting crime in its place: units of analysis in spatial crime research*. Springer, New York

Spatial Perspectives on Illegal Drug Markets

Lusine Tarkhanyan
Department of Security and Crime Science,
University College London, London, UK

Synonyms

[Drug corners](#); [Drug dealing places](#); [Illicit retail marketplaces](#); [Open drug market](#); [Street drug market](#)

Overview

Illegal drug markets have been the subject of considerable research in environmental criminology. Two main types of study have been conducted: evaluations of law enforcement strategies designed to police illegal drug markets and empirical studies which focus on the social and environmental context of places where drug distribution and consumption tends to occur. By addressing the dynamic relationship between drug trading, market procedures, environment,

and police interventions, both approaches share the common aim of informing more effective crime prediction and control strategies. Since the main purposes of illicit drug markets are to both provide and secure the supply chain of illicit commodities, understanding *how* and *where* buyers and dealers position themselves to engage into transactions is critical to crime prevention.

This entry addresses what is normally the last and lowest stage in the illegal drug supply chain – the drug market. Here, the term *drug market* refers to urban location(s) where illicit commodities are illegally traded. This entry considers the extent to which illicit drug markets depend on the specific geography of places and on their amenities. The discussion will highlight how the areas surrounding these places affect their suitability as drug markets and what makes them attractive from an economic perspective. One proposition that will be discussed is that since the demand for drugs displays a nonrandom spatial pattern across the urban environment, crime prevention strategies might usefully focus on targeting and altering specific locations that are conducive to drug crime, rather than or as well as, focusing on particular individuals or groups of people.

This entry is organized as follows: first, the theoretical context and a dilemma regarding the location of drug markets will be discussed. Next, a marketing perspective will be discussed that has been proposed to explain the reasoning behind locational choices. The subsequent sections will explore the relational links between drug markets, drug-related crimes, and law enforcement actions. And, finally, further questions and potential research areas will be highlighted.

Theoretical Rationale

An urban-dweller's routine consists of a number of activities distributed across the town or city, and the practice of committing crime is no exception. Environmental criminology identifies three factors which influence the probability of a crime event occurring: the "when," the "where," and the "how"

(Brantingham and Brantingham 1995). The "where" refers to the spatial location of a crime. Researchers suggest (Cohen and Felson 1979) that patterns of criminal activity are influenced by both the physical environment and a criminal's perception of what constitutes a good or bad opportunity for committing crime. Furthermore, it is proposed (Brantingham and Brantingham 1984; Cohen and Felson 1979) that the way in which a criminal moves around during their daily routine provides them with the knowledge about their surrounding environment and the opportunities it offers for criminal activity. Therefore, according to this perspective, crimes should be more highly concentrated within criminals' awareness spaces, and these are likely to be defined by the constellation of places they move around within their noncriminal daily routines (Cohen and Felson 1979; Brantingham and Brantingham 1984).

This rationale has been applied by scholars (Eck 1995; Rengert et al. 2005) to try to explain how both drug buyers and dealers identify places that are potentially suitable for transactions of the kind they involve themselves in. Moreover, it is proposed that rational decision making on the part of both participants – buyer and dealer – in the transaction influences the geographical distribution of drug dealing locations (Eck 1995). Depending on the method of transaction, drug markets can be very large and dispersed, or small and more concentrated (Eck 1995). Scholars (Rengert et al. 2005; McCord and Ratcliffe 2007) have identified that profitable drug markets are often situated in or near socially disorganized neighborhoods, where there is a lack of social resistance to the market's existence and from which there may be a ready supply of people willing to purchase drugs. However, the rationale for choosing these neighborhoods remains disputed: it is still unknown whether the locations are chosen initially due to their accessibility to a large number of drug users and their presence undermines the neighborhoods' social organization (Eck 1994) or whether the initial lack of social resistance provides optimal grounds for establishing drug markets (McCord and Ratcliffe 2007).

Additionally, it has been noticed that the success of trading locations is associated with types of land uses that are conducive to, or generate, other types of crime. For example, drug dealing is likely to happen close to facilities which inherently and routinely generate a large flow of people. These are mainly open public spaces, retail, entertainment facilities, and transport interchanges that are associated with low levels of adequate guardianship or place management (Eck and Wartell 1996). In their analysis, Rengert and colleagues defined two types of built environment facilities that may be associated with the locations of drug markets. First, there are those which indirectly increase the profits from drug sales, because they facilitate nonresidents' access to an area. An example of this would be transport interchanges, which provide easy access to the drug markets (Brantingham and Brantingham 1995). Second, there are those which generate opportunities for drug transactions because they are used routinely by potential drug buyers, for example, areas near to homeless shelters or pawnshops, where potential buyers can readily convert stolen goods to cash (Anderson 1999).

The Main Dilemma

In the same way that legitimate businesses do not select their location at random, drug market placement may reflect rational decision making (Eck 1995; Rengert et al. 2005): the main aim of both activities is to attract customers and supply products to them in order to make a profit. From a strategic perspective, when deciding where to locate their stores, retailers focus on finding an accessible spatial location which will attract many potential customers. However, their locational choices are also constrained by different planning regulations and environmental impact assessments, which are required by authorities to allow the placement of the shop.

Rengert (1996) suggests that drug dealers may follow a similar logic and try to identify potential profitable sites. As Rengert states, "the quality of the sale's location is directly related to the

quantity of profit for the illegal drug sales" (Rengert 1996). However, in contrast to legitimate trading, in the case of illegal markets, offenders have an additional goal: staying safe and unnoticed so as to avoid arrest (Reuter and MacCoun 1993). Thus, drug dealing locational choices are also constrained by the presence of legitimate and capable guardians who discourage criminal opportunities (Eck 1994). These include security guards, home and shop owners, and generally people who manage such places. In their study of drug and disorder problems, Mazerolle et al. (1998) found that at the level of street blocks, the place managers who engaged in crime prevention activities played an important role in guarding places from drug problems. Moreover, they found that place managers who engage their neighbors from the same street block in crime prevention efforts are more effective than individual efforts.

Eck (1995) has termed this specific aspect of drug markets the accessibility vs. security dilemma: "how to exchange illicit goods or service when the exchange process is very risky" (p. 71). Illicit drug markets typically face the conflict between needing to be accessible to many customers, including complete strangers, and avoiding the vulnerabilities associated with drug sales. Due to the illicit nature of the market, both customers and dealers are usually in a vulnerable position, since they run the risk of both legal intervention and being cheated or robbed by their counterpart during the course of a transaction. Of course, there is no means of securing the transaction through law enforcement or of resolving such conflict of interests through legal channels (Reuter and MacCoun 1993). Thus, violence is a very common means of regulating and resolving disagreements (Goldstein 1985), especially in street-based sales, where exchanges take place outdoors, often between anonymous participants. Both buyers and dealers will thus be motivated to limit their accessibility so as to reduce risk, seeking locations that they personally consider to be safest. Such places may be enclosed and familiar locations. For example, indoor sales may be made from fixed locations such as drug houses or deliveries made to indoor

locations specified by the customers (Curtis and Wendel; 2000). However, in both of these scenarios, one of the participants will always be at a greater risk than the other, since they will not be familiar with the location. In some cases, unfamiliarity may lead to them avoiding a location altogether.

Eck proposes two models of drug markets which can overcome the access-security dilemma, given the constant risk of police presence. The first is the *social network-based transactions model*, in which security is provided through a network of trusted people. The second is the *routine activity model*, in which both participants use their legitimate daily activities to search for places which are potentially appropriate for engaging in drug deals. In the first model, there is no attachment to a specific location: through a social network both parties can arrange a mutually accessible location, potentially based on their routine. This type of market tends to be closed in nature, and the market itself may be dispersed over a large geographical area, leading to a low spatial concentration of drug dealing incidents (Eck 1995). Although a network-based exchange offers security, it also limits the number of participants in the transaction, which in turn reduces the potential for profit.

The second model proposed assumes that the spatial-temporal patterns of both participants' legitimate daily activities determine where their drug transactions occur. These markets tend to be spatially clustered and to focus on locations which are familiar to both participants, thereby reducing their perception of potential risks. These markets also need specific operating conditions: when surrounded, for instance, by a large number of legitimate activities and a constant flow of foot traffic, it is easier to blend into the crowd and search for potential customers. This type of market is often an open street market with a high frequency of transactions between anonymous participants. The market permits equal access to all participants and is located near places with mixed land uses and a high concentration of activities, such as shopping centers, high streets, and transport interchanges. Eck proposes that place managers from the legitimate sphere who

control and manage these locations have a particularly important role in impeding this type of market.

It should be noted that not all locations which are next to shopping malls or transport facilities, for example, will form drug markets. According to this model, places which are more likely to have drug markets will be attractive from a retail perspective, offering a balance between the spatial distribution of those who demand the product and the distance they would be required to travel to the market. This concept will be discussed further in the following section.

State of the Art: Locational Choices and Spatial Economics

As with many crime types, open drug markets tend to concentrate geospatially: from all the available urban locations, there are very few which are well suited to illegal drug trading. Weisburd et al. (2004) found that a small number of street junctions in Jersey City (4.4 %) accounted for almost half of the drug sales arrests in the city. This demonstrates considerable spatial concentration of markets. According to Kleiman (1991 cited in Taniguchi et al. 2009) the reason for spatial clustering is that a large number of dealers operate in a single location, gaining security from arrest by spreading the risk of apprehension across all dealers. The same logic applies to customers who would prefer to be in a crowd than alone.

Rengert et al. (2005) propose that the exact location of open drug markets is likely to depend on the convergence of conditions that are most suitable for both participants involved in transactions and spatial patterns in demand. They suggest the model of "agglomeration economics" operates, as is the case with legal trade. That is, after a location becomes known as a site for specific goods, more customers will visit the area in the search of that good. At a certain point, the number of buyers that visit the area will be sufficient to support further suppliers of the product(s) and so more retailers locate there. In Wilmington, USA, researchers discovered that

the likelihood of large drug markets being located near to each other is quite high. They say “there are best places to sell illegal drugs because these are places where demand is focused spatially” (Rengert 1996). That is, in order to stay profitable, a drug market’s location should be attractive enough to a sufficient number of drug users. The attractiveness of the place is partially determined by the surrounding facilities, but above all by how far a buyer is prepared to travel to make a purchase (Pettiway 1995).

In legal trading, markets for goods which are highly valuable but purchased infrequently are usually found in very accessible urban locations, which may attract many potential customers from remote locations. In contrast, local markets tend to supply items, which consumers will wish to purchase frequently but will be prepared to travel only short distances to get them. In their study, Rengert and colleagues (et al. 2005) discuss concepts from the economics literature – *threshold population* and *range* – to frame a discussion on how the demand for drugs and the distance required to reach a market likely affect the location and stability of drug markets. Threshold is defined as a minimum number of customers required for a market to stay profitable. The concept of range concerns the distance that a buyer is prepared to travel to purchase a good. Pettiway (1995) suggests that the spatial range of a market that caters for vehicular movement will be larger than one that caters for pedestrian movement. Simply put, if the demand for a drug market is situated within the physical catchment area (an area to which they do or will travel) of many potential drug buyers, that market will remain stable or may even grow.

According to this rationale, dealers will try to establish markets at locations that reduce the total distance that customers will have to travel to reach them. In order to determine the demand for drug markets, US researchers have used methods of sociodemographic profiling to identify those types of neighborhoods with the characteristics that are associated with increased risk of drug use. They propose that if a drug dealer wants to sell drugs to a local community, he must first identify possible users. If the local demand is

not enough to sustain a drug market, the dealer must consider factors that would attract potential customers into the neighborhood. For example, the market should be situated in close proximity to transport facilities, which are used routinely by many potential drug addicts, or it should be accessible to modes of private transport.

Furthermore, since drug markets are established along routes which are used on a daily basis by many potential drug buyers, the location and retail characteristics of these markets can vary considerably. Depending on where users and dealers live relative to the market, drug markets are classified as *neighborhood*, *open regional*, *semi-open regional*, and *closed regional markets* (Reuter and MacCoun 1993; Eck 1994). Neighborhood markets are described as places in which both customers and dealers are from the vicinity: they might be neighbors or know each other. Open regional markets are usually established close to places which are routinely used by a large number of nonresidents, such as shopping malls. These markets are large enough to support several competing drug dealers and can secure a high frequency of transactions. Semi-open regional markets are typically established when both participants aim to lessen the risks involved in open street markets. They do not live in the same area and they only interact if they know each other or have been referred by a third person. Closed regional markets’ locations are determined by both participants of the transaction and are distributed over a wide area. The clientele of such markets are established through a network of friends and other trusted people. This type of market is the norm in wholesale drug dealing.

Alongside spatial clustering, drug markets are typically located in close vicinity to certain facilities: shopping centers (Eck 1995), high schools (Roncek and LoBosco 1983), bars (Roncek and Maier 1991), cash stores and pawnshops (Anderson 1999), transport links, train stations and highways (Eck 1994), and vacant homes (Rengert et al. 2005). Rengert and colleagues (2005) found that there are also facilities that discourage the establishment of drug markets in an area, such as police and fire stations or courts and federal

buildings. Given that accessible drug dealing locations should offer good retail potential, drug marketplaces can be further classified according to the level of pedestrian and traffic accessibility, which may bring potential customers to the area. It can be suggested that, depending on a market's geographical positioning in the city, the level of accessibility will vary – from locally to regionally accessible markets. For example, Eck (1994) found that in San Diego, outdoor drug markets formed at locations about two blocks away from major transportation arteries, suggesting that they were regionally accessible markets. Importantly, this suggests that although offenders aim to sell drugs from accessible locations, they do not tend to do so on the major roads (presumably as a way of reducing risk). That is, for regionally accessible markets, operating in close proximity to major roads may offer an acceptable balance of custom and safety. In comparison, in Philadelphia, Rengert and colleagues (2005) found a high concentration of drug markets located in the suburbs, located away from major roads, suggesting that these marketplaces are oriented to local rather than regional demand.

State of the Art: Drug-Related Crime

The drug-crime relationship is usually examined through the theoretical perspective of a “tripartite framework” advanced by Goldstein (1985), which categorizes drug-related crime into three groups. First, there is “psychopharmacological” crime in which drug use affects aggression, thereby increasing the chance of criminal behavior. Second, there is “economic-compulsive” crime, in which crime is committed to finance drug use. Third, there is “systemic” crime, in which criminal behavior occurs in order to secure the drug supply chain, protect profits, control locations, compete for customers, or resolve conflicts of interests. A fourth category has also been proposed (Blumstein 1995 – cited in MacCoun et al. 2003; Curtis and Wendel 2000; Rengert 1996), which focuses on the negative impact of drug markets on local communities, mainly in terms of how drug dealers' behavioral codes and

manners impact on community members who are not directly related to the market, for instance, by increasing gun carrying for the purpose of self-defense or conflict resolution.

Despite extensive research on the drug-crime nexus, little is known about how drug-related crime and drug markets relate to each other geographically. Based on crime pattern theory, it is suggested that drug markets can be both crime generators and crime attractors. The former may occur due to the inherent criminality of drug markets – with those who buy or sell drugs engaging in other types of crime while they are in an area, even though they may not have travel to such locations with the intention of so doing. Even though they do not travel to such locations to commit other forms of crime, this will nevertheless affect such crime rates in the area. Apart from potential drug buyers, drug markets can *attract* other offenders who purposely seek out criminal opportunities. In their research, Weisburd et al. (2004) found that drug-related crime clusters spatially around drug markets. Rengert and Wasilchick (1989) state that property offenses committed by drug users are spatially concentrated in close proximity to drug markets. Moreover, ethnographic research suggests that violence against dealers is not uncommon since they possess drugs and money.

Others (Eck 1994; Reuter and MacCoun 1993) suggest that depending on the retail nature of the market (street based, indoor from fixed locations, or delivery based), levels of violence and criminal behavior differ considerably. The indoor and delivery-based methods of transaction are safer than outdoor sales, since they enjoy an established and protected territorial boundary and/or security is provided by limiting sales to a network of trusted people.

The nature of a market – *local, export, public, and import* – is thought to affect levels of violence. According to Reuter and MacCoun (1993), local markets have the least violence, given that there is an established network of local social ties and all market participants interact on an ongoing basis within recognized territorial boundaries. Export markets also enjoy low levels of violence, since local dealers are interested in discouraging

violence in order to attract more customers from outside the local neighborhood. Public markets feature a high level of anonymous participants operating in public places and hence lack clear territorial boundaries. Due to the resultant competition and weak social ties between participants, which in turn promote potential mistrust and disagreement both between dealers and buyers and among dealers, these markets have considerable potential for violence. Import markets may feature a similar risk of violence, since neither dealers nor customers belong to the given neighborhood and the process of establishing territorial boundaries may cause dissatisfaction among local residents and tensions between dealers. As a result, conflict may arise between dealers and locals as well as between competing dealers.

“Systemic” violence, which occurs in drug markets as a means of controlling a location and competing for customers, has been investigated in relation to drug dealing gangs, who usually operate from the so-called street corners (Taniguchi et al. 2009). These street corners are financially more beneficial than other street segments due to the possibility of targeting a greater volume of passersby (Eck 1994). Consequently, gangs compete to defend and control the corners as a matter of prestige and economic gain. Scholars (Taniguchi et al. 2010) have found that there is more violence when multiple gangs operate from a single corner than when a single gang controls it. In the former scenario, violence between gangs arises from their competing interests in finding clients and establishing territorial boundaries. In comparison, when a single gang occupies a strategically positioned corner, it has an economic interest in keeping the area safe from violence in order to attract more customers (Cohen and Tita 1999; Goldstein 1985; Levitt and Venkatesh 2000 – quoted in Taniguchi et al. 2010).

Several authors (Eck and Maguire 2000; MacCoun et al. 2003) have noticed that drug markets are highly responsive and adaptive to law enforcement. They have suggested that violent behavior linked to competition may be caused by intensive law enforcement actions.

The link between police activities and drug markets will be discussed in detail in the following section.

State of the Art: Drug Markets and Law Enforcement

Apart from arresting drug dealers, one of the police’s main goals in targeting drug markets is to understand how the operational structure of these markets can be disrupted. Since drug markets cluster in relatively few locations, which also feature specific land uses which support them, crime prevention researchers (Taniguchi et al. 2010) suggest targeting specific locations instead of drug dealers. By applying the categorization of drug markets according to where dealers and buyers live, as discussed above, scholars (Reuter and MacCoun 1993) have identified the markets which are expected to be the most and least vulnerable to law enforcement activity. Local markets are, for example, the most resilient to police activities, because both participants in the transaction are from the same neighborhood with strong social ties. Consequently, if the market’s functioning is disrupted, the participants’ shared knowledge of the neighborhood allows the easy reestablishment of the market elsewhere. In contrast, public drug markets can be easily disrupted, because there will be relatively fewer locations which are well known to both dealers and buyers and which are suitable for drug transactions. Targeting these locations may be sufficient to keep drug dealers and buyers apart.

However, it has been claimed that disrupting the operations of drug markets will not solve the problem, since they are highly adaptive to law enforcement and might reopen nearby or simply start trading at different times. This amounts to claims of spatial or temporal displacement (Repetto 1974). Although displacement is a common criticism of geographically focused crime prevention efforts, this concern has not been substantiated by empirical research. Furthermore, a number of studies (Clarke and Weisburd 1994; Weisburd et al. 2004) have found the reverse effect, namely, a diffusion of

crime control benefits, whereby there is a reduction in the crime level in locations surrounding the area where police implemented crime prevention strategies. There are several explanations for this effect. From the perspective of agglomeration economics, “the immediate spatial diffusion is more likely than immediate spatial displacement” (Taniguchi et al. 2010). For example, it has been suggested that closing down the most profitable drug dealing location will make dealing in the neighboring locations less rather than more profitable. Dispersing the cluster of competing drug dealers will lead to a series of smaller marketplaces, each with fewer participants and drug transactions. According to this perspective, a forced change in spatial positioning can be sufficient in disrupting a drug market due to its detrimental effect on economic conditions.

Ethnographic research suggests (Weisburd et al. 2004) that after preventative interventions, drug dealers prefer to adapt their behavior and stay in the targeted area rather than move location. Both participants may implement less open modes of transaction, for instance, arranging the drug transaction time and place through mobile devices.

Weisburd and colleagues (2004) have studied the effects of geographically focused or hot-spot crime prevention strategies. Despite the widely held belief discussed above that hot-spot policing causes displacement, the researchers found no evidence on one of these. They strongly support the use of focused market targeting in police interventions since dealers resist moving away from a chosen area, as discussed above. They justify this support by emphasizing that dealers are part of the social and business community and when they relocate they may encounter violence caused by competition with established dealers in the new area. Moreover, existing networks of customers are not easily displaced to another location, since their limited geographical familiarity makes it harder to locate their dealer there. This can interrupt the supply-demand chain and hence affect the profitability of drug dealing at a particular place. This is one reason why focused hot-spot prevention may be beneficial to an area

since one consequence is that reestablishing the market is likely to require effort on the part of the dealer.

Overall, researchers (Rengert 1996; Weisburd et al. 2004) are in favor of hot-spot policing, believing it to produce strong crime prevention results. Since the change in economic and physical conditions makes the location less attractive for a dealer, they have strongly suggested paying more attention to the surrounding conditions which affect the marketplace’s economic attractiveness. However, caution should be taken in generalizing these results as several crackdown studies have found evidence that a diffusion of benefits to adjacent areas may be accompanied by partial displacement of the drug scene to indoor locations, other neighborhoods, or nearby metropolitan areas.

Future Research Directions

Overall, it can be noted that there is a particular environmental balance between physical and social characteristics which facilitate the optimal grounds for establishing a drug marketplace. As has been described above, the retail chain of drug supply operates differently depending on both the spatial structure of the area and the participants’ profile. Moreover, drug markets are more likely to be located in areas with low levels of social organization. It is hard, however, to establish whether the choice of location is determined by the level of social organization or whether the social organization is undermined by the presence of a drug market. Whatever the case, the ensuing lack of legal guardianship provides, from the dealers’ perspective, valuable protection from police enforcement.

For the future research directions, it may be worth looking at geography of separate crack or heroin markets alone and on crime concentration nearby. There is an opinion that crack markets tend to be more violent than markets for other drugs, but it is not supported with any empirical evidences. Additionally, it is useful to know whether markets that offer particular types of drug that is more expensive and difficult to source

are harder to stamp out for good and are more liable to displacement.

Although many researchers recognize the relationship between legitimate movement and drug markets, it is still unknown how much and what type of movement is required for the establishment of drug markets. Locational choices may, for example, be influenced by whether the market is frequented by pedestrian or vehicular traffic. A drug market which, for example, relies on vehicular traffic may need to be located in close vicinity to main roads with easy and quick access *to* and *from* the marketplace. In the case of pedestrian traffic, the market will probably be located near places where there are a sufficient number of legitimate activities which can conceal the drug transaction process. Presumably, there should be multiple escape routes, lookout points, and workarounds (Eck 1995). Since street networks influence the way people navigate towns and cities (Brantingham and Brantingham 1995; Hillier and Hanson 1984), and the street networks vary from place to another, caution should be taken when generalizing findings from one location to another. This is particularly true when comparing findings from studies conducted in automobile-based societies with regular grid-like street networks, such as America, with those from pedestrian-based societies with street networks which have grown organically, such as the UK. To be more explicit, all drug markets are not equal, and those factors that influence one type of market may be different to those that influence others. Understanding the factors that affect different types of markets would be a useful next step, and consideration of how the morphology of the street network might shape or support opportunities for drug crime would seem to be a logical starting point.

Another priority research area is the drug-crime connection. Rengert and Wasilchick (1989) discovered that offender's search behavior is oriented towards the location of drug marketplace. They showed that drug-dependent property offenders search for potential houses to burgle on their way to purchase a drug. Despite the fact that many scholars have pointed to a significant relationship between drug use and

crime, little is known about how drug-related crime patterns are associated geographically with drug markets: while those who use drugs may also commit crime, how are their spatial targeting decisions influenced by the location and type of drug markets? From a methodological perspective, it is necessary to apply appropriate crime data modeling techniques to examine the geography of criminal activities taking place in different locations and explore how potential drug users are positioned and move about within the city. Recent developments in the use of geographical information systems (GIS) allow the input, storage, merging, and analysis of geographical, statistical, and other types of datasets within a single file format. The high operational quality of these platforms enables the integration of many data formats and the construction of very large and detailed models. This means that it is possible, for example, to map drug dealing events in conjunction with sociodemographic statistics, data on land use, or crime statistics. This information can be aggregated according to different spatial units, such as the administrative boundary, ward unit, and building block. Remarkably, most of the research has focused on the patterns of drug dealing only at the area or grid-based unit of analysis. With few exceptions (e.g., Friedrich et al. 2009), drug markets have not been examined at the street segment level. Thus, in future work, it will be useful to conduct analyses using street segments as the unit of analysis. The advantage of detailing at such a fine scale of analysis allows more accurate data modeling – after all criminals do not navigate areas, they move along the streets. Moreover, from the detection and prevention perspective, it is quite useful to know the type of street segments that have a high probability of being or becoming drug dealing places and require more police attention.

Recent advances in computational methods allow very detailed modeling and analysis of data which should facilitate this sort of approach. For example, a group of researchers from Japan used a GIS platform to develop the Spatial Analysis along Networks (SANET) software. This provides a set of statistical tools that facilitate

the analysis of point data with reference to the street network. The network statistics used represent a modification of numerous traditional spatial but also graph theoretical techniques. For instance, the software enables the analysis of the density of crime incidents across street segments, the shortest path between points, and the enumeration of the fraction of clustered crime incidents and so on.

The morphology of street network per se is another factor that can be informative in understanding and analyzing crime incidents within a city. Space Syntax theory (Hillier and Hanson 1984) and research discuss how the spatial properties of the street network influence the distribution of pedestrian movement patterns. Space Syntax has developed analytical tools, which using computer modeling (Depthmap software; Turner 2001) provide graphic (and quantitative) representations of the probabilities of pedestrian and traffic movement along the street segments. Namely, it estimates the accessibility levels for every street segment in relation both to its immediate surroundings (local scale) and to the whole city network (global scale). The quantifiable differences in the level of movement accessibility between different locations enable testing statistically why particular locations are prone to crime.

In conclusion, it should be emphasized that illegal drug markets are difficult and multidimensional problems involving many social and spatial factors. Mechanisms of illicit retail may vary from country to country, and so the theoretical frameworks described in this entry require further testing to determine the extent to which they do and do not apply in other cultural contexts. Methodological and computational advances offer the opportunity to do so and will likely provide new insight into patterns of drug crime.

Related Entries

- ▶ [Crime Location Choice](#)
- ▶ [History of Geographic Criminology Part I: Nineteenth Century](#)

- ▶ [Rational Choice Theory](#)
- ▶ [Routine Activities Approach](#)
- ▶ [Spatial Models and Network Analysis](#)

Recommended Reading and References

- Anderson E (1999) *Code of the street: decency, violence, and the moral life of the inner city*. W.W. Norton, New York
- Brantingham PL, Brantingham PJ (1984) Burglar mobility and crime prevention planning. In: Clarke R, Hope T (eds) *Coping with burglary*. Kluwer-Nijhoff, Boston
- Brantingham PL, Brantingham PJ (1993) Environment, routine, and situation: toward a pattern theory of crime. In: Clarke RV, Felson M (eds) *Routine activity and rational choice*. Transaction Publishers, New Brunswick, pp 259–294
- Brantingham PL, Brantingham PJ (1995) Criminality of place: crime generators and crime attractors. *Eur J Crim Policy Res* 3:5–26
- Clarke RV, Weisburd D (1994) Diffusion of crime control benefits: observations on the reverse of displacement. *Crime Prev Stud* 2:165–184
- Cohen LE, Felson M (1979) Social change and crime rate trends: a routine activity approach. *Am Sociol Rev* 44:588–608
- Cohen J, Tita GE (1999) Diffusion in homicide: exploring a general method for detecting spatial diffusion processes. *J Quant Criminol* 15:451–493
- Curtis R, Wendel T (2000) Toward the development of a typology of illegal drug markets. In: Hough M, Natarajan M (eds) *Illegal drug markets: from research to policy*. Criminal Justice Press, Monsey, pp 121–152
- Eck J (1994) *Drug markets and drug places: a Case-control study of the spatial structure of illicit drug dealing*. Dissertation, University of Maryland, Ann Arbor, University Microfilms International
- Eck J (1995) A general model of the geography of illicit retail market places. In: Eck J, Weisburd D (eds) *Crime and place*. Criminal Justice Press, Monsey
- Eck JE, Maguire ER (2000) Have changes in policing reduced violent crime? An assessment of the evidence. In: Blumstein A, Wallman J (eds) *The crime drop in America*. Cambridge University Press, Cambridge, pp 207–265
- Eck J, Wartell J (1996) *Reducing crime and drug dealing by improving place management: a randomized experiment: report to the San Diego Police Department*. Crime Control Institute, Washington, DC
- Friedrich E, Hillier B, Chiaradia A (2009) Anti-social behaviour and urban configuration using space syntax to understand spatial patterns of socio-environmental disorder. In: Koch D, Marcus L, Steen J (eds) (Proceedings) 7th international space syntax symposium, Royal Institute of Technology (KTH), Stockholm
- Goldstein PJ (1985) The drugs/violence nexus: a tripartite conceptual framework. *J Drug Issues* 15:493–506

- Hillier B, Hanson J (1984) *The social logic of space*. Cambridge University Press, Cambridge
- Hillier B, Penn A, Hanson J, Grajewski T, Xu J (1993) Natural movement or, configuration and attraction in urban pedestrian movement. *Environ Planning (B): Planning Design* 20(1):29–66
- Levitt SD, Venkatesh SA (2000) An economic analysis of a drug-selling gang's finances. *The Quarterly Journal of Economics* 115:755–89
- MacCoun E, Kilmer B, Reuter P (2003) Research on drugs-crime linkages: the next generation. Toward a drugs and crime research agenda for the 21st century. Retrieved September 2003 from: <http://www.ojp.usdoj.gov/nij/drugscrime/194616.htm>
- Mazerolle L, Kadleck C, Roehl J (1998) Controlling drug and disorder problems: the role of place managers. *Criminology* 36(2):371–402
- McCord E, Ratcliffe J (2007) A micro-spatial analysis of the demographic and criminogenic environment of drug markets in Philadelphia. *Aust N Z J Criminol* 40:43–63
- Okabe A and Okunuki K and SANET Team SANET a spatial analysis along networks (Ver.4.0). Tokyo, Japan
- Pettit LE (1995) Copping crack: the travel behaviour of crack users. *Justice Q* 12(3):499–524
- Rengert G (1996) *The geography of illegal drugs*. West View Press, Boulder, Colorado
- Rengert G, Wasilchik (1989) Space, time and crime: ethnographic 87 drug sales, burglary, and neighborhoods insights into residential burglary. Final Report, National Institute of Justice, Grant # 88-IJ-CX-0013.
- Rengert G, Ratcliffe J, Chakravorty S (2005) Policing illegal drug markets: geographic approaches to crime reduction. Criminal Justice Press, Monsey
- Repetto TA (1974) *Residential crime*. Ballinger, Cambridge
- Reuter P, MacCoun R (1993) *Street drug markets in inner-city neighbourhoods*. Rand Corporation, Santa Monica
- Roncek D, LoBosco A (1983) The effect of high schools on crime in their neighborhoods. *Soc Sci Quart* 64:598–613
- Roncek D, Maier P (1991) Bars, blocks, and crime revisited: linking the theory of routine activities to the empiricism of 'hot spots'. *Criminology* 29:725–754
- Taniguchi TA, Rengert GF, McCord ES (2009) Where size matters: agglomeration economies of illegal drug markets in Philadelphia. *Justice Q* 26:670–694
- Taniguchi TA, Ratcliffe J, Taylor R (2010) Gang set space, drug markets, and crime around drug corners in Camden. *J Res Crime Delinq* 48(2011):327–363
- Turner A (2001) Depthmap: a program to perform visibility graph analysis. In: Proceedings 3rd international symposium on space syntax, Georgia Institute of Technology, Atlanta, Georgia, pp 31.1–31.9
- Weisburd D, Wyckoff J, Ready J, Eck J, Hinkle C, Gajewski J (2004) Does crime just move around the corner? A controlled study of spatial displacement and diffusion of crime control benefits. *Criminology* 44:549–592

Specialization and Sexual Offending

Danielle Arlanda Harris
Justice Studies, San Jose State University,
San Jose, CA, USA

Overview

The field of research into sexual aggression and offending is largely based upon the assumption that sexual offenders are inherently different from generic or nonsexual criminals. This position has permeated through more than just research and is now evident in the policies and practices that treat and manage sexual offenders and seek to prevent sexual abuse. This assumption of offense specialization proposes that sexual offenders commit sexual offenses persistently and exclusively (or at least predominantly) throughout their lives (Simon 1997; Zimring et al. 2007). It is often further implied that these offenders have a detectable degree of sexually deviant arousal and commit sexual offenses in a planned and repetitive way (Simon 2000; Zimring et al. 2007). Meanwhile, almost no other type of offender is viewed like this. Rather, they are considered to be criminally versatile, showing no apparent predilection for one crime over another.

This entry provides an overview of recent international literature regarding the nature and extent of offense specialization and versatility in the criminal careers of sexual offenders. A discussion of the many methodological considerations that arise from this research will be presented, followed by a brief overview of relevant theoretical perspectives. Finally, comments are provided regarding the implications for further research in criminology and criminal justice as well as prevention efforts aimed toward addressing sexual offenders.

Introduction

There are wide-reaching practical implications that stem from the view that sexual offenders

are either specialist or versatile in their offending. The sexual offender “profile” has become increasingly visible in recent years. This has likely been encouraged by a small number of horrific incidents that received extraordinary attention by the international media. As a consequence, individuals who are identified as sexual offenders are now subjected to more discretionary decision making in the criminal justice system than almost any other type of offender (Simon 1997). This has led to more frequent and aggressive prosecution of sexual offenders as well as the passage of various special legislative initiatives aimed exclusively at sexual offenders (Simon 1997). Judgments about an individual’s dangerousness and their perceived risk of reoffense are predicated upon the view that sexual offenders are “special cases” requiring or necessitating specific assessment, treatment, therapist certification, polygraphy protocols, probation models, protective custody, and so on.

If offenders do specialize in the types of crime they commit, then it makes sense for policy to be tailored toward specific individuals who have been determined likely to participate in a particular crime (Mazerolle et al. 2000). In this way, crime-specific incarceration policies would be influential in reducing the incidence of specific crime types. If knowledge of prior offenses could actually predict subsequent criminal events, then the identification of specialized offenders would certainly be useful in criminal justice decision making (Farrington et al. 1988).

If, on the other hand, offenders are predominantly versatile, crime-specific policies would have less impact. If versatility is the standard offending tendency, a reevaluation of many legislative initiatives and policies already in place might be necessary. One interpretation of this state of affairs would suggest that such initiatives, “fuelled by the assumption of specialization” (Simon 1997, p. 2), have evolved to cater to a sexual offender population that does not actually exist.

Competing Perspectives

The Specialist Perspective: Offense specialization is a central component of the criminal career

paradigm and refers to an individual’s tendency to engage in the same crime or type of crime on successive arrests or offending occasions (Blumstein et al. 1986). Specialist offenders are thought to become proficient at a certain crime, which they subsequently commit almost exclusively (Harris 2008). The consequence of this definition is that it is thought to be possible to predict any particular arrest in someone’s criminal career based on their most recent prior arrest (Britt 1994).

Compelling evidence of versatility first emerged in the 1920s (Harris 2008) which left the notion of specialization largely marginalized for decades. Despite this, books continue to be written, research projects funded, and theories constructed, all with an assumption that offenders specialize in their offending (Gottfredson and Hirschi 1990). The strong belief in specialist offenders – shoplifters, drug offenders, and white-collar criminals – conveniently allows “us” to separate ourselves from “them.” Nowhere is this trend more obvious than in cases of rapists or pedophiles.

The Versatile Perspective: The opposing position of offense versatility proposes a common construct of deviance that views all crimes as “acts of force or fraud undertaken in pursuit of self-interest” (Gottfredson and Hirschi 1990, p. 15). Gottfredson and Hirschi (1990) claim that all crimes are alike and argue that any distinction between them (such as trivial or serious, or victim or victimless) is irrelevant and misleading. It would follow then that distinguishing between sexual and nonsexual offending (or, e.g., between sexual offenses against adults and sexual offenses against children) is an equally futile endeavor.

Specialization and Versatility in Sexual Offending: Theoretical Context

Although offense specialization is probably best understood empirically, it is necessary to situate offending patterns within a meaningful theoretical context. Theory building in sexual offending research has largely occurred beyond the

boundaries of traditional criminology (Simon 2000). Meanwhile, criminological theories tend not to consider sexual offenders on their own. Some have observed that these developments have created an almost philosophical apartheid between conventional criminology and sexual offending theory (Soothill et al. 2000). There are, however, appropriate explanations that attend to both of these observations. The following discussion will focus explicitly on those theories of sexual offending that speak to specialization or versatility.

Sexual offenses are usually understood within the context of deviant sexual arousal and paraphilia. But a more generic model would consider a sexual offense within the context of general criminality. Here, a sexual crime would be seen as an expression of general antisocial behavior that likely occurs rarely and in a haphazard fashion within an offender's criminal career. Some researchers have concluded that generic models provide sufficient explanation of sexual offending, while others hold that sexual offenders warrant more specific theoretical consideration.

The propensity to offend can certainly be understood by either general or specialist theories. In particular, sexual offending can be understood as either an example of general crime or as a special case requiring the presence of sexual deviance. The Integrated Theory of Sexual Offending (Marshall and Barbaree 1990) and the Conditioning Theory of Sexual Offending (Laws and Marshall 1990), respectively, are proposed as two such examples.

The Integrated Theory of Sexual Offending. Marshall and Barbaree's (1990) Integrated Theory of Sexual Offending focuses on opportunity and self-control and can be seen as a sexual offending-specific version of Gottfredson and Hirschi's (1990) General Theory published in the same year. Marshall and Barbaree (1990) combined developmental elements with situational and environmental factors to create an integrated explanation of sexual offending (Ward 2002). The authors proposed that men are biologically equipped to learn to express their sexuality aggressively. They contend that "the acquisition of attitudes and behaviors during

childhood sets the stage for the developing male to respond to the sudden onset of strong desires characteristic of pubescence with a prosocial or an antisocial mental set" (Marshall and Barbaree 1990, p. 260). Similar to Gottfredson and Hirschi (1990), they illustrate the importance of parental attachment and argue that sexual offending is more likely in individuals whose childhoods are characterized by poor parenting or poor socialization (Ward 2002). Provided these preexisting variables are present, situational elements have the ultimate impact on sexual offending. Stress, intoxication, anger, presence of a potential victim, and the belief that detection can be avoided are the most important situational variables (Marshall and Barbaree 1990; Ward 2002).

Marshall and Barbaree's (1990) contribution does not address the issue of specialization versus versatility directly. For them, sexual offending remains independent of other behaviors and therefore warrants a unique explanation. They recognize that sexual offenders have the capacity to be criminally versatile and do engage in an array of antisocial behaviors. This represents a significant variation from other theories constructed to explain sexual offending which have essentially ignored this aspect of criminality. Ward et al. (2006) suggest that Marshall and Barbaree altered the theoretical landscape (within sexual abuse research) by simply acknowledging the component of nonsexual criminality.

Like the critics of Gottfredson and Hirschi (1990), researchers have also argued that the Integrated Theory is too general (Ward 2002). Their main criticism is that it does not account for important distinctions between types of sexual offenders (i.e., distinguishing, e.g., between rapists and child molesters) (Ward 2002). Drawing on typological traditions, Ward recommends that theorists should "develop more specific rather than general explanations" (p. 217). Evidently, influenced by the work of Gottfredson and Hirschi, the criminological tradition has moved in the opposite direction.

Conditioning Theory of Sexual Offending. Also published in 1990, Laws and Marshall's Conditioning Theory provides a coherent

explanation of specialization in sexual offending by emphasizing learning experiences and motivation. Like other well-regarded theories of sexual abuse, Laws and Marshall (1990) do not preclude a sexual offender from being versatile. Instead, they simply do not make any reference to a sexual offender's nonsexual offending. In this way, they seek to explain only how an individual with certain experiences will come to offend sexually.

Conditioning Theory provides a theoretical basis for the deviant sexual preference hypothesis (Ward et al. 2006). Although it has a specific sexual offender focus, the theory argues that sexual deviations are learned responses to possibly accidental experiences with sexually deviant behavior (Ward et al. 2006). That is, they could evolve in anyone.

Deviant arousal patterns are said to develop through the process of classical conditioning and the "elaborated use of deviant fantasy in masturbation" (Laws and Marshall 1990, p. 226). This is likely a gradual process occurring during masturbation to a memory which need not have been sexually stimulating at the time of the initial experience (Ward et al. 2006). Classical conditioning describes the unconscious and "repetitious or traumatic pairing of sexuality and some negative experience [which] produces some type of intensive emotional response that distorts subsequent sexual gratification" (Schwartz and Cellini 1996, p. 14). This explains the unconscious compulsion to reenact one's own abuse to gain mastery over the experience. This has been supported empirically in samples of sexual offenders, particularly for juveniles.

A Theory of Specialization and Versatility. At this point, two compelling empirical realities must be considered and reconciled. First, it is known that not all sexual offenders present with deviant sexual preferences or deviant sexual arousal (Simon 2000; Weinrott and Saylor 1991). Second, it is known that many people who commit sexual offenses also engage in a considerable amount of nonsexual criminal behavior in addition to their sexual offending (Lussier 2005; Simon 2000; Weinrott and Saylor 1991). What is missing from these perspectives is a sensible explanation of

two realities: sexual offending by men without deviant sexual interests and nonsexual offending committed by men who have explicitly been designated "sexual offenders."

Harris et al. (2009a) concluded that the components of Gottfredson and Hirschi's (1990) theory were appropriate for almost all (89 %) of the 506 sexual offenders in their sample. They indeed revealed the analogous behaviors that were predicted by the General Theory of Crime. At the same time, it still appeared that a small group of highly specialized offenders warranted the more sexually specific explanations such as those provided above (Marshall and Barbaree 1990; Laws and Marshall 1990). That is, for the remaining 11 % of the sample it was relevant to focus on such factors as sexual preoccupation; emotional congruence with children; and having a male, unrelated, or stranger victim. Essentially, the authors suggested that some aspects of sexual offending could be understood akin to the way Moffitt (1993) explained delinquency.

Borrowing Moffitt's (1993) framework amounts to a "Dual Taxonomy of Sexual Offending" where two distinct trajectories of offending are proposed. Thus, the observation of sexual offenses is understood to conceal two qualitatively distinct types of offending behavior: sexual offenses that are committed rarely or sporadically as part of the broader range of criminal activity in which an individual engages and sexual offenses that are committed by highly specialized offenders who seldom, if ever, engage in any other rule-breaking behavior.

The first category of offenders were referred to as "versatile" and made up the vast majority of offenders. They were characterized by general indicators of low self-control including substance abuse, adolescent antisocial behavior, elementary school problems, and unemployment (Harris et al. 2009a). They were more criminally versatile, engaged in a broader range of different criminal behaviors, and were more likely to offend upon release. This is in line with the results of a previous study (Meloy 2005) which concluded that the majority of sexual offenders could be described as situational or opportunistic, with versatile criminal histories.

The second category was “highly specialized” and represented 11 % of offenders, comparable in size to Moffitt’s life-course-persistents or Wolfgang’s chronic offenders. Contrary to their versatile counterparts, this group consisted mostly of child molesters and was characterized by sexual preoccupation, emotional congruence with children, and a greater likelihood of abusing known, related, and male victims (Harris et al. 2009a). Highly specialized sexual offenders were also statistically significantly more likely than versatile offenders to begin their offending in adulthood. This group best resembled the expectations of the theoretical perspectives that expect specialization (Laws and Marshall 1990; Marshall and Barbaree 1990).

It is definitely conceivable that more groups could be identified, particularly given the sample used in the study (Harris et al. 2009a) which is arguably biased toward sexual offenders in the direction of specialization. The group that is most notably absent in that study are offenders who have *not* been convicted of a sexual offense. Although the studies reviewed for this entry have demonstrated considerable similarities between convicted sexual offenders and the theoretical expectations of more generic nonsexual offenders identified in the literature, the opportunity certainly remains to assess this similarity empirically by including a control group of nonsexual offenders. This represents a worthwhile area of future inquiry.

Methodological Considerations in Specialization Research

Extant research on offense specialization is confounded by methodological inconsistencies which has an impact on current understanding of specialization. The most common concerns include the operational definitions of specialization and inconsistencies in the number and type of crime category classifications that are used (McGloin et al. 2007). Within the field of research on sexual offending, the added question of offender classification also impedes one’s ability to compare results across studies..

Definition of Specialization. Many different techniques exist to define and measure specialization. Although each one has limitations when used alone, a combination of measures can be used together to paint a clearer picture of specialization. The simplest measure, the Specialization Threshold (ST), is a crude cutoff that detects specialization if an arbitrary proportion of one’s arrests is for a particular offense or category of offenses (Cohen 1986; Harris et al. 2009b). Various thresholds have been used in prior research but Cohen’s (1986) definition is used most often and declares specialization if 50 % of a person’s arrests are for a particular type of crime. As the most simplistic measurement of specialization, the ST’s main challenge is its inability to capture change or crime switching over time. The Forward Specialization Coefficient (FSC) (Farrington et al. 1988) and Diversity Index (DI) attend to that limitation by taking into account transitions between offenses (Sullivan et al. 2006). These models generate elegant coefficients that can be compared across studies. Although the FSC relies on consecutive transitions only, the DI can be calculated on nonadjacent transitions as well (McGloin et al. 2007). For some researchers, the DI has been preferred over the FSC for its more intuitive application and interpretation (Harris 2008).

Number of Crime Categories. It is difficult to compare results across studies because researchers seldom apply the same or similar crime classification schemes. Although there is no consensus on which to use (Sullivan et al. 2006), the convention in criminology has been three or four offending categories: personal/violence, property, drugs, and other (where “drugs” and “other” are sometimes combined) (Mazerolle et al. 2000). Mazerolle and colleagues showed that the outcomes of using three or four categories are substantively similar and opted for the parsimonious model of combining “drugs” and “other.” So far, the emerging research on sexual offending has further integrated “personal” offenses into sexual violence and nonsexual violence.

Use of Official Data. The limitations of studying crime using official statistics have been

discussed elsewhere (Farrington et al. 1988; Rice et al. 2006). In addition to the actual nature of their behavior and whether they got caught, an individual's criminal record can also be influenced by a victim's willingness to report an offense, local law enforcement practices, and sentencing policies (Harris 2008). Further, in the case of sexual offending, official statistics tend to underestimate the "severity of sexually motivated violent offenses" (Rice et al. 2006, p. 525). Procedures such as plea bargaining often mean that the actual charges brought against a person no longer represent the sexual nature of the crime (e.g., rape being pled down to assault). Despite these disadvantages, official statistics remain the most common source of data for studies of specialization, but their subsequent results should be considered with these limitations in mind. A combination of official records and self-reported data is recommended.

Number of Previous Criminal Events. The concept of specialization implies the presence of another important element of the criminal career paradigm: persistence. Because it is meaningless to discuss specialization in one or two offenses, specialization research is often restricted to frequent and persistent offenders or "career criminals." Of course, analyzing only those offenders with more than five arrests, for example, greatly reduces one's sample of offenders. In addition to reduced statistical power, this introduces a possible sample bias by including only sufficiently persistent offenders. Although some researchers have limited their analysis to participants with at least five offenses or nine offenses, the general convention is to include participants with two or more (Farrington et al. 1988; Harris 2008; Harris et al. 2009b; Sullivan et al. 2006) or three or more separate arrests or sentencing occasions.

Use of Arrest or Conviction. Specialization studies either use arrests, convictions, or sentencing occasions as the primary unit of analysis. Because each occasion might include multiple counts or charges, the convention is usually to record only the most serious charge at each event (Farrington et al. 1988). Using charges will hide arrests that failed to result in

a charge and using sentencing occasions will hide convictions that failed to result in a sentence. The seriousness of most sexual offenses (compared with almost all other crime) means that recording only the most serious offense would conceal other co-occurring behaviors, thus inflating specialization. Although the opposite practice (of including every individual charge) has the complementary limitation of inflating versatility, given that the dark figure of sexual crime is so substantial, this method is considered the most appropriate for samples of sexual offenders (Harris 2008).

Classifying Sexual Offenders. Specialization is usually taken to mean the likelihood of committing the same offense on a subsequent occasion. Blumstein et al. (1986) notion of "offense clusters" broadened this definition so that rather than only detecting specialization in burglars who subsequently commit burglary, for example, (Britt 1994), an offense cluster of theft could be used to encompass such crimes as burglary, larceny, shoplifting, or fraud (Britt 1994). This approach is important for the present discussion because it is argued that some sexual offenders specialize *within* sexual offense categories (Soothill et al. 2000).

Consider a male child molester who is subsequently convicted for rape of a woman. Blumstein et al. (1986) would consider him a specialist sexual offender. A clinician, however, would consider the same individual quite differently from an exclusive child molester (who, e.g., might demonstrate the more typical characteristics of emotional congruence with children and difficulty in maintaining adult relationships). To Blumstein et al., he would be a "sexual offender," but to a clinician, although his offenses might all be sexual, he apparently shows no predilection for certain types of victims and might be considered a more challenging candidate for treatment. A mixed offending pattern (that includes both adult and child or male and female victims) might also indicate that he is potentially more likely to reoffend upon release, because his sexual crimes are unpredictable and his victim preferences are unclear. These characteristics are also those that would render an

individual at higher risk of reoffending (according to many risk assessment tools) than his more specialized counterparts.

Empirical Evidence of Specialization and Versatility in Sexual Offenders

Almost a quarter century of research on offense specialization is summarized by the findings of one of the first articles on the topic: there is a “small but significant degree of specialization in offending superimposed on a great deal of versatility” (Farrington et al. 1988, p. 461). Regardless of method or sample, there is a clear indication of versatility among offenders within the criminological literature. A modest amount of specialization is detected in some samples, and this is most notably observed in persistent offenders who use drugs.

Despite clinical assumptions to the contrary, research that has focused specifically on sexual offenders over the last 10 years has drawn similar conclusions: many sexual offenders do not restrict their criminal activities to sexual offenses (Lussier 2005; Miethe et al. 2006; Simon 2000; Smallbone and Wortley 2004; Sothill et al. 2000; Weinrott and Saylor 1991; Zimring et al. 2007). Further, notwithstanding possible differences in reporting, arrest, and conviction rates for sexual and nonsexual offenses, adult sexual offenders are about twice as likely to be convicted for nonsexual offenses as they are for sexual offenses, both before and after being convicted of a sexual offense (Smallbone and Wortley 2004). The following section details the findings of a selection of specific studies in this area.

As early as 1991, Weinrott and Saylor (1991) concluded that specialization in sexual crimes was relatively rare. The 99 convicted sexual offenders in their study collectively self-reported nearly 20,000 nonsexual crimes in just the 12 months prior to their incarceration. The majority of these offenses were minor and mostly included public intoxication and petty theft.

Smallbone et al. (2003) found substantial offending versatility in the official histories of

the 88 incarcerated adult male sexual offenders in their sample. The participants in their study included 33 rapists, 29 extrafamilial child molesters, and 26 incest offenders. Almost two thirds (60 %) of their total sample and 88 % of the offenders who recidivated reported having prior nonsexual criminal convictions.

Smallbone and Wortley (2004) later assessed the extent of persistence and versatility in the self-reported offending behaviors of 207 child molesters. Their sample was broken down further by victim relationship with 98 participants having committed incest exclusively, 72 having extrafamilial victims, and 37 having both intrafamilial and extrafamilial victims. Results from an extensive self-report questionnaire revealed considerable versatility with 69 % of the sample having been convicted of at least one nonsexual offense and 80 % having first been convicted for a nonsexual crime.

In an ambitious examination of almost 10,000 sexual offenders and some 24,000 nonsexual offenders released from prisons in 15 US states in 1994, Miethe et al. (2006) also found low levels of persistence and specialization among sexual offenders compared to nonsexual offenders. Sexual offenders had fewer arrests of all kinds than any other offender group except murderers. Although 70 % of the sexual offender sample had been arrested only once, some 95 % of those who had been arrested more than once had nonsexual offenses in their arrest histories.

Harris et al. (2009b) explored offense specialization in the officially recorded criminal histories of 506 sexual offenders referred for civil commitment in Massachusetts. They examined the sample by offense type (comparing rapists, child molesters, incest offenders, and offenders with mix-aged victims) and by referral status (comparing offenders who were observed and released with those who were observed and committed for treatment). Contrary to their hypotheses, their results indicated that treated participants were no more likely to specialize in sexual offenses than observed participants. When compared by offense type, as expected, child molesters were substantially more likely than rapists to specialize in sexual offenses. Child

molesters were also more likely to specialize in sexual offenses against children (than rapists were to specialize in sexual offenses against adults). As a whole, the sample was also more likely to specialize in sexual offenses than in the other offending categories of nonsexual violence, property, and other offenses.

Later, Harris et al. (2011) examined the post-release offending records of the same sample of 506 observed and committed offenders for a period of up to 10 years. Again, they found that the tendency to be criminally versatile remained the most compelling offending pattern across the sample. Few differences were detected between treated and observed participants, and almost no differences were found between rapists and child molesters with respect to recidivism or to post-release specialization. Rapists were more likely than child molesters to reoffend at all and to reoffend violently. There were no differences between groups on their likelihood of committing sexual offenses after release.

Interestingly, specialization prior to incarceration was not indicative of specialization in sexual offending after release. Of the 53 participants who specialized after release, approximately half had specialized criminal histories ($n = 27$) and half did not ($n = 26$) (Harris et al. 2011). The authors concluded that the predictive power of a sexually specialized criminal history was limited.

Harris and Knight (forthcoming) elaborated upon the findings from the previous two studies by using an addition of almost 300 more cases and attempting to establish the existence of an identifiable group of highly specialized sexual offenders that may have been masked by the liberal definition of specialization (50 %) used in prior works. A small group of 166 highly specialized sexual offenders existed out of a larger group of 748 offenders and was made up mostly of child molesters. This was consistent with both existing empirical evidence and the findings from Harris et al. (2009b) which concluded that child molesters were the most likely offender type to specialize in sexual offending.

A deeper analysis of the sample's responses to various items from risk assessment tools

demonstrated that the majority of the sample shared many of the general criminal attributes proposed by conventional criminology including the presence of analogous behaviors such as alcohol and drug use, problems in elementary school and adolescent behavior, and involvement in a range of different types of crimes. By comparison, the small group of highly specialized offenders was characterized instead by the variables identified by more specific theories of sexual offending. This study provided added empirical support for the Dual Taxonomy of Sexual Offending described earlier.

Specialization Within Sexual Offenses. As mentioned earlier, sexual abuse researchers distinguish between offenders who have sexually assaulted adults (rapists) and offenders who have sexually abused children (child molesters) (Meloy 2005; Simon 1997). Further distinctions are also made between familial and nonfamilial child molesters. A review of recent studies that have addressed the question of specialization *within* sexual offenses is provided here.

Rapists are repeatedly found to resemble violent nonsexual offenders most closely (Harris 2008). They are thought to be predominantly versatile in their offending and the commission of rape is seen to be part of a broader propensity to act in a generally antisocial manner (Lussier et al. 2005; Smallbone et al. 2003). Each of the 37 convicted rapists in Weinrott and Saylor's (1991) sample admitted to at least one nonsexual offense. Further, over half of the sample committed burglary and almost half committed robbery (Weinrott and Saylor 1991). The mean number of different offenses (out of a possible list of 22) that was self-reported by each rapist was 10.5 (Weinrott and Saylor 1991). Further, a third of convicted rapists admitted to having had sexual contact with a child (Weinrott and Saylor 1991).

Lussier et al. (2005) found that when compared with child molesters, rapists engaged in a larger variety of all crimes, had a higher frequency of property and violent crimes, and had also committed a "significantly greater variety of violent crimes than of sexual crimes" (p. 180). Addressing specialization using recidivism statistics, Langan and Levin (2002) found that 3

years after incarceration, drug offenders (41 %), larcenists (34 %), burglars (23 %), nonsexually violent offenders (23 %), defrauders (19 %), and robbers (13 %) were all much more likely than rapists (2 %) to be rearrested for the same type of crime for which they were originally convicted.

Child molesters, by comparison, are generally considered much less versatile than rapists (Harris 2008; Harris et al. 2009b). In fact, if a specialist, persistent offender exists, the stereotypical extrafamilial child molester is likely the most convincing example. Some evidence of crime switching has been detected in child molesters. For example, 12 % of the convicted child molesters in Weinrott and Saylor's (1991) sample admitted to attempting forced sex with an adult female and 34 % had abused children inside the family as well as outside the home. Lussier et al. (2005) also concluded that child molesters committed a variety of different sexual crimes. A quarter of Weinrott and Saylor's (1991) sample reported substance abuse and 20 % reported assault, theft, burglary, possession of stolen goods, and drug-related offenses in the year prior to incarceration.

Finally, incest offenders tend to most closely resemble non-offenders. When compared to rapists and child molesters, they display more social competence, having maintained steady employment and a stable marriage or marital-type relationship. Incest offenders are often thought to present with the lowest risk of reoffending, having engaged in a discrete and isolated event. But this idea has been challenged by evidence of versatility among samples of intrafamilial offenders (Harris 2008). Although a small sample, half of the 29 incest offenders in Smallbone et al. (2003) study "had at least one previous conviction for a nonsexual offense" (p. 57). All of Weinrott and Saylor's (1991) also small sample of 18 incest offenders self-reported at least one nonsexual offense. Other studies also find that incest offenders have reported committing nonsexual violent crimes and sexual crimes outside the home (Harris 2008). For example, half of Weinrott and Saylor's (1991) sample of incest offenders admitted to abusing extrafamilial children.

A Lingering Methodological Concern

In the methodological concerns raised above, the "offense cluster" of sexual offending was seen as irrelevant for sexual offenders. The research presented here, however, indicates that a substantial number of sexual offenders commit sexual offenses *outside* their assumed "specialty." The observation that identified rapists offended against children and identified child molesters against adults poses a unique problem for specialization studies that has not yet been evaluated properly. Much is made in the theoretical and treatment literature about the empirically significant and clinically valuable differences between rapists and child molesters.

The response to this conundrum is not a simple one because the distinctions between offender types are clearly valuable in practice. But, within the context of research, the considerable amount of crime switching results in the label of "rapist" becoming meaningless for a man who later admits to having child victims. Likewise, calling someone a "child molester" is similarly misleading if he has also offended against adults.

Policy Implications

Men convicted of sexual crimes (particularly against children) are now subject to specific, expansive, discretionary, and controversial legislation that has impacted almost every stage of the criminal justice system (Meloy 2005). Policies include community registration and notification, residency restrictions, mandatory specialized treatment, and civil commitment. Taken together, the studies reviewed above support existing criminological perspectives regarding versatility and challenge the assumption of specialization. By concluding that sexual offenders are largely indistinguishable from general criminals, these findings cast considerable doubt upon the value of these policies.

The focus of many sexual offender laws is to address the most chronic and persistent offenders. Evidently, specialized, persistent, and predatory sexual offenders do exist but they are an

identifiable minority. So, the laws and regulations aimed specifically at them amount to a substantial financial commitment toward policies that addresses the rarest of circumstances. Further, it is now well established that children are much more likely to be abused by those who are known or related to them. Thus, a law that notifies a community of an offenders' name and address or that restricts offenders from living within a certain distance of a school or park is fundamentally flawed. These laws have the unfortunate consequence of increasing homelessness for an already disenfranchised population and reducing employment opportunities, thus elevating the likelihood of reoffending. The considerable burden that these laws place on police, parole, and probation officers should also be taken into account.

Concluding Remarks

Research now indicates that many sexual offenders neither restrict their criminal behavior to sexual offenses nor their sexual offending behavior to specific sexual offense subtypes (Soothill et al. 2000; Weinrott and Saylor 1991). Even accounting for the bias inherent in clinical samples of sexual offenders, or the difference between official statistics and self-reported data, the ultimate message is the same: a broad trend of versatility is maintained across the studies. When compared by offender classification, child molesters are consistently found to be more likely (and rapists less likely) to specialize in sexual offenses. Although incest is predominantly seen as a distinct behavior, existing studies of these men indicate that they also offend outside the home.

Various methodological considerations were examined. While it is clear that no approach is without limitations, researchers should manage these obstacles by using multiple measures of offending, multiple definitions of specialization, and mixed methods of analysis. The construct of offense specialization versus versatility should be considered in terms of the extent both to which sexual offenders commit nonsexual offenses and

to which they commit sexual offenses outside their presumed sexual offense subtype.

Broader prevention efforts clearly need to do much more than rely on the selective policing and incapacitation of identified sexual offenders because there is now sufficient evidence to indicate that sexual offenders are generally criminally versatile. Individual offending tendencies are likely dynamic and could be influenced by a range of factors that remain underexplored. How these offending patterns emerge and perhaps change over time requires further examination. Such inquiries could in turn inform policy makers, criminal justice practitioners, and politicians alike with empirically valid evidence-based alternatives from which decisions regarding community safety could be made.

Related Entries

- ▶ [Criminal Careers](#)
- ▶ [Measuring Crime Specializations and Concentrations](#)
- ▶ [Offense Specialization: Key Theories and Methods](#)

Recommended Reading and References

- Blumstein A, Cohen J, Roth J, Visher C (1986) *Criminal careers and career criminals*. National Academy Press, Washington, DC
- Britt C (1994) Participation and frequency. In: Hirschi T, Gottfredson M (eds) *The generality of deviance*. Transaction, New Brunswick
- Cohen J (1986) Research on criminal careers: individual frequency rates and offense seriousness. In: Blumstein A, Cohen J, Roth J, Visher C (eds) *Criminal careers and career criminals*. National Academy Press, Washington, DC
- Farrington D, Snyder H, Finnegan T (1988) Specialization in juvenile court careers. *Criminology* 26(3):461–487
- Gottfredson M, Hirschi T (1990) *A general theory of crime*. Stanford University Press, Stanford
- Harris DA (2008) *Offense specialization and versatility in a sample of male sexual offenders referred for civil commitment*. Unpublished PhD dissertation, Griffith University
- Harris DA, Mazerolle P, Knight RA (2009a) Understanding male sexual offending: a comparison of general and specialist theories. *Crim Justice Behav* 36(10):1051–1069

- Harris DA, Smallbone S, Dennison S, Knight RA (2009b) Specialization and versatility in sexual offenders referred for civil commitment. *J Crim Just* 37:37–44
- Harris DA, Knight RA, Dennison S, Smallbone S (2011) Post release specialization and versatility in sexual offenders referred for civil commitment. *Sex Abuse* 23:243–259
- Langan P, Levin D (2002) Recidivism of prisoners released in 1994. US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, Washington, DC
- Laws D, Marshall W (1990) A conditioning theory of the etiology and maintenance of deviant sexual preference and behaviour. In: Marshall W, Laws D, Barbaree H (eds) *Handbook of sexual assault: issues, theories and treatment of the offender*. Plenum Press, New York
- Lussier P (2005) The criminal activity of sexual offenders in adulthood: revisiting the specialization debate. *Sex Abuse: J Res Treatment* 17(3):269–292
- Lussier P, LeBlanc M, Proulx J (2005) The generality of criminal behaviour: a confirmatory factor analysis of the criminal activity of sex offenders in adulthood. *J Crim Just* 33:177–189
- Marshall W, Barbaree H (1990) An integrated theory of the etiology of sexual offending. In: Marshall W, Laws D, Barbaree H (eds) *Handbook of sexual assault: issues, theories and treatment of the offender*. Plenum, New York
- Mazerolle P, Brame R, Paternoster R, Piquero A, Dean C (2000) Onset age, persistence, and offending versatility: comparisons across gender. *Criminology* 38(4):1143–1172
- McGloin J, Sullivan C, Piquero A, Pratt C (2007) Local life circumstances and offending specialization/versatility: comparing opportunity and propensity models. *J Res Crime Delinq* 44(3):321–346
- Meloy M (2005) The sex offender next door: an analysis of recidivism, risk factors, and deterrence of sex offenders on probation. *Crim Justice Policy Rev* 16(2):211–236
- Miethe T, Olson J, Mitchell O (2006) Specialization and persistence in the arrest histories of sex offenders: a comparative analysis of alternative measures and offense types. *J Res Crime Delinq* 43(3):204–229
- Moffitt T (1993) Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychol Rev* 100:674–701
- Rice M, Harris G, Lang C, Cormier C (2006) Violent sex offenses: how are they best measured from official records? *Law Hum Behav* 30:525–541
- Schwartz B, Cellini R (1996) The sex offender: new insights, treatment innovations and legal developments. Civic Research Institute, Kingston
- Simon L (1997) The myth of sex offender specialization: an empirical analysis. *N Engl J Crim Civ Confin* 23:387–403
- Simon L (2000) An examination of the assumptions of specialization, mental disorder, and dangerousness in sex offenders. *Behav Sci Law* 18:275–308
- Smallbone S, Wheaton J, Hourigan D (2003) Trait empathy and criminal versatility in sexual offenders. *Sex Abuse: J Res Treatment* 15(1):49–60
- Smallbone S, Wortley R (2004) Onset, persistence, and versatility of offending among adult males convicted of sexual offenses against children. *Sex Abuse: J Res Treat* 16(4):285–298
- Smallbone S, Marshall W, Wortley R (2008) Preventing child sexual abuse: evidence, policy and practice. Willan, Cullompton
- Soothill K, Francis B, Sanderson S, Ackerley E (2000) Sex offenders: specialists, generalists – or both? *Brit J Criminol* 40:56–67
- Sullivan C, McGloin J, Pratt T, Piquero A (2006) Rethinking the “norm” of offender generality: investigating specialization in the short-term. *Criminology* 44(1):199–233
- Ward T (2002) Marshall and Barbaree’s integrated theory of child sexual abuse: a critique. *Psychol Crime Law* 8:209–228
- Ward T, Polaschek D, Beech A (2006) *Theories of sexual offending*. Wiley, West Sussex
- Weinrott M, Saylor M (1991) Self-report of crimes committed by sex offenders. *J Interpers Violence* 6:286–300
- Zimring F, Piquero A, Jennings W (2007) Sexual delinquency in Racine: does early sex offending predict later sex offending in youth and young adulthood? *Criminol Public Policy* 6(3):507–534

Specialization in Juvenile Offending

Jesse Cale
Key Centre for Ethics, Law, Justice and Governance, Griffith University, Brisbane, QLD, Australia

Synonyms

[Crime repetition by youth; Patterns in youthful offending](#)

Overview

Whether and to what extent juvenile offenders specialize in committing certain types of crimes

during their delinquent careers has substantial implications for theory and policy. The question of offender specialization arguably first emerged with the advent of the positivist school in criminology in the late nineteenth century when Cesare Lombroso sought to identify criminal types. With the introduction of specialized juvenile justice in laws in the early twentieth century and their fundamental reforms in many countries in the second half of the twentieth century and into the initial decade of the twenty-first century, there have been considerable theorizing, research, and debate about specialization in youthful offending. Political debates in countries such as Australia, Canada, England-Wales, and the United States concerning the reform of youth justice laws have been influenced, in part, by the considerable research on specialization youthful offending. The main theoretical debate has been focused on the attempts to specify potential distinguishing characteristics of certain types of violent juvenile delinquents such as serious sexual offenders, murderers, and gang members, compared to less serious violent offenders, property offenders, and drug users/traffickers, for example. Many diverse labels have been applied to young offenders, such as serial fire-setter/rapist/murderer, career criminal, life-course-persistent/adolescence-limited offender, chronic/nonchronic offender, and adolescent psychopath/conduct disorder, that typically include attempts to specify types and patterns of crime. Oftentimes these labels or categories have been introduced into the media, political, and legislative circles, and debates emerge about whether to incarcerate, punish, and/or rehabilitate juvenile delinquents depending on what "type" of offender they are.

The essence of the theoretical debate is whether juvenile offenders can be classified into an "offender type" based on the qualitative nature, and/or quantity of their offending, versus whether there is more simply a single type of juvenile offender who engages in diverse and multiple antisocial and criminal behaviors. Not surprisingly, a considerable empirical research theme over the past half a century has focused on the nature and extent to which juvenile

offenders demonstrate specialization in offending. The related theoretical debate has centered on considering that if specialization is not evident, then a single or general theory of juvenile delinquency could be sufficient to explain this phenomenon. In turn, criminal justice responses, therefore, could be directed toward a central set of causes or risk/protective factors to prevent juvenile delinquency from beginning as well as reoccurring. In contrast, if research demonstrates specialization or a diversity of types of juvenile delinquent offending patterns, then typologies require far more complex and unique theoretical explanations along with specialized juvenile justice responses, including treatment and intervention strategies, depending on the type of juvenile delinquent or young offender. To better understand how this debate on specialization has evolved over the years, it is necessary to describe the key theoretical controversies historically.

Key Issues and Controversies: Theoretical Debates

Typologies of adult and youth offenders gained popularity in the 1960s with the assertion that offenders could be classified into conceptually valid types based on fundamental differences in the crimes they committed. The most basic and obvious crime-centered typologies generally consisted of nominal categories of specialized criminal behavior consisting of property versus violent offender types. More types emerged as well including white-collar criminals, organized crime, street gangs, and political terrorists, for example. It was then theorized that these criminal types were explained by different sets of causal factors.

By the 1980s, there was a wide range of theories to explain ever-increasing types of offenders. In turn, typologies and related theories generated intense debates concerning basic definitions of deviance, delinquency, and criminality that are essential to assessing youthful offending specialization. For example, the assertion that delinquency was overwhelmingly associated with

youth from low-income groups along with those belonging to certain ethnic/racial minorities was challenged by self-report studies of delinquency (e.g., Elliott and Ageton 1980). These studies demonstrated that minor delinquency was common for youth from all income and ethnic groups. A second criticism of initial crime-centered typologies was the reliance on offense records whereby the presence of a particular type of offense (e.g., sexual assault) determined a youth's "specialization" in offending. Critically, such a snapshot-based examination of offending patterns did not consider variations in offending over time. The longitudinal cohort studies of the 1970s and subsequent major cross-sectional studies provided a more valid basis for establishing typologies and specialization in youthful offending because they took into consideration changes in the frequency of offense types over the entire course of adolescence. Finally, a third key criticism involved arbitrary and income/ethnic group biases of juvenile justice laws regarding charges and convictions up to the 1970s. For example, juvenile justice laws from the early 1900s to the early 1970s virtually criminalized many childhood and adolescent antisocial and deviant behaviors. Typologies ranged from the enormous diversity of trivial antisocial behaviors such as status offenses (e.g., swearing and drinking alcohol) to the most violent crimes (e.g., rape and murder). However, because juvenile justice laws governed the ages between 8 and 21 depending on the country and within-country jurisdiction (e.g., state, province, region), it was abundantly clear that attempts to assess the nature and extent of specialization in juvenile offending had to be based on more sophisticated approaches.

What emerged from both longitudinal cohort and major cross-sectional research was that simple typologies based on the presence of certain offenses in a juvenile's history were inadequate for assessing specialization for several reasons. For example, status, property and violent crimes committed by juveniles varied substantially in not only degree of seriousness but frequency. To assess specialization, theoretically driven and complex equation-based indices often

consisting of ratios of different types of crime such as property compared to violent crimes became central to investigating specialization in youthful offending. Seriousness of types of crimes, for example, based on the extent or amount of damage in dollars involved with property crimes, and the nature and extent of injuries for violent crimes also were incorporated into indices determining offender types. Another dimension was the variation in the extent of switching to different types of crime in certain developmental periods, such as childhood, adolescence, and adulthood. In other words, early crime-centered typologies were unable to address rapidly emerging evidence that delinquent careers were characterized by significant heterogeneity involving different patterns in the type, seriousness, and frequency of offenses committed over childhood and adolescence.

Several of the key concepts from the criminal career paradigm, particularly the concept of the career criminal, evoked intense theoretical and empirical debate. Klein (Klein and Malcolm 1979; Klein 1984) maintained that an abundance of empirical evidence demonstrated that juvenile offending was predominately unpatterned and reflected the notion of "cafeteria-style" offending. Subsequently, Gottfredson and Hirschi (1990) introduced their General Theory of Crime based on Hirschi's (1969) original research and his Bonding Theory of social control. This theory asserts that specialization does not exist and instead delinquents engage in a wide range of delinquent behaviors, stemming broken bonds to family and other societal institutions such as the church, school, and community. The absence of strong prosocial bonds further explains the time-stable personality trait of low self-control, which in turn is caused by lax and/or inconsistent parenting practices. Low self-control is established in childhood and remains stable throughout life. It is manifested by the pursuit of easy and instant gratifications inherent in deviant behaviors such as smoking, excessive drinking, short-term and multiple sexual conquests, recklessness, and crime, when opportunities are present. In effect, the General Theory of Crime

predicts that juvenile offending is best understood as involving substantial offense versatility.

By the 1990s the developmental theoretical perspective of crime introduced more complex predictions about patterns in juvenile offending such as specialization and versatility. This perspective identified risk and protective factors for several forms of antisocial behavior throughout the life course beginning with early childhood and into early adulthood. Loeber (1990) asserted that children who exhibited recklessness, problems with authority, aggression, and coercive/manipulative behaviors were more likely to follow a more violent pattern in adolescence but also engage in diverse crimes. In contrast, children who were more withdrawn and antisocial were more likely to engage in covert patterns and drug use as adolescents.

The most prominent developmental theory relevant to specialization involved Moffitt's (1993) developmental taxonomy. She predicted that more sustained aggressive and violent life-course patterns were based on the age of onset of offending. Moffitt proposed two distinct types of offenders in the population: 'life-course-persistent' (LCP) offenders and "adolescence-limited" (AL) offenders. LCP offenders are characterized by the interaction, and accumulation, of individual deficits and environmental adversities that begin early in childhood and cascade over subsequent developmental periods. In effect, these individuals begin offending early in childhood/adolescence, engage in a wide range of deviant and criminal acts including bullying and other violent behavior, and persist into adulthood. In adolescence, therefore, this developmental group was predicted to exhibit frequent, versatile, and serious offending including violence. In contrast, AL offenders are predominately characterized by risk factors that emerge in adolescence. The primary risks involve status frustration driven by the expanding maturity gap between biological and social maturities (i.e., the observation that while adolescents experience physical or biological maturity at earlier ages when compared to the early twentieth century, their social maturity as evidenced by the age they assume adult social roles such as husband, breadwinner, father is

delayed), as well as exposure to deviant or delinquent peers and role models. AL offending was predicted to be more specialized than LCP offending, in the context of offenses reflecting the effect of the maturity gap such as vandalism, public order, substance abuse, and status offenses.

As with all classic debates in criminology, measurement of key constructs are often central to resolving competing hypotheses, in this instance, about specialization versus versatility in youthful offending.

The Measurement of Specialization in Youthful Offending

Historically, methodological and measurement innovations have had to take into consideration the inherent limitations associated with data sources and especially official juvenile justice information recording. Most importantly, there is a consensus that official data significantly underrepresents the total amount of crimes committed, and even more for official juvenile offending data. During the past 30 years, measures based on informal and formal diversion programs have been introduced in many Western countries. Often, informal diversions from formal justice proceedings have not been recorded. In addition, there has been concern over the variations in the offense recording and charging process across jurisdictions even within countries. When only the most serious offense is proceeded with in terms of offense recording, trial, and/or sentencing, both the range and seriousness of actual offenses are under-recorded. This affects the accuracy of officially recorded crime records. Another limitation has been the concern with variations in the number of police charges that are recorded when police practices include deliberately over charging in order to enhance their ability to gain information about other crimes and to an advantage in plea bargaining to gain a confession of guilt. As well, police and prosecutors utilize their ability to increase or decrease the seriousness of the charges since criminal codes include wide discretion concerning the

degree within a criminal offense category. For example, a violent assault resulting in death ranges from manslaughter to first-degree murder. All of these concerns raise questions over the validity of measures when using, for example, charges versus convictions in assessing specialization, especially across different juvenile justice jurisdictions included in comparative research designs. Similar validity concerns have been a source of debate with self-report methodologies utilized to assess specialization. Most importantly and especially when youth are involved, memory recall is often problematic in terms of overstating offenses, and the timelines in which they occurred in their past. Nonetheless, despite these limitations, most research on specialization in youthful offending has used a combination of methodologies to identify typologies based on a variety of specific measurement criteria. One advantage with official data is that a more complete delineation of potentially different offending trajectories over time is possible because this data can record patterns across childhood, adolescence, and adulthood. Similarly, complex cohort designs typically include multiple data recording approaches (e.g., including both official and self-report data) to measure the types, frequency, and temporal patterns of offending.

The seminal Philadelphia Birth Cohort Study (Wolfgang et al. 1972) was the most widely recognized initial methodological innovation in juvenile offender specialization research. Wolfgang and colleagues examined crime type patterns from offense to offense as a series of transitions using transition matrices. This method assessed the probability of juvenile offenders having received charges for similar crimes across adjacent offending events. Transition matrices identified whether specialization in certain crime types was evident by evaluating whether the probability of being charged with the same offense consecutively was greater in value than the probability of being charged with a different offense. Importantly, these techniques were also adopted to assess specialization using major cross-sectional research that included the complete court histories of charges and

convictions for juvenile offenders (e.g., Farrington et al. 1988).

Related to the transition matrix methodology was the development of statistical tests in the 1980s to measure whether the probability of observed transitions to similar offense types was greater than could be expected on the basis of chance alone. Coefficients representing the level of specialization for specific types of offending in a sample became a standard measure. For example, Bursik (1980) introduced the Adjusted Standardized Residual (ASR) that determined whether the ratio between observed and expected transition values was statistically significant. Farrington et al. (1988) further introduced the Forward Specialization Coefficient (FSC) to control for the influence of infrequent cases that would distort the basic observed to expected ratio. The FSC represents a single value summary of specialization for different offense types based on the sequence of similar offenses in transition matrices. Fisher and Ross (2006) more recently asserted that the FSC is the most widely utilized measure of specialization to date because of its theoretical sophistication and predictive validity.

Despite widespread use, the transition matrix approach generally, and the FSC specifically, has several limitations (e.g., see Britt 1996; Fisher and Ross 2006; LeBlanc and Frechette 1989; Nieuwebeerta et al. 2011; Sullivan et al. 2009). Most importantly, the FSC describes specialization aggregately for a group of delinquents. This aggregate measure is theoretically appropriate for the explanation of the shifting from different types of crime within groups. However, it does not measure shifts in crime for specific individuals. In effect, the aggregate nature of the FSC does not produce information about the individual extent of specialization.

Within-individual change was central to the concept of criminal careers in the 1980s (Blumstein et al. 1986) and the individual-level examination of specialization in juvenile offenders. From this perspective, the measurement focus is on the proportions of similar offense types in individual criminal histories. Here, specialization was defined by the absence of diversity of different types of crime in an

individual's offense history. Regarding a specific measurement, for example, Wikstrom (1987) calculated the total proportion of crimes belonging to a single category of offenses in the histories of offenders, which allowed for the utilization of different percentage thresholds for determining specialization. Wikstrom (1987) asserted that if, regardless of the type of offense, two-thirds or more were of one type, then the offender's criminal offending history is considered specialized. However, this measurement methodology is not applicable to occasional or limited offending. In a Swedish sample of delinquents, Wikstrom (1987) identified approximately one-third of juvenile offenders as specialists. Theoretically, different thresholds can be asserted as establishing specialization. In addition, this threshold measurement methodology raises validity concerns depending on the number of types of crimes available for analysis. For example, certain studies can and have calculated a threshold based on three categories of crime, such as property, violent, and status offenses. Others have used even more specific crime categories, making cross-study comparisons problematic because the two-thirds threshold used for three crime types does not appear an appropriate threshold for a larger number of categories. In short, determining specialization based on percentage threshold approaches is somewhat arbitrary and ultimately does not easily capture the nature of offense specialization.

More recently, the extent of specialization in an offender's history has been calculated using the diversity index which also focuses on the proportion of different offenses committed by individuals but also provides a measure of the probability that any two observed offenses drawn randomly from a criminal history belong to dissimilar offense types. In effect, by calculating the proportions of offenses across a specified number of crime categories, this approach provides an individual-level measure of specialization based on the complete offense patterns, rather than the sequence of offenses. This measure calculates the overall amount of similar/dissimilar offenses and allows for the comparison of specialization patterns for both individuals and

groups. Again, however, when studies used different numbers of crime categories to calculate the amount of diversity, cross-study comparisons become problematic. In addition, some scholars have asserted that the diversity index does not provide precise information about the type of offense specialization, such as the frequency of different offense types (Sullivan et al. 2009).

Several recent studies have applied new and innovative approaches to measure specialization in juvenile offending (and more generally). These approaches typically involve complex multivariate techniques that have the ability to simultaneously capture the nature and extent of specialization, something that past techniques typically accomplished independently (i.e., the FSC compared to the diversity index). One such approach is Latent Class Analysis (LCA). While not designed specifically for the study of offense specialization, LCA is useful for identifying homogeneous subgroups of offenders that cluster around specific types of offenses within a sample. LCA has been used in the classification of offenders' criminal career trajectories that are based on the varying frequency of total offending over time (Francis et al. 2004). In effect, this technique has the potential to distinguish groups according to both the frequency and type of offenses committed.

Research Findings

The original empirical studies using transition matrices were characterized by mixed findings because of differences in sample composition, data sources, and several methodologies. Nonetheless, most of these studies identified considerable versatility in juvenile offending behavior, which therefore, provided initial support for the General Theory of Crime and the prediction of "cafeteria-style" offending patterns of juvenile offenders (e.g., Klein and Malcolm 1979; Klein 1984). However, several of the initial studies indicated some support for specialization for certain types of minor offending, such as theft (Wolfgang et al. 1972). Subsequent studies in the 1980s found evidence of specialization,

again, for less serious crimes such as property and status offenses. For example, Farrington et al. (1988) demonstrated that the most specialized offenses involved status offenses (e.g., running away, liquor, incorrigibility, curfew, and truancy) and property offenses (e.g., burglary, vehicle theft) as well as drug use. As well, there was some indication of a modest tendency toward crime switching to more serious forms of offending, particularly from status to property and property to violent offenses (e.g., LeBlanc and Frechette 1989; Smith et al. 1984). The same pattern of mixed support for specialization was evident in other studies. For example, females were more likely to specialize in status offenses compared to males, who were comparatively more likely to specialize in serious offending (e.g., Kempf 1986, Farrington et al. 1988). Yet still, other studies did not report such differences. Even regarding ethnicity/racial groups, versatility was more prevalent than specialization despite some studies that indicated black youth were slightly more likely to specialize in more serious property offenses than white youth (e.g., Lattimore et al. 1994). Again, several studies did not report this pattern.

By the late 1990s, studies based on developmental theories' predictions concerning specialization in juvenile offending revealed even more complex findings. Mazerolle et al. (2000) examined Moffitt's typology and found AL offenders were more likely to specialize in nonserious offenses compared to LCP offenders. The latter also were more violent (see also Piquero et al. 1999). Subsequent developmental theory-based studies expanded into the adulthood stages revealing that offending trajectories over the life course confirmed the adolescence period had the least amount of specialization with specialization typically emerging later in adulthood. Most importantly, both onset and desistance occurred more often in the adulthood stages than originally predicted. In effect, as offenders aged into adulthood, their offending became more patterned but less frequent.

In sum, three main empirical themes regarding specialization in youthful offending were evident. First, adolescence generally involved the

least amount of offense specialization compared to later periods. Second, a small but significant amount of specialization occurred during adolescence. This variation however, typically consisted of the less serious end of the offending spectrum, such as status offenses and theft. However, while this small but significant amount of specialization occurred, versatility was far more prevalent. Third, a small minority of juvenile offenders was responsible for much of the versatile offending committed in adolescence and, most importantly, for more serious violent offending.

Juvenile Justice Policy Implications Related to the Specialization Debate

The original juvenile justice laws reflected early crime-specific theories and crime-centered typologies that presupposed specialization in juvenile delinquency and, in particular, status offenses. Given the apparent extent of nonserious and more specialized offending, welfare/rehabilitation models of juvenile justice were dominant and focused on the family-based causes of delinquency through rehabilitation interventions in juvenile courts. In effect, status offenses were a key target for policy intervention with youth, and in many Western jurisdictions, formal status offender programs proliferated and were based on the hypothesis that youth status offenders would escalate to more serious forms of offending if formal intervention and treatment was not pursued. In other words, one rationale for these programs was that early and formal intervention/treatment was warranted if "status offenders" represented a distinct group of youth who specialized in nonserious misbehavior to prevent them from becoming involved in delinquency and more serious offending. Proponents of this view, therefore, subscribed to the belief that youthful offending was (a) characterized by some degree of specialization and (b) escalated to more serious forms of delinquency in the absence of intervention.

By the mid-1970s, however, the welfare model-based juvenile justice laws and

assumptions about specialization in youthful offending were challenged theoretically, empirically, and politically. One reform theme was that diversion of youth outside of the formal juvenile justice system was necessary to avoid the hypothesized delinquent effects of labeling. In other words, the common view was that minor versatile offending such as status offenses and vandalism escalated into more diverse and serious offending types when juvenile justice systems labeled minor offending youth as delinquents. In effect, specialization was conceptualized more in qualitative terms such as status offenders versus serious criminal, property, and violent offenders. However, beyond the theoretical debates, research on specialization in juvenile offending in the 1970s and 1980s did not support the assertion that, in the absence of intervention, the majority of youth who specialized in status offenses escalated to more serious offending types.

The reform debates in several jurisdictions, especially Canada and the United States, focused on specialization themes involving the identification of certain types of offender categories including career criminals, chronic offenders, serious and violent offenders, and sexual offenders. In effect, the political and research debates did not just center on more contemporary themes and measures of specialization but rather on more broad conceptualizations focused on distinguishing types of minor offenders, types of property offenders, types of violent offenders, and combinations of offender types. Youth justice laws were and are still based on this conceptualization of juvenile delinquent or young offender specialization. In many jurisdictions, different sets of procedural processes are utilized depending on the above specialized categories as well as versatile types. For example, as stated above, minor property offenders and minor violent offenders in certain jurisdictions in Australia, New Zealand, Canada, the United Kingdom, and the United States are completely diverted away from juvenile justice court proceedings. In contrast, in Canada, the United States, and most other jurisdictions, the most serious violent juvenile offenders, such as murderers, are either processed in adult criminal courts or subject to adult length

custodial sentences. In addition, patterns of offending, especially predominately minor violent offenses, automatically subject such offenders to more punitive sentences. The key assumption informing many of these sentencing options is that the minor violent offending specialization pattern typically indicates the potential for such youthful offenders to escalate into specialized violent and chronic offending.

The mixed results regarding the existence of specialization in youthful offending belie the above assumptions that there is a strong correlation between chronic minor juvenile offending and long-term chronic violent offending. While there is more support for the relationship between serious chronic violent juvenile offending and long-term violent offending trajectories, only an extremely small proportion of all juvenile offenders have been identified within this specialized type. The concern is that juvenile justice laws have become over focused on this rare offender category to the detriment of programming for the overwhelming number of less serious juvenile offenders. In other words, juvenile justice reforms have been influenced by the debate in research concerning specialization, arguably, in a progressive way in some countries (e.g., Australia, New Zealand, Canada, the United Kingdom) and in a regressive manner in several US states, for example, those with life sentences with no parole and formally capital punishment for the most serious juvenile offenders. Even more recently in Canada, a conservative federal government introduced punitive reforms to the national youth justice law based on the assertion that specialized violent and specialized property offenders require more punitive sentences.

Conclusion

There is a consensus that versatility typifies most juvenile delinquents or young offenders. There also is considerable evidence that for a small proportion of juvenile offenders, certain types of offenses distinguish them from other offense patterns. For example, sexual offenses, while not constituting the largest proportion of a young

offender's offending profile, nonetheless, distinguishes such an offender from another young offender who has no sexual offenses in their offending profile. In effect, the conceptualization of offender specialization remains open to different operational definitions. The key to the different conceptualizations is the theory or policy rationales. Major theoretical insights, for example, have emerged from developmental theories of offending that have asserted certain types of specialization as discussed above. The related theoretical debates and empirical research has propelled further theorizing and research that has several benefits. First, multiple measures and indices involving increased methodological sophistication have emerged and enhanced the understanding of juvenile offending trajectories. Second, the specialization research has revealed far more complex and diverse patterns of offending than were originally hypothesized. Third, it is vitally important that sophisticated research designs continue to be employed in assessing specialization because the ongoing juvenile justice policy debates concerning how to react to different types of offenders are best informed by valid and empirically supported categories of offending as opposed to impressions and emotion regarding what constitutes serious youthful offending.

Related Entries

- ▶ [Crime Specialization, Progression, and Sequencing](#)
- ▶ [Juvenile Violence](#)
- ▶ [Measuring Crime Specializations and Concentrations](#)
- ▶ [Offense Specialization: Key Theories and Methods](#)

Recommended Reading and References

Blumstein A, Cohen J, Roth JA, Visher CA (1986) *Criminal careers and 'career criminals'*, vol 1. National Academy Press, Washington, DC

Britt CL (1996) The measurement of specialization and escalation in the criminal career: an alternative modeling strategy. *J Quant Criminol* 12:193–222

Bursik RJ (1980) The dynamics of specialization in juvenile offences. *Soc Forces* 58:851–864

Cloward R, Ohlin L (1960) *Delinquency and opportunity: a theory of delinquent gangs*. Free Press, New York

Colvin M, Pauly J (1983) A critique of criminology: toward an integrated structural-Marxist theory of delinquency production. *Am J Sociol* 89:513–551

Elliott DS, Ageton SS (1980) Reconsidering race and class differences in self-reported and official estimates of delinquency. *Am Sociol Rev* 45:95–110

Farrington DP, Snyder HN, Finnegan TA (1988) Specialization in juvenile court careers. *Criminology* 26:461–485

Fisher G, Ross S (2006) Beggarman or thief: methodological issues in offender specialization research. *Aust N Z J Criminol* 39:151–170

Francis B, Soothill K, Fligelstone R (2004) Identifying patterns and pathways of offending behaviour: a new approach to typologies of crime. *Eur J Criminol* 1:47–87

Gottfredson MR, Hirschi T (1990) *A general theory of crime*. Stanford University Press, Stanford

Hirschi T (1969) *Causes of delinquency*. University of California Press, Berkeley

Kempf K (1986) Offense specialization: does it exist? In: Cornish D, Clarke R (eds) *The reasoning criminal: rational choice perspectives on offending*. Springer, New York

Klein MW (1979) Deinstitutionalization and diversion of juvenile offenders: a litany of impediments. In: *Crime and justice*, vol 1. University of Chicago Press, Chicago, pp 145–201

Klein MW (1984) Offense specialization and versatility among juveniles. *Brit J Criminol* 24:185–194

Lattimore PK, Visher CA, Linster RL (1994) Specialisation in juvenile careers: Markov results for a California cohort. *J Quant Criminol* 10:291–316

Le Blanc M, Loeber R (1998) Developmental criminology updated. In: *Crime and justice*, vol 23. University of Chicago Press, Chicago, pp 115–198

LeBlanc M, Frechette M (1989) *Male criminal activity from childhood through youth*. Springer, New York

Loeber R (1990) Developmental and risk factors of juvenile antisocial behavior and delinquency. *Clin Psychol Rev* 10:1–41

Mazerolle P, Brame R, Paternoster R, Piquero A, Dean CW (2000) Onset age, persistence, and offending versatility: comparisons across gender. *Criminology* 38:1143–1172

Moffitt TE (1993) Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychol Rev* 100:399–420

Nieuwbeerta P, Blokland AAJ, Piquero AR, Sweeten G (2011) A life-course analysis of offence specialization across age: introducing a new method for studying individual specialization over the life course. *Crime Delinq* 57:3–28

Piquero AR, Paternoster R, Mazerolle P, Brame R, Dean CW (1999) Onset age and offence specialization. *J Res Crime Delinq* 36:275–299

- Smith DR, Smith WR, Noma E (1984) Delinquent career lines: a conceptual link between theory and juvenile offences. *Sociol Quart* 25:155–172
- Sullivan CJ, McGloin JM, Ray JV, Caudy MS (2009) Detecting specialization in offending: comparing analytic approaches. *J Quant Criminol* 5:419–441
- Wikstrom P-OH (1987) Patterns of crime in a birth cohort: age, sex, and class differences (Project metropolitan: a longitudinal study of a Stockholm cohort, no. 24). University of Stockholm, Department of Sociology, Stockholm
- Wolfgang ME, Figlio RM, Sellin T (1972) Delinquency in a birth cohort. University of Chicago Press, Chicago

Specialty Courts

- ▶ [Problem-Solving Courts](#)

Specific Deterrence

- ▶ [Rational Choice, Deterrence, and Crime: Sociological Contributions](#)

Stalking

Troy McEwan¹ and Michele Pathé²

¹Centre for Forensic Behavioural Science, Monash University and the Victorian Institute of Forensic Mental Health, Melbourne, Australia

²Queensland Forensic Mental Health Service and Griffith University, Brisbane, Australia

Synonyms

[Harassment](#), [Threat assessment](#)

Overview

Stalking affects hundreds of thousands of people around the world every year, and criminal justice and mental health professionals are often

required to manage or treat stalkers or their victims. This entry provides an overview of the phenomenon of stalking, its legal status, explanatory models, and victimology and impact. A detailed description of best-practice assessment procedures for perpetrators and victims and treatment/management considerations for both groups is also provided.

Fundamentals

What Is Stalking?

Stalking is a problem behavior characterized by repeated and unwanted intrusions inflicted by one person upon another in a manner that causes reasonable fear and/or distress. Intrusions typically come in the form of unwanted communications, such as those via telephone (including SMS), electronic media (e-mail, social networking websites), letters, or graffiti. In many stalking episodes, intrusions will also take the form of unwanted contacts, including loitering near the victim, following, maintaining surveillance, and making approaches. In addition to these core stalking behaviors, a range of other forms of harassment are often present, including property damage, theft, malicious use of the Internet (e.g., creating defamatory websites), sending or leaving unsolicited materials (ranging from gifts to items intended to frighten or implicitly threaten, such as dead animals), inveigling others to harass the victim, vexatious complaints, uttered threats, and assault.

The defining features of stalking behavior are that it is *repeated*, is *unwanted*, and *causes significant distress*. Even within this broad definition, stalking has been shown to occur in two basic patterns (Purcell et al. 2004). The first is a short-lived burst of self-limiting intrusions that occur over the course of a day or two and usually involve unwanted approaches by a stranger or distant acquaintance. The second involves both unwanted communications and contacts, and frequently other associated behaviors, tends to be perpetrated by an ex-intimate or acquaintance, and persists for weeks, if not months. The latter type tends to have a far greater psychological and

social impact on the victim and is more frequently associated with threats and physical assault, so is more likely to attract criminal justice or mental health attention. In recent years, stalking behaviors have also been observed in a purely virtual environment, a phenomenon labelled “cyberstalking.” Those who engage in cyberstalking use the mechanisms of the Internet to harass their target(s), but do not engage in the aforementioned behaviors commonly seen in “off-line” stalking. Studies of this behavior are few and bedevilled by problems with definition and methodology. The best investigation to date concluded that, while Internet-based harassment is common, pure cyberstalking is actually rare and most victims who experience stalking via the Internet are also targets of off-line stalking behavior (Sheridan and Grant 2007).

Legal Definitions

Stalking is a relatively new crime in most jurisdictions; the Californian law of 1990 was the first legislation specifically intended to prevent the behavior. By the end of the 1990s, stalking laws had been introduced in most English-language jurisdictions, with many countries in mainland Europe following suit throughout the first decade of the 2000s. The first wave of anti-stalking laws (1990–1999) usually conceptualized stalking as a precursor to assault, and so the purpose of the law was to prevent physical attack. More recently, stalking has been conceptualized as harmful in and of itself, and the introduction of the second wave of laws (those implemented since 2000) has tended to focus less on violence, although this is still a concern. Anti-stalking legislation is relatively uncommon outside of Australasia, Europe, and North America. Where present, it is usually included in domestic violence laws and so is typically poorly defined and largely restricted to female victims of ex-intimate partners.

Depending on jurisdiction, anti-stalking legislation typically requires three elements to prove the offence. The behavior or conduct element is usually defined as a “course of conduct,” or conduct engaged in on more than one occasion (intentionally broad so as to allow for early

intervention). There is no requirement that the individual instances of behavior be unlawful (e.g., repeated nonthreatening telephone calls would satisfy a course of conduct, although none would be illegal in isolation). Some jurisdictions leave the behavior element at this point without further defining the nature of the acts that constitute harassment or stalking, while others go on to define the types of conduct that are prohibited. The second necessary element is perpetrator intent. Most anti-stalking laws require either that the perpetrator intended to harass or cause harm to the victim or that they ought to have known that their behavior could have had that effect (akin to being reckless as to the effect of their behavior). The final element, which is present in many North American and European jurisdictions, although not in Australasia, is that the victim experiences the behavior as intimidating and feels fear or apprehension as a result. Two levels of proof are often required: a subjective standard (this victim felt fear) and an objective standard (a reasonable person would have felt fear). In many jurisdictions, stalking is divided into misdemeanor or felony offences, depending on the nature of the behavior engaged in. Furthermore, many anti-stalking laws provide for the civil option of a protection order for the victim, breaches of which are a criminal offence. Stalking laws have provoked considerable discussion, which is summarized by Kapley and Cooke (2007) and Mullen and colleagues (2009a).

Prevalence

Estimates of the prevalence of stalking vary according to the definition, sampling, method of enquiry, and willingness of participants to respond and answer candidly. Nonetheless, estimates are remarkably consistent across location and time, with lifetime victimization rates of 17–30 % reported for women and 4–12 % for men (Purcell et al. 2002; Smith et al. 2011; Stieger et al. 2008). Reported prevalence rates in the USA tend to be lower than other countries (7 % for women and 2 % for men) as the national survey included only those victims who experienced the stalking as dangerous or

life-threatening (Basile et al. 2006). Victims are largely female (70–80 %) and stalkers are predominantly male (80–85 %), with 20–25 % involving same-gender stalking, usually male-male.

Explaining Stalking Behavior

Stalking, like any complex human behavior, is a product of a range of psychological, social, and cultural influences. People stalk out of anger, hope, lust, ignorance, or a combination of these factors. The influence of psychopathology, discussed later, is also relevant in many cases of stalking, either as a direct or indirect causal factor. To date there are no comprehensive theories to explain the variety of stalking behavior that is observed. Some single factor theories have been described, most commonly disrupted attachment (Meloy 2007). This theory posits that stalking is a consequence of insecure and, specifically, preoccupied attachment. Research has supported aspects of this theory, finding that stalkers do indeed often present with insecure adult attachment, although there is little evidence for a specifically preoccupied style. Proponents of stalking as an expression of disrupted attachment tend to focus on those seeking relationships and do not address other possible trajectories into stalking behavior. Moreover, the mechanisms by which insecure attachment produces stalking in some contexts but not in others, even within the same individual, have not been adequately explained. Nonetheless, there does seem to be some evidence for chronically disturbed attachment style in some stalkers.

Where attachment theory examines distal dispositional characteristics that may lead to increased proclivity to stalk, Spitzberg and Cupach's (2007) application of relational goal pursuit theory to relationship-based stalking suggests more proximally relevant factors. In this theory, the goal of possessing a particular relationship becomes entwined with the perpetrator's value system and is integral to their sense of self-worth and life happiness. In this context, pursuit of the goal (a relationship with the victim) takes on disproportionate significance. The failure to achieve the goal in spite of persistent attempts

results in rumination and consequent negative emotional states, which the perpetrator attempts to escape by further efforts to achieve the desired goal. This is perhaps the most well-explicated theory of stalking behavior, although it also does not account for the variety of trajectories seen in stalking behavior, and to date it remains largely untested. A significant gap in the stalking literature remains, awaiting a truly comprehensive multifactor theory that adequately explains this complex behavior.

Classification Systems

In the absence of such a theory, many classification systems have been developed in an attempt to reduce the heterogeneity of stalkers into manageable categories that can assist decision-making. These taxonomies vary depending on the needs of the authors and audience. Those developed by clinicians emphasize the goal of treatment and management, while those for law enforcement tend to be used more as shorthand for management and referral options. In both realms, classification schemes tend to focus on three variables: mental disorder, prior relationship, and motivation. The earliest classification schemes focussed heavily on psychopathology, reflecting the fact that stalking behavior was seen as a manifestation of emotional and psychological disturbance. By the mid-1990s, prior relationship and affective state were also considered useful ways of differentiating stalkers. More recently, Mohandie and colleagues (2006) proposed a scheme based on the nature of the prior relationship and the context in which the stalking occurred. This scheme in particular may be useful for law enforcement personnel as it is relatively easy to apply based on minimal information. Many typologies have been proposed, and a useful review is provided by Pinals (2007).

A Multiaxial Classification Scheme

One of the most commonly used typologies, endorsed by the Group for the Advancement of Psychiatry Committee on Psychiatry and the Law, is that of Mullen et al. (2009). This typology, based originally on a sample of 145 Australian stalkers, took the unique step of proposing

classification on multiple axes to determine type. The first axis is the nature of the prior relationship between stalker and victim; the second, the initial motivation for and context in which contact with the victim occurred; and the third, the presence and nature of any psychiatric diagnoses. Using these three axes, five stalker types were identified, which are described below. More lengthy descriptions can be found in Mullen et al. (2009).

The Rejected: These stalkers begin to stalk following the dissolution of an intimate relationship, however brief, in an attempt either to reclaim the relationship or to exact revenge against their former partner for leaving. In many cases, both of these motives are present at different times. Not typically affected by severe mental illness, these stalkers tend to present with antisocial, narcissistic, borderline, or paranoid personality traits, depression, and/or substance use disorders. In a minority of cases, dependent and obsessive-compulsive personality traits are evident.

The Resentful: These stalkers target an acquaintance or stranger, but invariably someone who they perceive has mistreated them, either in their own right or because they are representative of some organization that has provoked the stalker's ire. The stalking usually begins with the aim of righting the perceived wrong and over time becomes the only way to assuage feelings of resentment and regain their lost sense of power and control. These stalkers often suffer from paranoid or narcissistic personality disorders and occasionally present with paranoid delusions incorporating the stalking victim(s).

The Incompetent Suitor: These individuals target strangers or acquaintances with the initial intention to pursue a friendship or date. The stalking is an inept and often self-centered attempt to achieve that outcome. Over time it may take on an angry or aggressive tone as the stalker's wishes are thwarted, but this group typically does not stalk for an extended period, instead transferring their interest elsewhere. While many stalkers who fall into this category do not experience any form of mental disorder, this behavior does emerge in the context of impaired social skills and interpersonal deficits,

and intellectual disability or developmental disorders are not infrequent. Also found in this group are narcissistic individuals who are assertive, overbearing, and simply can't imagine why any woman would refuse the opportunity to date them.

The Intimacy Seeker: These stalkers target acquaintances or strangers in an expression of their love and seeking love in return. Usually the stalker persists in an affectionate and amorous way, but very occasionally they can become angry and vengeful if they realize that the longed-for relationship is under threat. Often driven by a severe mental illness, these individuals may believe that they already have a mutually loving relationship with the victim, even in the absence of any true connection. Conversely, they may be intensely infatuated and believe that the idealized relationship is possible if they persist. Overvalued ideas of intimacy are almost ubiquitous in this group, arising either out of a severely disturbed borderline personality or at the more extreme end, a delusional disorder or schizophrenia with erotomanic delusions.

The Predatory: These men most often target strangers with the goal of achieving some form of sexual gratification, usually deviant. They use stalking as part of a wider repertoire of sexually inappropriate behaviors, and their stalking is usually characterized by surreptitious loitering, following, spying, and, in many cases, sexual violence. The predominant diagnoses in this group are paraphilias, depression, and substance misuse.

Psychopathology

Stalking is a set of behaviors, not a diagnosis, but studies suggest that at least 80 % of community referred stalkers have some form of psychiatric condition (McEwan et al. 2009). Personality disorders, schizophrenia, and major mood disorders are the primary presenting problems (McEwan et al. 2009; Rosenfeld 2003). In their retrospective file review of a forensically derived sample of stalkers, Mohandie and colleagues (2006) found 46 % had a probable diagnosis of mental disorder. Referral bias limits the generalizability of these findings, in that forensic samples are

weighted towards more serious offending and severe mental illness.

There is a higher prevalence of major mental illness among those who stalk public figures (James et al. 2009). Over 80 % of individuals who made inappropriate communications or approaches to the British Royal Family were found to have a major mental disorder. In a study of individuals who pursued and ultimately physically attacked a Western European politician, over half were mentally disordered, most of whom were psychotic.

Suicidal ideation is also common among stalkers. Using retrospective file review, Mohandie and colleagues (2006) identified suicidality in 25 % of 763 North American stalkers, while McEwan and colleagues (2009) observed that one in ten of their Australian sample reported suicidal ideation. In a 3-year follow-up, McEwan and colleagues (2010) found that 3 of 138 stalkers (2.2 %) had committed suicide. Though small, these numbers represent substantially higher rates, indeed six times greater, than would be expected among psychiatric patients or community-based offenders. In all three cases, the stalker had a history of suicidal ideation or intent. This finding is of particular concern because suicidal stalkers pose an enhanced risk of violence towards their victim and third parties. Stalking risks will be discussed further in a later section.

Particular psychopathology is found more commonly in certain types of stalker, as noted in the multiaxial classification scheme presented earlier. For instance, personality disorders or traits predominate in Rejected stalkers though morbid jealousy of delusional intensity is also encountered in this group. Intimacy seekers are far more likely to have serious psychopathology, in the form of psychotic disorders, and female stalkers have higher rates of psychosis than their male counterparts (Meloy and Boyd 2003). As elaborated later, the risk of serious violence in the stalking situation is greater when the stalker has a nonpsychotic as opposed to a psychotic condition.

Earlier assumptions that stalkers have above average intellectual functioning have been called into question by a prospective study of stalkers

referred to a clinic specializing in their assessment and treatment (MacKenzie et al. 2010). Intelligence testing found a mean intelligence quotient (IQ) in the average range (91.59 [SD 16.2]), with a verbal IQ significantly lower than performance IQ. These findings have important implications for psychological intervention programs for stalkers, which must reflect their cognitive abilities.

Victims and the Impact of Stalking

Stalking is an event that has a perpetrator and a victim. The victim is pivotal to stalking because stalking is, in effect, a victim-defined crime. Behaviors that are inappropriate and gauche are transformed into behaviors which are harmful and illegal by the apprehension that is evoked in the victim. While the line between the inappropriate and the fear inducing is in some respects influenced by the tolerance and sensitivities of the victim, most stalking that comes to the attention of the criminal justice system could be expected to provoke fear or disquiet in the majority of people.

Stalking has been viewed as a product of failed relationships and a form of domestic violence perpetrated by men against women. Earlier in the evolution of our understanding of stalking, the concept generalized from behavior directed at the famous to that of males who battered their partners. While lending attention to a long neglected issue, the concept of stalking is problematic when used to describe behaviors that occur in the context of an ongoing relationship. To begin with, the focus on stalking as a form of domestic violence has predominantly been on women as victims and men as abusers, while those who stalk former intimates can be female and victims can also be male. More importantly, while behaviors akin to stalking undoubtedly emerge among cohabiting partners, especially in relationships characterized by jealousy, this usually comprises behaviors that are aimed at information gathering (e.g., monitoring, surveillance) or coercive physical or verbal abuse and threats. A distinction should be made between such conduct and stalking. The aim of stalking is to make one's presence felt *where it would not otherwise*

exist. Where there is an ongoing relationship, however conflicted and potentially damaging, the behavior should be classified as a form of terrorizing and controlling intimate partner violence, not as stalking. Confusing the two is semantically and practically problematic as approaches to assessing and managing stalkers will be ineffective in situations where the stalker and victim continue to cohabit.

A number of classifications have been proposed for stalking victims, all of them based upon the preexisting relationship, if any, between victim and stalker. The following is a relational classification of stalking victims, derived from the typology proposed by Mullen and colleagues (2009):

1. *Former sexual intimate*. The most common victim profile is a woman who previously shared an intimate relationship with the stalker, usually male, though women can be perpetrators and stalking can also arise in the context of same-gender relationships. This category includes only those cases in which the relationship has been explicitly terminated. Acquaintances, who can be:
 2. *Family/friend*. Rejected stalking may follow the breakdown of a close friendship or estrangement from a family member. Occasionally, Resentful stalking can emerge in relation to some perceived injustice (e.g., dispute over a family estate).
 3. *Professional relationship*. As below, those in the health, law enforcement, and teaching professions are especially at risk.
 4. *Casual contact*. Intimacy seekers, Incompetent suitors, and Predatory stalkers may all target their victim in the context of a casual social encounter. Neighbor stalking usually falls into this category, most often involving a Resentful stalker incited by disputes over noise or a fence line.
 5. *Workplace contact*. Where the victim meets their stalker in a workplace context. The stalker may be in a subordinate or superior position, or in some cases a client or customer. Those without any real connection to the victim:

6. *Stranger*. Some victims have no prior knowledge of their stalker. The stranger may identify themselves, but some remain anonymous.
7. *The famous*. As noted below, prominent figures such as politicians and media personalities attract a range of stalker types, particularly Intimacy seekers and Resentful stalkers. Some celebrities suffer the attentions of multiple stalkers.

To this can be added secondary victims, who are impacted by stalking as a consequence of their relationship to the primary victim (e.g., the victim's partner or other family members, co-workers, or neighbors). Stalking tends to have a "ripple effect" on the victim's social network. In some instances, stalkers pose a greater risk to secondary victims than their primary target.

It has long been observed that stalking can have damaging, if not devastating, consequences and it is the reason, ultimately, that stalkers have attracted the interest of mental health professionals. Stalking impacts upon the victim, the third parties, the stalker themselves, and the community as a whole. Such impacts have received systematic attention over the past decade. The initial studies were mainly descriptive and highlighted the psychological, occupational, and social impacts of these behaviors (Pathé and Mullen 1997). Subsequent large epidemiological studies (e.g., Purcell et al. 2002; Purcell et al. 2005) found that stalking was prevalent in Western nations and also confirmed the higher rates of psychiatric morbidity among victims of stalking.

Victims of stalking commonly report a profound sense of violation and loss of control over their lives. Many experience a pervasive mistrust of others that can persist long after the stalking has ceased. In their epidemiological study that compared matched controls, victims of brief harassment, and victims of protracted stalking (>2 weeks), Purcell and colleagues (2002, 2005) found that the rates of general psychiatric morbidity were significantly higher among those who reported a history of protracted stalking (36 % vs 19 % for controls and 22 % for brief harassment). Those subjected to severe stalking were also more likely to report recent suicidal ideation (10 %) and clinically significant

levels of posttraumatic symptomatology (three times more likely than the briefly harassed group).

More recent studies of stalking victims have considered groups who are particularly vulnerable to stalkers. These include health professionals (Whyte et al. 2011) and public figures (Mullen et al. 2009), but there is also anecdotal evidence of higher rates of stalking victimization in other professions such as teaching and law enforcement. Studies to date suggest that the lifetime risk for a health professional being stalked by a client or patient is 20 % (Purcell et al. 2005), and those working in the mental health field, by virtue of the patients they encounter, may be more susceptible to such behaviors. One of the commonest motivations for stalking health practitioners is morbid infatuation, where a lonely, disordered patient misconstrues sympathy and attention as romantic love. These patients-turned-stalkers can be categorized as Intimacy seekers. Resentment is another common motivation, stemming from some supposed injury or dereliction. These stalkers fall into the Resentful category. A smaller group is pursuing reconciliation or revenge following the termination of a (usually long-term) therapeutic relationship (Rejected stalkers). The impact of stalking on health professionals can be significant, with over 70 % of affected psychologists in one study (Purcell et al. 2005) modifying aspects of their personal and professional lives, some of whom relocated their practice and/or their private address. Almost a third contemplated leaving their profession altogether, this figure rising to nearly 50 % among those whose stalker made malicious complaints to their professional registration board.

Assessment Guidelines

Assessments of stalkers or victims often take place in a stressful context. For the stalker, this may be as part of a presentence or parole report, or in a mandated treatment situation. For victims, help is often sought when the pressures of the stalking become overwhelming and they feel that they can no longer cope on their own. Stalkers tend to deny, minimize, rationalize, and justify their behavior. Victims often minimize the

experience of stalking and overemphasize their responsibility for the harassment. In both scenarios, obtaining collateral information about the stalking episode is essential. Where available, police or court documents summarizing the behavior are invariably useful, as are accounts by those associated with the victim or stalker. When possible, it is usually best to avoid having contact with both the victim and stalker. Victims may experience attempts to obtain collateral information as simply another unwanted intrusion by someone acting on behalf of the stalker, while for stalkers, having a professional make contact with the victim can be understood as implicit support for an ongoing “relationship” or as an opportunity to obtain new information about the victim.

Assessing a Stalker

It is usually possible to form hypotheses about the motivation for stalking reasonably quickly. Information about the prior relationship between victim and stalker and about the apparent initial motivation can be gleaned from collateral sources or identified very early in an assessment. Often the best way of determining motivation is to simply ask the stalker what drove them to this behavior or what they originally wanted to achieve. A central aspect of the assessment is conducting a functional analysis of the stalking. This involves examining the common situational, emotional, and behavioral antecedents to and consequences of each intrusion. An analysis of this kind can provide useful information upon which to base hypotheses about the cognitive and emotional states associated with offending and the skill deficits that might lead them to stalk rather than use a more adaptive strategy to achieve their goals.

For mental health clinicians, particularly those with forensic expertise, the assessment will be like most others, with slightly more emphasis on the individual’s experience of other, similar interpersonal situations. For example, for a Rejected stalker, additional questions about prior relationships and breakups would be relevant, and for a Resentful stalker, enquiry into how they have handled previous injustices or disputes would be

useful. Where sexually deviant behavior is present, a full sexual behavior assessment is appropriate, as for sex offenders. If the clinician is not trained in assessment and diagnosis of psychopathology, referral to an appropriately qualified expert is strongly recommended.

Assessing a Victim

Some stalking victims present for mental health intervention after the stalking has ceased. They are commonly seeking help to deal with post-trauma symptoms and reconstitute their life. An appraisal of the risk of recurrence of the stalking may also be appropriate. Many more stalking victims seek assistance when the stalking is ongoing, and the priority in these instances is to ensure all reasonable steps are taken to protect the victim's safety and privacy. Some may be engaging in behaviors which place them and other parties at risk (e.g., retaliating against their stalker), necessitating urgent guidance. Around 10 % will have suicidal ideation requiring assertive intervention.

As a consequence of their experiences, many stalking victims are distrustful, and some will have previously encountered negative attitudes from the "helping" professions. In engaging stalking victims, it is essential that the assessment is conducted by a professional familiar with stalking issues and the available interventions. It is important at the outset that stalking victims are reassured that the stalking is not their fault and that they are not alone in their ordeal.

Stalking victims require a comprehensive assessment, including a full psychosocial history that encompasses past victimization and a detailed account of the stalking. This includes the nature of the victim's prior relationship with the stalker, a description of the onset of the stalking, and the frequency and nature of the intrusions, as well as any strategies employed to end the stalking. How effective were these measures, and has anything exacerbated the behavior? Have they involved the police, and if not, why not? Are there any current protective injunctions, or is the stalker subject to any other legal directives? Is the stalker known or suspected to be mentally ill, and if so, are they known to the

mental health system? Some victims, especially those stalked by strangers, may have limited information, but victims should be encouraged to bring as much documentation as possible to the assessment (including any evidence of the stalking in the form of letters, e-mail, phone records, photographs, court orders, or witness statements).

Stalking victims should be questioned about suicidal inclinations, any homicidal intent towards the stalker, and any measures they may have taken to defend themselves (e.g., keeping weapons in the house). One should endeavor to gauge the victim's current level of support, others affected by the stalking, the victim's employment situation, and mental health status. Are they suffering current anxiety or depressive symptoms, and how have these been managed? Maladaptive coping mechanisms such as substance abuse or gambling are not uncommon and may require specific interventions.

Risk Assessment

In any stalking situation, a primary concern for the victim is the likelihood of further damaging behavior. Often their worries focus on the likelihood of physical violence, the likelihood that the stalker will persist, or if they have stopped, that they will start again. Therefore, for those charged with managing stalking situations, risk assessment is a primary focus. For those working with stalkers, an additional concern is the potential for psychological and social harm to the perpetrator and what this means for other areas of risk.

While the stalking risk assessment literature continues to lack strong retrospective or prospective studies of representative samples, a sizeable body of published work now exists, and relevant information can also be gleaned from the much larger body of research into violence and offending risk assessment more generally. Two structured professional judgement tools for assessing stalking have been published and are undergoing validation trials – the *Stalking Risk Profile (SRP)* (MacKenzie et al. 2009) and the *Guidelines for Stalking Assessment and Management (SAM)* (Kropp et al. 2008). Where these tools are not available, clinicians and law

enforcement personnel must rely on clinical experience and integrating the literature themselves.

Risk of Persistence or Recurrence Stalkers that persist for at least 2 weeks are at increased risk of continuing for months. Those with a prior relationship with the victim tend to stalk for longer, and as a rule, women tend to be more persistent than men. By far the strongest predictor of extreme persistence is psychosis, and these stalkers are unlikely to desist without mental health intervention. Recurrent stalking appears to be the domain of the personality disordered and those who share the victim's children, property, or locale, although a lack of research makes firm conclusions difficult (Rosenfeld 2003).

Risk of Violence As a group, ex-intimate partners present the greatest risk of violence, with up to half in this category using physical violence against their victim or a related third party (McEwan et al. 2007). Identified risk factors for violence for all stalkers include:

- *Demographic variables.* It has been repeatedly shown that the gender of the stalker or the victim has no impact on the prevalence of either threats or assault, and same-gender stalking victims are at equal risk of violence (Strand and McEwan 2011).
- *Stalking behaviors.* Stalkers who follow, loiter near, accost the victim, or enter the victim's home are at increased risk of violence. Escalations in the intrusiveness or intensity of the stalking have not been subject to research, but should be treated as periods of increased risk.
- *Threats.* Threats are common in stalking situations and have been shown to be a risk factor for violence, particularly for ex-intimate partner violence (McEwan et al. 2007).
- *Prior violence and criminal behavior.* The evidence is inconsistent; however, on balance, it seems that prior violence in particular is a predictor of stalking-related violence (McEwan et al. 2007).
- *Mental disorder.* Psychotic stalkers have consistently been found to be less violent than nonpsychotic stalkers; however, the presence

of other risk factors should not be discounted because the stalker is psychotic (McEwan et al. 2009). The relationship between personality disorder and violence in stalkers is unclear, although findings from the wider risk assessment literature indicate that antisocial personality disorder in particular is related to violence.

- *Substance misuse.* Substance misuse is associated with stalking violence, likely having a compounding effect on other risk factors.

Formulation and Treatment Planning for Stalkers Clinical formulation links assessment information with a systemic treatment and management plan. A good formulation will explain the mechanisms underlying the stalking and direct interventions to stop it. For the clinician assessing a stalker, the central focus of formulation is the *stalking behavior* itself rather than the presence of psychopathology or other matters. A detailed explanation of how to formulate stalking behavior is beyond the scope of this entry, but some guidelines are available in McEwan et al. (2011).

Treatment and Management

Stalkers

Management of stalkers often involves the use of the criminal justice system, which is described further below. Presented here are options for interventions designed to address the underlying causes of the behavior in conjunction with such environmental management strategies. The information provided in this section assumes the availability of an appropriately trained and supported mental health practitioner and a client that is willing and/or mandated to attend treatment. Whenever possible, treatment of stalkers should be undertaken by a clinician experienced in working with offenders, who has training in structured risk assessment, familiarity with the stalking research literature, and an understanding of local stalking laws. Those who work with stalkers (or their victims) should also be cognizant of the risks of such work and should take appropriate steps to ensure their own safety and security (Pathé 2002).

There is limited information available about the efficacy of offence-specific psychological interventions for stalkers, with Rosenfeld and colleagues' (2007) trial of dialectical behavior therapy the only published study. Treatment strategies described by these and other authors (e.g., MacKenzie and James 2011) take a cognitive-behavioral perspective and focus on helping the stalker to develop skills that they can apply in situations where they have previously used stalking (e.g., emotional regulation, problem solving, and communication skills are all targets of psychological intervention). Which particular skill deficits are most relevant to an individual stalker, and the prioritization of one skill over another, should be evident following a comprehensive formulation as outlined above. In addition to skill development, another central aspect of treatment is helping the stalker recognize potential triggers for stalking behavior, identify high-risk thinking and behavioral patterns, and implement alternative strategies.

As with any offender treatment, attention must also be paid to the client's readiness to desist from their behavior. Treatment readiness can be defined as having appropriate motivation, capacity, and ability to engage in treatment so as to maximize the likelihood of therapeutic change. Where motivation to engage is low, time should be spent helping the client identify personal goals that would be assisted by avoiding stalking in the future, encouraging the view that they are able to change their behavior, and reducing emotional distress that might be provoked by thinking or talking about their offence.

Maximizing treatment readiness may require other issues to be dealt with prior to beginning any offence-specific interventions. Perhaps the most common factors that interfere with a stalker's capacity and ability to engage in offence-related treatment are active symptoms of mental illness and substance misuse. As discussed previously, these may also be risk factors for further stalking behavior, and so appropriate treatment can both reduce risk and increase the client's opportunity to engage in other offence-related interventions. Mullen et al. (2009) provide a detailed description of the

management of mental disorders in stalkers. In essence the appropriate treatments are no different to those offered to non-stalkers, although consideration of risk may lead to increased emphasis on the safety of the client and others and so greater impetus for involuntary treatment. Similarly, where substance misuse is present, interventions offered to individuals who do not stalk are usually appropriate, although with due consideration of the need to protect the stalking victim while treatment for substance misuse is undertaken.

Victims

Strategies to Combat Stalking These approaches need to take account of the individual circumstances, including the prior relationship between victim and stalker; the presence of shared children, property, or workplace; the methods of harassment; and the assessed risk to the victim and/or other parties. In most situations, the victim should be encouraged to inform relevant people of the stalking, so they are alert to the problem and better equipped to protect the victim and themselves. It is important to advise victims at the outset to avoid all contact with their stalker. Attempts to appease the stalker through "one last meeting," or to explain, yet again, why the relationship ended, are, in effect, perpetuating the "relationship" and reinforcing the stalker's efforts.

Using Anti-stalking Laws As stalking is a criminal offence in many jurisdictions, informing the police may not only offer a resolution but also open an official paper trail documenting the stalker's course of conduct. To maximize the likelihood of police action, victims should be encouraged to keep a clear chronological record of harassing intrusions and any tangible evidence of the stalking. While there are few surveys of police responses to stalking complaints, those that exist show that in the majority of cases, stalkers are charged with offences other than stalking (e.g., trespass), even where specific anti-stalking legislation exists. Anecdotal evidence suggests that police often recommend that the victim obtain a protective injunction,

enabling them to prosecute a breach; however, this strategy may carry some risks (see below). Storey and Hart's (2011) investigation of 32 cases showed that legal intervention is one of the more effective tactics to stop stalking. This includes everything from a formal police warning to a conviction. Informal warnings from police were shown to be ineffective in this study.

Protective Injunctions Also termed restraining, non-molestation, apprehended violence, protection, no contact, or intervention orders, their use against stalkers remains a contentious issue. These orders should not be used indiscriminately with stalkers as they have the potential to inflame the situation, provide a false sense of security, and exacerbate the victim's distress. Protective injunctions are more likely to be effective if they are properly policed, and the perpetrator is an Incompetent suitor or a Rejected stalker who has less invested in the relationship (e.g., a brief relationship with no children) and with a nonviolent, law-abiding background. These orders are less likely to succeed with Intimacy Seekers, especially those with erotomantic delusions, and Rejected stalkers who have a strong investment in the relationship (long-standing, shared children and assets) and/or an antisocial personality disorder, morbid jealousy, or a history of domestic violence.

Privacy, Safety, and Security There are a range of measures that can enhance the victim's security. These are beyond the scope of this entry but are detailed in Pathé (2002). Victims and potential victims such as health practitioners should try to ensure that as far as possible all personal information is removed from the public domain. This includes delisting their telephone and fax numbers, obtaining a post office box address, and removing their name from the electoral roll and property titles. Stalking victims should be cautious in their use of social networking websites such as Facebook and Twitter and take appropriate steps to protect their privacy.

Clinical Management of Stalking Victims Cognitive-behavioral approaches have

been recommended with stalking victims. As a consequence of their stalking experiences, a victim's previously held beliefs in their strength and resilience are replaced by feelings of extreme vulnerability, and their previous assumptions about the reasonable and predictable nature of the world are shattered. Cognitive-behavioral therapies aim to correct the unhelpful assumptions that underlie their anxiety and depression and the avoidance behaviors that frequently emerge in this context. Priority should always be given to the client's safety, and if the stalking is ongoing, some avoidance behaviors may be entirely appropriate. For stalking victims presenting with chronic or complex posttraumatic stress disorder, dialectical behavior therapy (DBT) may prove efficacious (Mullen et al. 2009b).

Although there is a dearth of randomized clinical trials for pharmacological agents in the treatment of stress-related syndromes, those most frequently used with victims of stalking, as an adjunct to psychological approaches, are the selective serotonin reuptake inhibitor (SSRI) antidepressants. These agents have demonstrated efficacy for stress-related symptoms, panic, and depression. Care should be exercised in prescribing benzodiazepines for anxiety, as the symptoms, and thus the need for treatment, may be protracted, increasing the risk of drug dependence. Furthermore, the benzodiazepine agents such as Valium may diminish alertness and self-control, which can exacerbate the victim's feelings of vulnerability.

Other Therapies Wherever possible, the victim's partner and other relevant family members and supports should be involved in their management and the development of a safety plan. They frequently share in the victim's distress and may have been more directly impacted than the primary victim. They often seek information to try to make sense of the stalking, and educating significant others can alleviate pressure on the victim, by stemming criticism and unhelpful advice. Some may require referral for separate counselling.

Group therapies and support organizations can provide some validation for stalking victims and

assist in reducing the victim's sense of isolation and self-blame. They are a useful source of information, support, and advocacy. Stalking victim groups need to be mindful of security at the meeting venue and the confidentiality of their members.

Conclusions

Stalking is a complex and damaging behavior that has commanded the attention of behavioral scientists, mental health professionals, the legal system, and the general public. Stalkers vary according to their motivations and psychopathology, and this determines their level of risk in multiple domains, and their management. While theories of stalking are not well understood, this problem behavior is amenable to psychological and environmental interventions. The mental health and criminal justice systems have a key role to play in bringing stalking episodes and victim suffering to an end.

Related Entries

- ▶ [Police Family Violence Services](#)
- ▶ [Risk Assessment, Classification, and Prediction](#)

Recommended Reading and References

- Basile KC, Swahn MH, Chen J, Saltzman LE (2006) Stalking in the United States: recent national prevalence estimates. *Am J Prev Med* 31:172–175
- James DV, Mullen P, Pathé M, Meloy J, Preston L, Darnley B et al (2009) Stalkers and harassers of royalty: the role of mental illness and motivation. *Psychol Med J Res Psychiatry Allied Sci* 39:1479–1490
- Kapley DJ, Cooke JR (2007) Trends in antistalking legislation. In: Pinals DA (ed) *Stalking. Psychiatric perspectives and practical approaches*. Oxford University Press, New York, pp 141–163
- Kropp RP, Hart SD, Lyon DR (2008) *Guidelines for stalking assessment and management (SAM)*. ProActive Resolutions Inc, Vancouver
- MacKenzie RD, James DV (2011) Management and treatment of stalkers: problems, options, and solutions. *Behav Sci Law* 29:220–239
- MacKenzie RD, McEwan TE, Pathé MT, James DV, Ogloff JRP, Mullen PE (2009) *Stalking risk profile: guidelines for assessing and managing stalkers*. StalkInc & Centre for Forensic Behavioural Science, Monash University, Melbourne
- MacKenzie R, James D, McEwan T, Mullen P, Ogloff J (2010) Stalkers and intelligence: implications for treatment. *J Forensic Psychiatry Psychol* 21:852–872
- McEwan TE, Mullen PE, Purcell R (2007) Identifying risk factors in stalking: a review of current research. *Int J Law Psychiatry* 30:1–9
- McEwan TE, Mullen PE, MacKenzie R, Ogloff JRP (2009) Violence in stalking situations. *Psychol Med J Res Psychiatry Allied Sci* 39:1469–1478
- McEwan TE, Mullen PE, MacKenzie R (2010) Suicide amongst stalkers. *J Forensic Psychiatry Psychol* 21:514–520
- McEwan TE, Pathé M, Ogloff JRP (2011) Advances in stalking risk assessment. *Behav Sci Law* 29:180–201
- Meloy JR (2007) Stalking: the state of the science. *Crim Behav Ment Heal* 17:1–7 [Editorial]
- Meloy JR, Boyd C (2003) Female stalkers and their victims. *J Am Acad Psychiatry Law* 31:211–219
- Meloy S, Hoffman (2008) *Stalking, threatening and attacking public figures*. Oxford University Press, New York
- Mohandie K, Meloy JR, McGowan MG, Williams J (2006) The RECON typology of stalking: Reliability and validity based on a large sample of North American stalkers. *J Forensic Sci* 51:147–155
- Mullen PE, James DV, Meloy J, Pathé MT, Farnham FR, Preston L et al (2009a) The fixated and the pursuit of public figures. *J Forensic Psychiatry Psychol* 20:33–47
- Mullen PE, Pathé M, Purcell R (2009b) *Stalkers and their victims*, 2nd edn. Cambridge University Press, New York
- Pathé M (2002) *Surviving stalking*. Cambridge University Press, Cambridge
- Pathé M, Mullen PE (1997) The impact of stalkers on their victims. *Br J Psychiatry* 170:12–17
- Pinals DA (2007) Stalking: classification and typology. In: Pinals DA (ed) *Stalking. Psychiatric perspectives and practical approaches*. Oxford University Press, New York, pp 27–60
- Purcell R, Pathé M, Mullen PE (2002) The prevalence and nature of stalking in the Australian community. *Aust N Z J Psychiatry* 36:114–120
- Purcell R, Pathé M, Mullen PE (2004) Editorial: when do repeated intrusions become stalking? *J Forensic Psychiatry Psychol* 15:571–583 [Editorial]
- Purcell R, Pathé M, Mullen PE (2005) Association between stalking victimisation and psychiatric morbidity in a random community sample. *Br J Psychiatry* 187:416–420
- Rosenfeld B (2003) Recidivism in stalking and obsessional harassment. *Law Hum Behav* 27:251–265
- Rosenfeld B, Galietta M, Ivanoff A, Garcia-Mansilla A, Martinez R, Fava J et al (2007) *Dialectical behavior*

- therapy for the treatment of stalking offenders. *Int J Forensic Ment Health* 6:95–103
- Sheridan LP, Grant T (2007) Is cyberstalking different? *Psychol Crime Law* 13:627–640
- Smith KE, Coleman K, Eder S, Hall P (2011) Homicides, firearms offences and intimate violence 2009/10. Home Office, London
- Spitzberg BH, Cupach WR (2007) The state of the art of stalking: taking stock of the emerging literature. *Aggress Violent Behav* 12:64–86
- Stieger S, Burger C, Schild A (2008) Lifetime prevalence and impact of stalking: epidemiological data from Eastern Austria. *Eur J Psychiatry* 22:235–241
- Storey J, Hart S (2011) How do police respond to stalking? An examination of the risk management strategies and tactics used in a specialized anti-stalking law enforcement unit. *J Police Crim Psychol* 26:128–142
- Strand S, McEwan TE (2012) Violence among female stalkers. *Psychol Med* 42:545–556
- Whyte S, Penny C, Christopherson S, Reiss D, Petch E (2011) The stalking of psychiatrists. *Int J Forensic Ment Health* 10:254–260.

State Crime

Dawn L. Rothe
Department of Sociology and Criminal Justice,
Old Dominion University, Norfolk, VA, USA

Overview

State crimes have been one of the foremost social problems of the past 100 years, with wide-reaching costs and levels of victimization. During the course of the twentieth century, the state crimes of Turkey, Nazi Germany, Stalinist Russia, Pol Pot and the Khmer Rouge, and Maoist China were especially large-scale, dramatic examples. More current examples include the continued possession of nuclear weapons by some states in violation of the Nuclear Nonproliferation Treaty; the US wars of aggression in Iraq and Afghanistan; genocides and crimes against humanity that include the Serbian attack on Bosnian and Croatia, Rwanda, Yugoslavia, and Sudan's actions against Darfur; and the Democratic Republic of Uganda and the Democratic Republic of Congo's wars and stealing of natural resources by states and corporations alike. Other

charges of state crime occurring around the world as of 2011 include the political, economic, and militarized oppression involving Egypt, Israel, the United States, China, Russia, Chechnya, and North Vietnam all of which have resulted in the victimization of tens of millions that lost their lives, were rendered homeless, imprisoned, and psychologically and physically damaged through the illegal or socially harmful actions of governments.

Despite the gravity, costs, and extensiveness of crimes committed by states, the field of state crime remains understudied relative to conventional street crimes in the field of criminology and criminal justice. Nonetheless, over the last two decades, considerable theoretical, conceptual, and empirical progress has been made by criminologists to better specify the nature, extent, distribution, causal variables, and issues associated with state violence. However, there remain two primary areas of debate within the field of state crime: the standard by which to define state crime and the issue of controlling state criminality for practitioners, politicians, citizens, as well scholars of state crime. This entry provides an overview of the history of state crime and subsequent costs, followed by a discussion of the contentious issues that remain within the field. The entry concludes with a potential philosophical change that could prove to be the most fruitful means in which to constrain and control state criminality.

The Fundamentals

History and Overview of the Field of State Crime

The criminological study of state crime can be traced back to Edwin Sutherland (1939), who called attention to a then-neglected form of crime, namely, the crimes of respectable people in the context of a legitimate occupation and of corporations. His extension of the concept of crime, beyond its conventional parameters, provided an important foundation built upon by several later scholars of white-collar crime. However, one can argue that William Chambliss'

1988 American Society of Criminology Presidential address on state-organized crime provided the more direct and immediate inspiration for more systematic attention to crimes of the state. Exploring crimes such as piracy and smuggling, Chambliss showed how states can be crucial in the organization and support of activities that violate their own laws and international laws when doing so fulfills their broader political and economic objectives. A number of criminologists, particularly critical criminologists, quickly adopted the concept, broadening and enriching the field. Their early work focused not only on crimes tacitly supported or organized by a sovereign polity but also on actions committed on behalf of states themselves. However, the early research on state criminality was plagued by definitional issues and generated much debate regarding whether the individual or the state (organization) was culpable for acts deemed state crime and what standards should be used to define state criminality. These two contested areas cut to the core of the field of criminology in general; thus, it was not surprising that this debate influenced the early development of the field of state crime and, in some cases, continues today and will be discussed further in the following sections.

Regardless of the remaining contentious issues associated with standards and definitions, since the onset of criminological inquiry concerning state criminality began the field has grown exponentially. For example, literature has been produced on state crime with topics ranging from the US invasion of Iraq, the illegal use of and threatened use of nuclear weapons, the ongoing genocide in Darfur, crimes against humanity in Uganda, the treatment of illegal aliens, the US role in and lack of response to Hurricane Katrina, to the many cases of state-corporate crime such as the Challenger, Imperial Foods, ValuJet cases, and more recently cases involving Halliburton, BP, and Abu Ghraib. Additionally, there are now two comprehensive texts on state crime, Penny Green and Tony Ward's (2004) *State Crime* and Dawn L. Rothe's (2009) *State Criminality: The Crime of all Crimes*, and six edited anthologies on state crime, including *Resistance to State*

Crime, edited by Elizabeth Stanley and Jude McCulloch (2012); *State Crime in the Global Age* edited by William Chambliss, Raymond Michalowski, and Ronald Kramer (2010); *State Crime, Current Perspectives*, edited by Dawn L. Rothe and Christopher W. Mullins (2010); *State-Corporate Crime: Wrongdoing at the Intersection of Business and Government* edited by Raymond Michalowski and Ronald Kramer (2006); and *Varieties of State Crime and its Control*, edited Jeffrey Ian Ross (2000). Additionally, John Hagan's (2010) *Who Are the Criminals? The Politics of Crime Policy from the Age of Roosevelt to the Age of Reagan* examines both state policies towards street crime as well as acts of state crime within and beyond the borders of the United States. Other cognate areas of state crime include political crimes, political white-collar crimes, environmental crimes, finance crimes, and the recently added crimes of globalization (see entries in this volume).

Costs of State Crime

Committing the most harmful of crimes are entities and individuals acting on the behalf of, or in the name of, the state. The types of costs include physical costs such as death, dismemberment, torture, ill health, and other costs associated with interpersonal violence directly and indirectly; environmental costs that include natural resources; economic costs; and psychological costs that are associated with a host of victimization and perpetrator mental and emotional health issues. For example, crimes of the state also can lead to the destruction of infrastructures, resulting in additional devastating harms (e.g., the 2002 sinking of the Senegalese ferry *Le Joola* with a loss of more than 1,000 lives) and environmental destruction with lasting effects on generations of citizens (e.g., the 1986 disaster at Chernobyl). Perhaps most importantly, state-committed crimes rip asunder the social trust between a state and its citizenry as well as trust between states (e.g., invasion and occupation of sovereign states, including most recently Iraq). When states violate such trust, domestically or internationally, they threaten the security of the global order, peace, and their own

legitimacy. Given the above, it is impossible to estimate an overall cost associated with state crime, yet, we do know it far outweighs that of street crime.

Key Issues and Controversies

History of Contention: Definitions and Standards

Dating back to Sutherland, the notion of an organization being criminally liable had consistently been met with resistance by criminologists until the late 1970s through the mid-1980. During this period some criminologists began to incorporate ideas from organizational sociologists' research as they argued that social scientists needed to move beyond focusing on the individuals who make up an organization and to recognize that the aggregate whole functions as an entity. As such, it was suggested that organizations, as social actors, can and should be the primary focus of analysis in state and corporate crime.

Others strongly objected to the notion of a state, as a social actor, in an analysis of state or corporate crime. Furthermore, within criminology, the idea of a state being criminally liable was met with significant resistance. There were those that denied state criminality was possible. However, within the international legal arena, the notion of a state as an actor that could be held accountable was already well underway as the concept of a state, as an entity possessing individual rights and subject to criminal liability, emerged back in the mid-1900s. Likewise, several legal scholars observed that there was a connection between individual criminal responsibility and state criminal responsibility under international law.

On the other hand, scholars who supported the idea of state criminality were divided upon the standards to be used to define such acts as criminal between those who favored a legalistic frame and others who favored a broader frame ranging from social harms to human rights. The tensions and debates over defining state crime, however, reflect a broader debate within criminology itself. Utilization of state-produced legal codes has long

been the stated and unstated norm. Such a reliance on state-produced definitions has caused tensions within street crime and white-collar crime studies, with critically orientated criminologists rejecting state-produced definitions. The political nature of law production has long been the main rationale for this rejection given that one cannot separate the nature of the political process that guides legislatures (and legislators) from the legislation produced. Further, states have an inherent drive to fulfill their own self-interest and not define harmful and problematic behavior as criminal (especially their own). Consequentially, within the field of state crime in particular, alternative formulations have been advanced, including the standard of basic human rights precepts, socially analogous harms, and social harm framework, to the perceptions of the state's citizens, each not without their own criticisms, strengths, and weaknesses.

Standards

The substance of a debate about definitions goes beyond merely critiquing the source or the substance of a given definition as alternative formulations were put forth. For example, in 1970, Herman and Julia Schwendinger suggested using a humanistic approach that would draw from objectively identifiable harms to humans and violations of human rights as the core definition of crime. Others have advocated that crimes are any socially injurious actions, regardless of the actor in question. Still some scholars have advocated that state crime should be defined by a social audience that recognizes the act as deviant. More recently, some criminologists have called for the abandonment of the concept of crime entirely in favor of Zemiology, the study of harm, thus, a social harm standard. In general, two positions on standards to be used to classify state actions as criminal remain within the broader definitional debate among scholars of state crime: crime as a social harm definition and a legalist approach.

The legalist approach includes a state's own domestic law as well as the broader umbrella of international public law (customary law, treaties, charters, and the newly emerged criminal law). This framework includes other approaches and

standards such as human rights and social and economic harms. Furthermore, international criminal law covers individuals as well as states, thus resolving any enduring reservations of the state as actors versus individuals. Additionally, the legalists' use of extant statute identifies an external reference point, while other approaches are said to use a more amorphous and relativistic definitional rubric. Those that accept the use of a legalistic standard suggest that it adds legitimacy to the field's definition. Legalists argue that if a critique of state crime studies is that they are not truly scientific but rather politically inspired diatribes, establishing the illegality of such actions under a legal code is a fitting response to such critiques. Nonetheless, this approach has been criticized as another example of continuing to use law as a "tool of the state" to control the very entities that create it.

The social harm approach begins with the realization that crime has no ontological reality and it is subjectively defined by states within the context of broader issues of power and political and economic interests. In the case of crimes of the powerful, harmful activities are rarely defined as criminal. Many acts and behaviors that cause serious harm are not part of the domestic or international criminal law, thus being omitted by those using a stricter rubric such as the legalist standard. Those advocating this framework suggest that harm be defined as physical, financial and economical, psychological and emotional, sexual, and cultural. This includes the observable forms of harm, but also those seemingly more "natural" including death and illness caused by starvation, untreated treatable illnesses due to lack of medical care, those who lose their traditions and communities due to economic displacement and relocation, and a host of other ways millions of humans suffer and experience harm as a result of state policies and actions. Most importantly, this approach focuses on the origins of the harm rather than merely the actors or states involved or the act itself. This is not to say the other approaches do not do this in their theoretical and analytical assessments; however, those working from a social harm perspective begin with this focus.

There are pros and cons to using either standard to determine what is to be considered a state crime. For example, the use of international law provides clarity and precision in the definitional processes. One need not negotiate the problematic aspects of defining "harm" per se. One of the major issues with using the legalist framework is that both state-produced law and international law are the result of a questionably legitimate political process. Additionally, given the objective nature of law as defined, there is the criticism that using it as a frame to define behavior as criminal serves to maintain particular power relations at both the state and international level. In fact, international law often fails to be created or is created in a nonjusticiable way, to express politicization of its construction. While states not only have the ability to define crime, they are more often than not powerful enough to resist definitions of crime that label their own behavior criminal. The above issues are also related to a separate but equally important problem: that of enforcement.

On the other hand, the social harm framework allows for various forms of negligence and actions that are harmful yet do not meet any extant legal definition of crime to be examined and decried. This perspective allows criminologists to develop their own determination of their subject matter rather than rely on predetermined laws and human rights to dictate the field of inquiry or make the claim that a particular state behavior is "wrong," thus removing the external political influences. The social harm perspective then avoids the formal institutionalized problematic way in which crime is defined and potential additional levels of harm that could be generated by controlling crime in general through a formal system of response. Nonetheless, using a social harm frame has its own weaknesses. While this framework might add a conceptual and ideological purity to definitional processes, it takes an already broad subject matter and casts the net wider to what could be perceived as nearly the entirety of individual and institutional behavior within contemporary societies. Additionally, as with any other standard, the question becomes who then defines what is or is not harm?

Given the strengths and weaknesses of the above standards, the field of state crime now incorporates both positions. For example, while a legalist may well stick to the standard of international public law, there are times when such a standard is not acceptable. Consider the research on state crime involving, explicitly and implicitly, international financial institutions (IFIs) or crimes of globalization. Here, the intersection of these institutions not only results in immense harm, but the harms are not covered through the application of international public law, save for human rights violations in some cases. Although the policies and implicit and explicit actions of these IFIs can result in or facilitate state criminality, they are not covered under international public law, thus requiring a broader standard for defining such actions as "criminal." On the other hand, the use of a social harm standard can be complementary to the legalist approach when such harms include human rights violations. In both of the examples noted, the intermixing of these standards provides researchers of state criminality with the objective foundation as argued by the legalists within the limitations of the legal perspective. The value of both positions should not be recognized as the field of state crime research continues to grow.

The Quagmire of Controlling State Criminality

Resisting or controlling state criminality is bound to some degree with the definitional quagmire. Those that draw from the legalist standard promote formal and informal accountability mechanisms from social condemnation to prosecution. On the other hand, scholars working within the social harm framework generally see resistance and controls not in terms of accountability but in terms of policy responses. There are additional forms of harm that result from relying on a crime control industry. According to those who embrace a social harm standard, it is far more important to address not just the harms committed but also the underlying structural conditions that facilitated them. Accordingly, they suggest that this requires debates about policy and

resources rather than handing the problem over to another arm of the broader power structure that, more often than not, further facilitates social harms: a criminal justice system. It requires using something beyond "the masters' tools" to confront, constrain, and control state crime. This can include movements from below, including social movements as has been witnessed in 2011 in Egypt, Tunisia, Libya, Bahrain, Yemen, and other Middle Eastern countries and with the Occupy Wall Street Movement.

Other scholars of state crime view international law as fundamental to the potential control of crimes of the state and mechanisms of accountability. These mechanisms include an international court such as the International Criminal Court and other states' domestic criminal justice systems. Additionally, impunity for crimes of the elite has long been an issue. Granting them near immunity for their actions in favor of extensive definitional and policy-related debates would suggest a reinforcement of impunity. Further, controlling state crime and holding heads of state or other high-ranking government officials accountable have long been met with contradictions and controversy from the perspective of state crime scholars as well as juristic practitioners. Consider that current geopolitical and international legal structures offer no threat of consequences. International law can be violated without threat of prosecution simply because there is no empowered institution to do so, and when it does occur, it is riddled with issues of selectivity. This reality has led many criminologists to deduce that the legalist standard and subsequent domestic and international laws are meaningless and not a valid framework for identifying or controlling state criminality. Indeed, for the legalist standard, the greatest obstacle is to actually control state criminality in some meaningful way. This includes ending impunity in the face of *realpolitik* and power differentials at the international level.

Discussions of resistance to complex forms of state criminality range from the formal to the informal, from the individual to the local to the state level, and to the international level. At the individual level, there are cases of

dissenters and whistleblowers, as examples. At the local, organizational, and state levels, there are NGOs, media outlets, social movements, and political and civil groups that attempt to change existing conditions and expose state criminality. At the international level, there are political pressures from other states and international intergovernmental organizations and international laws that serve as deterrents to state crimes or after-the-fact accountability. Within each of these categories, resistance fluctuates and occurs at various points of time in relation to specific state criminality and the resources of organizations. It is through all of these mechanisms that awareness of the harms produced and illegalities of states' actions can come to be recognized and generate appropriate responses.

Future Directions

The term *cosmopolitan* is an ideology expressing that all of humanity belongs to a single moral community: an ideology for “global citizenship.” Such an ideology is not new. For centuries philosophers have speculated on the conditions for achieving enduring peace between nations, one of which includes the notion of cosmopolitanism or universalism. Consider Kant’s thoughts in his 1795 essay *Perpetual Peace*, where he suggests *ius cosmopolitanicum* (cosmopolitan law/right) be a guiding principle to protect people from war – grounded in the belief of universal hospitality. In more contemporary times, the idea of cosmopolitanism reemerged in both formal and informal political realms that included the development of a United Nations, the global citizen, and an international community. While there remains a distinct contradiction between this ideology and the current state of international affairs, one vision for the future would be a reality of the ideal type where global citizens become the primary consideration in foreign policy and state actions. There is no magic answer or quick policy fix for state crimes: the worst crimes. It is an ongoing problem that requires diligence, commitment, and a change in ideology, praxis, and relations.

Related Entries

- ▶ [Crimes of Globalization](#)
- ▶ [Crimes of the Powerful](#)
- ▶ [International Responses to Victims in Criminal Justice](#)
- ▶ [Resistance to State Crime](#)
- ▶ [State-Corporate Crime](#)

Recommended Reading and References

- Chambliss W (1990) State organized crime. *Criminology* 27(2):183–208
- Chambliss W, Michalowski R, Kramer R (eds) (2010) *State crime in a globalized age*. Willin Press, Devon
- Green P, Ward T (2004) *State crime: governments, violence and corruption*. London: Pluto Press
- Hagan J (2010) *Who are the criminals? The politics of crime policy from the age of Roosevelt to the age of Reagan*. Princeton University Press, Princeton
- Michalowski R, Kramer RC (eds) (2006) *State-corporate crime: wrongdoing at the intersection of business and government*. Rutgers University Press, Piscataway
- Ross JI (ed) (2000) *Varieties of state crime and its control*. Criminal Justice Press, Mosney
- Rothe DL (2009) *State criminality: the crime of all crimes*. Lexington/Roman and Littlefield, Manheim
- Rothe DL, Mullins CW (eds) (2010) *State crime, current perspectives*. Rutgers University Press, Piscataway
- Stanley E, McCulloch J (2012) *Resistance to state crime*. Routledge Press, London
- Sutherland E (1939) *White collar criminality*. Presidential address to the American society of sociology. Reprinted 1940. *American Sociological Review* 5:1–12

State-Corporate Crime

Elizabeth A. Bradshaw¹ and Ronald C. Kramer²
¹Department of Sociology, Anthropology and Social Work, Central Michigan University, Mt. Pleasant, MI, USA

²Department of Sociology, Western Michigan University, Kalamazoo, MI, USA

Overview

State-corporate crime has been defined as the illegal or socially injurious actions that occur

when one or more institutions of political governance pursue a goal in direct cooperation with one or more institutions of economic production and distribution (Kramer et al. 2002). Originating from a series of papers and articles produced by Ronald Kramer and Raymond Michalowski in the early 1990s (see Michalowski and Kramer 2006), the concept of state-corporate crime unites two parallel streams of criminological literature to draw attention to the mutually dependent interorganizational relationships between governments and private corporations. Scholarship on state-corporate crime seeks to breach the conceptual wall between economic crimes and political crimes in order to create a new lens through which criminologists can examine the ways illegal acts and social harms often emerge from intersections of economic and political power. While the concept of state-corporate crime could be applied to illegal or other socially injurious actions in societies ranging from private production systems to centrally planned political economies, most of the early research focused on state-corporate crimes within the private production system of US capitalism. State-corporate crimes within a global capitalist economy involve the active participation of two or more organizations, at least one of which is in the civil sector and one of which is in the state sector.

Antecedents

The concept and theory of state-corporate crime evolved within the broader tradition of white-collar crime research in the field of criminology, and it introduces a third form of organizational crime to examine in addition to the separate focus on either corporate or state crime. The roots of the concept can be traced to Edwin Sutherland's analysis of corporate offenses in his classic work, *White-Collar Crime* (1949). Over time, two divergent approaches to white-collar crime research emerged. One approach embraced the occupational dimension of white-collar crime, while the other stressed the importance of focusing on organizational offenders. Occupational

crimes consist of offenses committed by individuals for their own personal gain in an occupational setting, while corporate crimes are illegal or socially injurious acts committed by corporate officials acting on behalf of the corporation for the benefit of the corporation (Clinard and Quinney 1973). The organizational turn in white-collar crime research retained the importance Sutherland placed on studying crimes of the powerful and paved the way for an organizational analysis of corporate crime (Clinard and Yeager 1980; Kramer 1982).

Although the broader study of white-collar crime also includes offenses committed by government officials, the concept of state crime more explicitly draws attention to the illegal or harmful actions committed by the state as an organizational entity. Highlighting the important relationship between politics and the economy, state crime is often motivated by the need for capital accumulation by modern nation states. Defined as acts prohibited by criminal law and committed by state officials in the pursuit of their jobs as representatives of the state, "state-organized crime" includes acts such as piracy, smuggling, and illegal spying on citizens (Chambliss 1989). As with corporate offenses, deviant behavior by state actors can take the form of both crimes of commission and of omission (Kauzlarich et al. 2003).

To sum up, after the organizational turn in the field of white-collar crime, two nearly independent bodies of research had developed. Theory and research in the area of corporate crime had concentrated primarily on organizational deviance within private business corporations. Paralleling that work but seldom intersecting with it, others had examined crimes by governments. Despite its ubiquity, the structural relations between corporate and governmental organizations had been relatively peripheral to the study of organizational crime. Kramer and Michalowski suggested that many forms of organizational deviance are generated at the interstices of corporations and government and they introduced the term state-corporate crime to denote these particular forms of organizational deviance.

In addition, rather than limiting inquiry to acts in violation of criminal law, white-collar and organizational crime studies broadened the scope of criminological concern to include violations of other forms of law, such as regulatory law for the study of corporate crime and international law for the analysis of state crime. With regard to crimes of the state, since governments select which behaviors to criminalize, socially harmful actions committed by the state are often exempt from any legal scrutiny. For this reason, some criminologists argue that it is necessary to move beyond the confines of law in favor of approaches that focus on violations of human rights or other forms of social harm (Schwendinger and Schwendinger 1970; Hillyard et al. 2004). The study of state-corporate crime fully supports this expansion of criminological inquiry to encompass violations of regulatory and international law, as well as other legal, but socially injurious behavior, as legitimate phenomena for analysis.

Development

Initially, four case studies (Kramer 1992; Kauzlarich and Kramer 1993; Aulette and Michalowski 1993; Matthews and Kauzlarich 2000) established the foundation for understanding two distinct forms of deviant state-corporate interactions: state-initiated corporate crime and state-facilitated corporate crime. State-initiated corporate crime occurs when corporate entities undertake organizational deviance at the behest or direction of, or with the tacit approval of, government institutions. Illustrating this concept, the case studies of the *Challenger* explosion and the US production of nuclear weapons both draw attention to explicit role of the state in initiating deviant interorganizational actions. Alternatively, state-facilitated corporate crime occurs when government institutions of social control fail to restrain deviant business actions, either due to direct collusion between corporations and government or because they adhere to common goals that would be thwarted by forceful regulation. Both the investigation of the fire at the Imperial Food Products chicken processing

plant and the crash of *ValuJet* 592 highlight the failure of the state to effectively constrain corporate criminality.

The original development of the concept of state-corporate crime stems from a case study of the deviant state and corporate intersections that resulted in 1986 *Challenger* space shuttle explosion. At first glance, the explosion of the *Challenger* appeared to simply be an accident. However, Kramer's (1992) examination demonstrated how the interactions between a government agency, the National Aeronautics Space Administration (NASA) and Morton Thiokol, Inc. (MTI), a private business corporation, led to risky decisions and unsafe actions that resulted in the death of six astronauts and one school teacher. While the technical cause of the disaster was the failure of the O-ring seal on the solid rocket booster, larger structural and organizational forces increased the likelihood that a dangerous outcome would occur. Budgetary compromises, political pressure to launch, and what (Vaughn 1996) later identified as the "normalization" of deviant practices within the organizational culture at NASA led to a catastrophic explosion.

Also focusing on the historical, political, and contextual factors contributing to deviant state and corporate interactions is Kauzlarich and Kramer's (1993) study of the environmental damage caused by the US nuclear weapons manufacturing industry. For more than 50 years, the Department of Energy (DOE) and the Atomic Energy Commission contracted with private multinational corporations such as Westinghouse, DuPont, General Electric, and Martin Marietta to produce nuclear weapons. While DOE owned the equipment and oversaw the production of nuclear weapons and materials around the country, the corporations were responsible for daily manufacturing operations. Producing nuclear weapons results in enormous amounts of radioactive and nonradioactive hazardous waste. This waste was improperly disposed of causing irreversible environmental damage. For example, two of the most environmentally harmful nuclear weapon facilities, the Hanford facility and the Savannah River plant which both produce

plutonium and tritium, have released billions of gallons of liquid waste contaminating the local air, soil, groundwater, rivers, and watersheds. Encouraged by Cold War cultural beliefs and lacking interorganizational oversight, the nuclear weapons industry placed production of defense materials above the environmental consequences of war-head production (Kauzlarich and Kramer, 1993).

Another foundational case study of state-corporate crime by Aulette and Michalowski (1993) details the 1991 fire at the Imperial Food Products chicken processing plant in Hamlet, North Carolina, that resulted in the deaths of 25 workers and injured an additional 56. Although the technical cause of the fire was a rupture in the hydraulic line near the deep fryer that sent a wave of fire throughout the plant, it was discovered that Imperial had deliberately locked the fire doors to prevent employee theft, thereby denying them access to a safe exit. Beyond the actions of Imperial, a complex pattern of regulatory failure was revealed. Facilitated by a long history of privileging business interests over labor, North Carolina's neglect to fund the state's Occupational Safety and Health Program severely weakened regulatory oversight designed to protect workers. In contrast to the examinations of the *Challenger* explosion and the US manufacturing of nuclear weapons that demonstrate the direct role of the state in the commission of corporate wrongdoing, the Hamlet fire study identified a different type of relationship in which the state indirectly creates the conditions for corporate crime to occur (Aulette and Michalowski 1993).

Matthews and Kauzlarich's (2000) examination of the crash of *ValuJet* Flight 592 in the Florida Everglades on May 11, 1996, helps to further define the role of the state in facilitating corporate crime. While the explicit cause of the crash that killed all 105 passengers and five crewmembers was the explosion of oxygen generators in a cargo compartment that resulted in fire, government investigations also identified the failure of both *ValuJet* and *SabreTech* (an airline maintenance company) to comply with numerous regulations as important factors. Dually tasked with the conflicting mandates of regulating the safety of the airline industry while

simultaneously promoting it, the Federal Aviation Administration (FAA) refused to implement safeguards and guidelines that could have protected passengers in favor of the economic interests of the airline industry. By ignoring two specific recommendations by the National Transportation Safety Board (NTSB) to place smoke detectors in cargo holds (the exact area the fire started in *ValuJet* 592), as well as to reclassify cargo holds to prevent the spreading of fire to the rest of the plane, the FAA indirectly sets the stage for the crash to occur (Matthews and Kauzlarich 2000).

A Theory of State-Corporate Crime

State-corporate crime has three useful characteristics as a sensitizing concept. First, it refutes the notion that organizational deviance is a discreet act by illuminating the relationships between social institutions. Second, by embracing the relational character of the state, the concept of state-corporate crime demonstrates how the horizontal interactions between political and economic institutions contain the potential for illegal and social injurious actions to occur (Wonders and Solop 1993). Finally, adopting a relational approach to the state not only allows for a consideration of horizontal interactions but of the vertical relationships between different levels of organizational action: political-economic, organizational, and interactional.

Corresponding to each level of social action are three theoretical approaches to the study of corporate crime: differential association theory, organizational theory, and political economy. Developed by Edwin Sutherland, differential association theory is a social psychological theory that seeks to establish the processes by which individuals learn deviant behavior. While differential association has been criticized for failing to consider the institutional context, the organizational perspective helps to link the external political-economic environment with the work-related thoughts and actions of individuals occupying structural positions in the social organization of work. Motivated by pressure to achieve organizational goals in

a competitive setting, there exists an inherent inducement for the organization to participate in crime (Kramer 1982). The perspective of political economy demonstrates the implicit criminogenic potential of pressures for profit maximization within capitalist markets. By drawing attention to the structural relations between political and economic institutions, the political-economic perspective recognizes that many crimes of corporations and of governments are ultimately crimes of capital arising from the ownership or management of the accumulation process (Michalowski and Kramer 2006).

Independent of one another, each of these three theoretical approaches is inadequate to explain organizational deviance. In an effort to overcome these shortcomings, state-corporate crime scholars propose an integrated theoretical model of organizational crime that combines the three levels of analysis with three catalysts for action: motivation, opportunity, and control (Michalowski and Kramer 2006). The first catalyst for action concerns goal attainment. As the emphasis on goal attainment by political-economic institutions, organizations, and individuals increases, corporations and state agencies become more susceptible to engaging in organizational deviance. The second catalyst for action – opportunity – assumes that organizational deviance is more likely where legitimate means are scarce relative to goals. Finally, the third catalyst for action examines the presence or absence of social control at all three levels of analysis. Organizations subjected to a high operationality of social controls are more likely to cultivate organizational cultures that favor compliance with laws and regulations, and those organizations that are not subject to such controls are more likely to develop cultures of resistance. By investigating the linkages between levels of analysis and catalysts for action, a more nuanced understanding of state-corporate crime can potentially be developed.

Recent Research Trends

Beyond its early incarnations, the concept and theory of state-corporate crime has been

increasingly used to understand diverse forms of organizational deviance engaged in by state and corporate actors. In addition to the early empirical and theoretical research on state-corporate crime, *State-Corporate Crime: Wrongdoing at the Intersection of Business and Government* (Michalowski and Kramer 2006) includes case studies on topics such as globalization, women and state-corporate crime, corporate facilitation during the Holocaust in Nazi Germany, the National Highway Transportation Safety Administration (NHTSA), Ford Motor Company and the Bridgestone-Firestone tread separation case, the Exxon Valdez oil spill, Enron era economics and economic democracy, violations of the treaty rights of indigenous peoples, the invasion of Iraq, as well as the role of Halliburton in Iraq. In addition to these early case studies, more recent research on state-corporate crime has emerged, particularly concerning international and environmental issues.

International Issues

A number of criminologists have paid increasing attention to the changing nature of state power in the global neoliberal economy and have begun to focus on the international arena. One theme among numerous recent case studies of state-corporate crime concerns the organizational deviance that has occurred due to the US invasion of Iraq. A second prominent theme is the relationship between private military firms and governments and the implications this has for state-corporate deviance. Moreover, other case studies on the state and corporate interactions surrounding international conflict over scarce natural resources, as well as the antecedents of the most recent global financial meltdown, have also been undertaken.

Kramer and Michalowski (2005) argue that the 2003 invasion and occupation of Iraq were illegal under international law and can therefore be classified as an instance of state crime. Laying the foundation for international relations is the United Nations (UN) charter, which at its core prohibits the use of aggressive force between nations with only two exceptions: self-defense and humanitarian interventions. Initially, the

Bush administration sought to implicate Iraq in the September 11th terrorist attacks in an effort to justify war as self-defense. After being forced to concede that there was in fact no connection between Iraq and 9/11, the Bush administration argued for preemptory war due to allegations that Iraq possessed weapons of mass destruction, thereby constituting a direct threat requiring defensive military action. As a final strategy, the Bush administration attempted to convince the UN Security Council that it was necessary to go to war in order to protect Iraqi citizens from human rights abuses by Saddam Hussein. Despite all the efforts to legitimate war in Iraq, none met the legal requirements of international law and were unable to receive necessary approval from the UN Security Council. In the end, the March 2003 decision to go to war in Iraq was undertaken in clear violation of international law and paved the way for continued state-corporate crimes as a result (Kramer and Michalowski 2005).

Whyte (2007) explores how the overarching principle behind the US invasion and occupation of Iraq was the creation of a new rule of law based on the opening up of the economy to privatization by Western, and particularly US, corporations in breach of international law. Facilitating this transition, the Coalition Provisional Authority (CPA) was given the power to rebuild the Iraqi economy over a 14-month period. Guided by neoliberal economic principles, the CPA spent over \$20 billion in Iraqi oil revenue to restructure the economy, much of which was disbursed to US corporations with scant oversight or documentation. Through 100 legally binding administrative orders and decrees, the CPA set the foundations of Iraq's new economy, justice system, and political structure based on the idea of "trickle down" economics. The Development Fund of Iraq (DFI) provided the CPA with immediate money from Iraqi oil revenues to be dispersed for the humanitarian needs of the people. Since the DFI funds were disbursed in bundles of cash, CPA transactions did not leave a paper trail and contracts frequently went undocumented entirely. Moreover, the Bush administration in conjunction with the CPA took steps to thwart audits and investigations into the dispersal of funds. A US

government appointed auditor has since established that an unknown sum of the DFI funds have disappeared as a result of bribery, overcharging, embezzlement, product substitution, bid rigging, and false claims. These forms of state-corporate corruption under the CPA have been essential in achieving neocolonial dominance in occupied Iraq (Whyte 2007).

In the course of the occupation and invasion of Iraq, privatized security has been deployed on a scale never before seen. Rather than abdicating its authority to private military firms (PMFs), Welch (2009) contends that governments work in direct cooperation with them, creating a situation of fragmented power. Although it has enjoyed close ties with the Bush administration, even recruiting former officials from the CIA and the Pentagon, Blackwater has received criticism for its lethal actions in Iraq. For example, in May 2007, Blackwater employees opened fire on the streets of Baghdad twice in 2 days, including a standoff with Iraqi security forces. Another incident, labeled "Baghdad's Bloody Sunday," occurred on September 16, 2005, when Blackwater guards shot and killed between 8 and 20 innocent civilians and wounded dozens of others in Nisour Square. A documented repeat offender, Blackwater has also been investigated for at least six other episodes of excessive force. Despite this pattern of criminality, the US government has not only neglected to prosecute Blackwater employees and other contractors but has extended immunity from wrongdoing altogether with respect to the reconstruction of the Iraqi economy. In this manner, the decoupling of police and government forces enables private contractors to escape accountability for war crimes and human rights abuses (Welch 2009).

Part of a broader trend of privatizing military functions, private military companies (PMCs) today have become a legitimate industry involved in a wide range of activities including protecting governmental and nongovernmental organizations during humanitarian missions, in addition to protecting corporate interests such as the extraction of oil and mining. Whyte (2003) argues that the unfolding of the PMC market can be understood as transference of law from

international prohibition treaties and national criminal law to civil contracts as the principal means of legal regulation. Far from a reduction of state sovereignty in the era of neoliberal globalization, private military markets are dependent on the consent and support of governments for their livelihood. Moreover, governments have also come to recognize the benefits to be gained by forging a state-corporate alliance in the private military market. Accepting the proliferation of PMCs as not only inevitable but also desirable, the United Kingdom has embraced a “soft-touch” regulatory mix that strengthens state-corporate military relations in a number of ways. First, such an approach enhances British industrial competitiveness in weapons manufacturing and other industries. Second, encouraging the growth of PMCs also allows governments to monitor and influence remote territories while remaining ostensibly neutral, thereby conducting foreign policy by proxy. Finally, the difficulty of prosecuting PMCs under international law and the reluctance of the UK government to support the UN convention on mercenaries create a law-accountability gap. By expanding the PMC market, the opportunity structure for state-corporate crime is increased as states and corporations are able to engage in high risk or politically sensitive conflicts while evading accountability for their actions. Absent any new criminal legal controls, state-corporate crime in the PMC market is only likely to accelerate (Whyte 2003).

Providing a revised theoretical framework of state-corporate criminality, Rothe and Ross (2010) analyze how anomie and social disorganization, resulting in a lack of regulation, are significant factors in explaining the criminal propensity of private military companies (PMCs). Private forces (such as Bechtel, Blackwater, CACI International, DynCorp, Halliburton and subsidiary Kellogg, Brown, and Root, Logo Logistics, and Titan) provide a wide range of services including direct tactical military assistance, military consulting (strategic advisory and training), and logistic, intelligence, and maintenance services to armed forces. Within each of these sectors, however, there are variations in the types of crimes committed including murder,

fraud, and war profiteering. PMCs are not held to the same rules of engagement as the military and have an unclear legal status that is undefined by international law. Lacking internal and external constraints at every level, PMCs operate in anomic (lawless or normless) conditions that cultivate criminogenic behavior. Social disorganization occurs when communities are no longer able to create and enforce informal mechanisms of social control. As they operate in war ravaged areas, PMCs function in disorganized environments with uncertain mandates, high rates of employee turnover, and little social support. These factors converge to produce an environment in which regulation and social control breaks down and PMCs are free to engage in deviant behavior without consequence (Rothe and Ross 2010).

In the face of recent US actions that violate international law such as the war on terror, the war in Afghanistan, and the war in Iraq, a large segment of the citizenry nevertheless shows support for these policies. Part of the ideology of military aggression, Klein and Lavery (2011) argue, is perceived in-group victimization. A corporately owned political and cultural institution, the mainstream media functions as a tool by which politicians emphasize and validate the national experience of collective victimization. In an attempt to legitimate their illegal actions, elites and organizations use the media to amplify their claims of victimization and thereby neutralize criticism. At least in part, fear of being a terror victim seems to be related to the content of news media. Moreover, there also seems to be support for the notion that terrorism victimization in the media can contribute to public acceptance of state-corporate criminality (Klein and Lavery 2011).

Confronted with widespread political and economic violence for much of the past century, the experience of the Democratic Republic of Congo (DRC) cannot solely be explained by internal factors experienced in the transition from colonial to postcolonial rule. Rather, Mullins and Rothe (2008) argue that neighboring states, transnational corporations, and international organizations all play a significant role in the continuing

violence within the DRC. Seeking control of large concentrations of valuable mineral fields including gold, diamonds, silver, copper, and coltan, transnational corporations have been more than willing to negotiate access with warlords and militias controlling the hinterlands despite widespread human rights abuses. While demand from European markets is an important factor spurring participation by transnational corporations and their Western trading partners, states and international organizations also play a crucial role in facilitating the sale of illegally misappropriated resources. Capitalizing on the chaos surrounding the Second Congolese War in 1995 by enacting patterns of illegal resource appropriations particularly for gold and diamonds, Uganda acknowledged to the World Bank (WB) that production levels might reflect exports usurped from the DRC. The WB and the International Monetary Fund praised Uganda and Rwanda for their increases in production, despite the illegal means by which it was achieved. While it is largely recognized that there is no indigenous production of gold in Uganda, Metalor Technologies, a Swiss refinery, has been responsible for purchasing the Congolese gold from Ugandan sources. Although it had been demonstrated that over 100 companies from over a dozen countries are benefiting from illegal mineral exploitation, there has been reluctance on behalf of governments to intervene. The institutional structure of the larger international community therefore plays an important role in perpetuating the political and economic violence in countries engulfed by genocidal civil war such as the DRC (Mullins and Rothe 2008).

As an investigator for a 2002 Royal Commission examining fraud within the Dutch construction industry, Van Den Heuvel (2005) helped reveal that rather than an isolated incident of collusion, the entire sector had engaged in illegal practices including fraud, undercutting the market, unjustified subsidies, monopolization resulting in higher prices, and bribery of politicians and public servants. At the core, the Commission identified the industry-wide pandemic to be due to multiple forms of collusion or a secret agreement for a fraudulent purpose.

One form of collusion occurred between contractors themselves, such as illegal price fixing. A second type involved collusion between authorities and contractors, such as favoring particular contractors over others. Finally, a third type of collusion identified by the Commission occurred at the individual level in the bribing of public servants, for example, providing generous gifts and favors. The Commission concluded that these forms of collusion within the Dutch construction industry were so interconnected that they constituted a culture that placed contractors above the law and in control of authorities. Because of the pervasive collusion between contractors and authorities, it seems necessary to enact stricter rules governing state-corporate interorganizational relationships (Van Den Heuvel 2005).

Liederbach (2010) contends that while the state established the framework for the 2008 mortgage default crisis, corporate malfeasance was at its core. From the 1930s through the 1970s, there were stringent government regulations on mortgages which restricted lending practices such as issuing variable rate loans, capped interest rates, and prohibited partial monthly payments (known as negatively amortized mortgages). However, statutory changes and deregulation by Congress during the early 1980s removed these restrictions. Coasting on the economic success of the 1990s, an effort to increase national homeownership led by the US Department of Housing forced the government-sponsored entities, Fannie Mae and Freddie Mac, to engage in riskier subprime mortgage markets since the targeted borrowers could not qualify for prime loans. Setting the conditions for the crisis to thrive, during the early 2000s, the Federal Reserve expanded the availability of credit by setting and maintaining surprisingly low interest rates and decreasing the amount of money banks were required to keep in reserve. All of these factors culminated in an economic meltdown that resulted in millions of home foreclosures and an unprecedented \$750 billion taxpayer-funded bailout to rescue distressed financial institutions (Liederbach 2010).

Environmental Issues

Another emerging trend within the research on state-corporate crime is case studies examining the environmental harm resulting from organizational deviance. Of increasing significance are the state and corporate interactions contributing to the critical problem of global warming and climate change.

Smandych and Kueneman (2010) contend that Canadian officials and oil corporations are guilty of manipulating environmental regulations as well as the practice of “greenwashing:” disinformation provided by an organization attempting to present an environmentally friendly public image, in this case surrounding the Alberta tar sands project. Criticized for its concentrations of toxic waste and a major source of greenhouse gas emissions, the aggressive state-corporate development of the Alberta tar sands project began in 1990s. Effectively surrendering sovereignty over the nation’s resources, the 1989 Canada-USA Free Trade agreement obliges Canada to share its oil and gas resources with the USA. The two means of extracting the oil from a mixture primarily comprised of sand are strip mining and underground “in situ” mining, both of which are environmentally disastrous leading to the devastation of pristine boreal forests, contamination of the Athabasca watershed, and depletion of natural resources. In this regard, the USA accrues the benefits while Canada suffers pollution of its land, air, and water. Working hand in hand with the oil industry, specifically US-based multinational oil corporations, the conservative governments of the Province of Alberta and the federal government of Canada have suppressed attempts to define the harm caused by the tar sands as criminal activity (Smandych and Kueneman 2010).

Lynch et al. (2011) applied the concept of state-corporate crime to global warming by examining the politicization of global warming under the G.W. Bush administration. Although industry leaders are often selected for governmental positions, appointing corporate leaders from the oil and mining industries to crucial environmental policy positions was a dominant trend in the Bush administration that strengthened state-corporate ties and deterred action

on climate change. Despite the overwhelming consensus of climate change scientists, the Bush administration colluded with corporations in the fossil fuel industries in an effort to discredit and suppress science on the human causes of global warming. White House officials deliberately sought to undermine science on global warming in a number of ways including blocking publications, editing government reports, altering federal policy, and pressuring climate change scientists to delete references to global warming and climate change in government-sponsored research. Furthermore, the corporate strategy for undermining information on global warming was to fund and create front organizations to disseminate misinformation about climate change to the public. In this manner, the policies of the Bush administration privileged the interests of the fossil fuel industry over the interests of the public (Lynch et al. 2011).

Four forms of state-corporate crime shape the social and environmental harms caused by global warming. Kramer and Michalowski (2012) argue that state and corporate actors produce these harms by (1) denying that global warming is caused by the actions of humans (anthropogenic), (2) thwarting attempts to reduce greenhouse emissions, (3) excluding from the political arena ecologically just adaptations to climate change, and (4) responding to the social conflicts that arise as a result of climate change with militarism and violence. While anthropogenic global warming is the result of over 200 years of industrialization and fossil fuel consumption, it is possible to identify state-corporate relationships that caused knowable and predictable harm and that could have been avoided. Therefore, the failure of state institutions to mitigate or reduce carbon emissions in the private and public sectors should be understood as a state-corporate crime of omission. More than just a failure to act, however, the orchestrated denial of climate change despite overwhelming scientific evidence to the contrary constitutes a state-corporate crime of commission. Designed to cast doubt on the evidence for anthropogenic global warming, the global warming denial counter-movement has been directed, organized, and funded by

corporations and conservative think tanks. States and corporations have also refused to seriously consider socially just adaptation policies despite the increasing number of social conflicts resulting from climate change. This too, it is argued, constitutes a state-corporate crime of omission (Kramer and Michalowski 2012).

Conclusion

The study of political economy has been around for a long time, and the idea that corporations and states often act together in ways that have serious social consequences is not new. However, the concept and theory of state-corporate crime filled a gap in the evolving study of organizational crime within criminology. By examining social harms that result from the interaction of political and economic organizations, scholarship on state-corporate crime has made an important contribution to the field. The concept has generated a substantial body of research, and some argue, “the approach developed by state-corporate crime scholars is a significant advance toward developing a powerful integrated theoretical model” (Green and Ward 2004, p. 51). Judged by the growing number of criminologists who find the concept state-corporate crime useful in their work, research and theory in this area have the potential to make additional contributions to the study of organizational offending, particularly with regard to international and environmental crimes.

Related Entries

- ▶ [Corporate Crime Control](#)
- ▶ [Corporate Crime Decision-Making](#)
- ▶ [Crimes of Globalization](#)
- ▶ [Crimes of the Powerful](#)
- ▶ [Green Criminology](#)
- ▶ [Nuclear Weapons and State Crime](#)
- ▶ [Political Crime](#)
- ▶ [Resistance to State Crime](#)
- ▶ [State Crime](#)
- ▶ [Victims of State Crime](#)

Recommended Reading and References

- Aulette J, Michalowski R (1993) Fire in Hamlet: a case study of state-corporate crime. In: Tunnell K (ed) *Political crime in contemporary America: a critical approach*. Garland, New York, pp 171–206
- Chambliss W (1989) State-organized crime- the American society of criminology, 1988 presidential address. *Criminology* 27:183–208
- Clinard M, Quinney R (1973) *Criminal behavior systems: a typology*, 2nd edn. Holt, Rinehart and Winston, New York
- Clinard M, Yeager P (1980) *Corporate crime*. The Free Press, New York
- Friedrichs D (2010) *Trusted criminals: white collar crime in contemporary society*, 4th edn. Wadsworth, Belmont
- Green P, Ward T (2004) *State crime: governments, violence and corruption*. Pluto Press, London
- Hillyard P, Pantazis C, Tombs S, Gordon D (eds) (2004) *Beyond criminology: taking harm seriously*. Pluto Press, London
- Kauzlarich D, Kramer R (1993) State-corporate crime in the US nuclear weapons production complex. *J Hum Justice* 5:4–28
- Kauzlarich D, Mullins C, Matthews R (2003) A complicity continuum of state crime. *Contemp Justice Rev* 6:241–254
- Klein J, Lavery C (2011) Legitimizing war by victimization: state-corporate crime and public opinion. *Crime Law Soc Chang* 56:301–317
- Kramer R (1982) Corporate crime: an organizational perspective. In: Wickman PW, Daily TB (eds) *White collar and economic crime*. Lexington Books, Lexington, pp 75–94
- Kramer R (1992) The space shuttle challenger explosion: a case study of state-corporate crime. In: Schlegel K, Weisburd D (eds) *White collar crime reconsidered*. Northwestern University Press, Boston, pp 214–243
- Kramer R, Michalowski R (2005) War, aggression and state crime: a criminological analysis of the invasion and occupation of Iraq. *Br J Criminol* 45:446–469
- Kramer R, Michalowski R (2012) Is global warming a state-corporate crime? In: White R (ed) *Climate change from a criminological perspective*. Springer, New York
- Kramer R, Michalowski R, Kauzlarich D (2002) The origins and development of the concept and theory of state-corporate crime. *Crime Delinq* 48:263–282
- Liederbach J (2010) Pass the trash: the mortgage default crisis as state-corporate crime. *Sociol Crime Law Devianc* 15:17–41
- Lynch M, Burns R, Stretesky P (2011) Global warming and state-corporate crime: the politicization of global warming under the bush administration. *Crime Law Soc Chang* 54:213–239
- Matthews R, Kauzlarich D (2000) The crash of ValuJet flight 592: a case study in state-corporate crime. *Sociol Focus* 33:281–298

- Michalowski R, Kramer R (eds) (2006) *State-corporate crime: wrongdoing at the intersection of business & government*. Rutledge University Press, New Brunswick
- Mullins C, Rothe D (2008) Gold, diamonds and blood: international state-corporate crime in the Democratic Republic of the Congo. *Contemp Justice Rev* 11:81–99
- Rothe D, Ross JI (2010) Private military contractors, crime, and the terrain of unaccountability. *Justice Q* 27:593–617
- Samandych R, Kueneman R (2010) The Canadian-Alberta tar sands: a case study of state-corporate environmental crime. In: White R (ed) *Global environmental harm: criminological perspectives*. Willan Publishing, Portland, pp 87–109
- Schwendinger H, Schwendinger J (1970) Defenders of order or guardians of human rights? *Issues Criminol* 5:123–157
- Sutherland E (1949) *White collar crime*. Dryden Press, New York
- Van Den Heuvel G (2005) The parliamentary enquiry on fraud in the Dutch construction industry: collusion as concept between corruption and state-corporate crime. *Crime Law Soc Chang* 44:133–151
- Vaughn D (1996) *The challenger launch decision: risky technology, culture, and deviance at NASA*. University of Chicago Press, Chicago
- Welch M (2009) Fragmented power and state-corporate killings: a critique of blackwater in Iraq. *Crime Law Soc Chang* 51:351–364
- Whyte D (2003) Lethal regulation: state-corporate crime and the United Kingdom government's new mercenaries. *J Law Soc* 30:575–600
- Whyte D (2007) The crimes of neo-liberal rule in occupied Iraq. *Br J Criminology* 47(2):177–195
- Wonders N, Solop F (1993) Understanding the emergence of law and public policy: toward a relational model of the state. In: Chambliss W, Zatz M (eds) *Making law*. Indiana University Press, Bloomington, pp 204–225

STDs

- ▶ [Drug Abuse and Alcohol Dependence Among Inmates](#)

Stigma

- ▶ [Rational Choice, Deterrence, and Crime: Sociological Contributions](#)

Sting Operations

- ▶ [Law of Undercover Policing](#)

Stolen Auto

- ▶ [Motor Vehicle Theft](#)

Stolen Vehicle

- ▶ [Motor Vehicle Theft](#)

Stop Now and Plan (SNAP[®]) Model

Leena Augimeri¹, Margaret Walsh², Kathy Levene², Karen Sewell² and Erin Rajca²

¹Centre for Children Committing Offences, Child Development Institute, Toronto, ON, Canada

²Child Development Institute, Toronto, ON, Canada

Synonyms

[Centre for Children Committing Offences \(CCCO\)](#); [Earls court Child and Family Centre \(ECFC\)](#); [Earls court Girls Connection \(EGC\)](#); [SNAP[®]](#); [SNAP[®] Boys](#); [SNAP[®] GC](#); [SNAP[®] Girls](#); [SNAP[®] ORP](#); [Under 12 Outreach Project \(ORP\)](#)

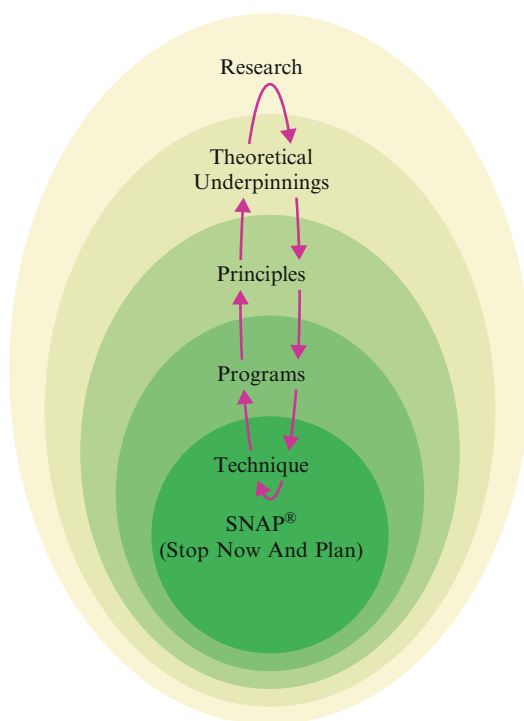
Overview

This entry presents a comprehensive summary of the SNAP[®] (Stop Now and Plan) model including its related programs which are part of a crime prevention strategy for at-risk children under the age of 12. At Child Development Institute (CDI), a multiservice, not-for-profit children's mental health organization, the mandate for two key

evidence-based clinical programs, SNAP[®] Boys and SNAP[®] Girls, and the Centre for Children Committing Offences (CCCO) promotes effective services for these “forgotten” children who have or are at-risk of having police and/or child welfare contact for their disruptive behavior. The comprehensive strategy includes (1) police-community referral protocols; (2) gender-specific risk assessment using Early Assessment Risk Lists (EARLs); and (3) gender-specific interventions, the evidenced-based SNAP[®] programs (for further details, see Koegl et al. 2008). The overarching goal of SNAP[®] is to keep high-risk children in school and out of trouble. SNAP[®] has been found to reduce aggression, delinquency, and antisocial behavior; increase social competency; prevent further and future delinquency; improve academic success by decreasing behavioral issues at school; engage high-risk children and their families in service; increase effective parent management skills; and connect children and parents to community-based resources. Authors discuss key aspects of the SNAP[®] model including SNAP[®] principles, theoretical foundation, model framework, and research summary of findings to date.

Fundamentals of SNAP[®]

Over 27 years, CDI has developed its expertise in responding to children with disruptive behavior problems and their families. In 1985, with the decriminalization of children under the age of 12 in Canada, CDI (with support from the provincial Ministry of Children and Youth Services in Ontario, Canada) developed SNAP[®] programs in response to the need of mental health services for this age group of high-risk children and families. The SNAP[®] model is based on a comprehensive framework (Fig. 1, SNAP[®] Model Framework) for effectively teaching children with serious behavior problems emotion regulation, self-control, and problem-solving skills. Parents also learn SNAP[®] skills as well as researched cognitive-behavioral parenting techniques. Children and families learn how to stop and think in order to find solutions that “make their problems smaller, not bigger.” The



Stop Now and Plan (SNAP[®]) Model, Fig. 1 SNAP[®] model framework

SNAP[®] Model Framework depicts the interconnectedness and relationship of the theoretical underpinnings, principles, programs, and technique and how research plays a role in each of these key areas. The SNAP[®] programs have been developed with the *technique* as the cornerstone of the program components, have been continuously informed by theory and research, and are delivered through adherence to the SNAP[®] principles. This dynamic model provides feedback loops which allow for fluidity between the elements to influence and inform ongoing development. The evidence-based SNAP[®] programs have been adapted to different populations and settings: SNAP[®] Boys, SNAP[®] Girls, SNAP[®] Schools, SNAP[®] for Children with Asperger Syndrome, SNAP[®] for Aboriginal Communities, SNAP[®] Youth Leadership, and SNAP[®] for Youth in Custody.

Technique

There is robust evidence that early childhood interventions focused on enhancing self-control

are likely to bring greater return on investments (Moffitt et al. 2011). Further support for this approach is identified in other investigations which highlight that children tend to be good candidates for learning self-control strategies, especially before the age of 10 (Piquero et al. 2010). The SNAP[®] technique is a cognitive-behavioral emotion regulation, self-control, and problem-solving strategy intended to help children *stop and think before they act* and come up with socially appropriate plans to address their problems – helping to control impulsivity, challenge cognitive distortions, and think about the consequences of their behavior. A key aspect of the SNAP[®] technique is to help children identify their bodies' physiological responses (*body cues*), thoughts, feelings, and triggers (things that make them feel angry/sad/worried) and help them to make the connection between their body cues, feelings, thoughts, and what they can do to effectively regulate arousal levels and help their bodies calm down (*Stop*), so that they can come up with an effective *Plan*. As discussed in the article, *Rolling Out SNAP[®] – An Evidence-Based Intervention: A Summary Of Implementation, Evaluation, and Research* (Augimeri et al. 2011a), SNAP[®] was first developed in the former Earls court Child and Family Centre's day treatment classroom for children with behavioral problems in the late 1970s, and the technique underlies the entire foundation of the SNAP[®] Model Framework (see Fig. 1). This was then formalized with the creation and publication of program manuals (Earls court Child and Family Centre 1990a, b, 2001a, b, 2002; Levene 1998) and trademarked in 1998.

As noted in the SNAP[®] program manuals, there are a number of steps to the SNAP[®] technique that have been mapped onto the image of a stoplight – red light (*Stop*), yellow light (*Now and*), and green light (*Plan*). These steps are used to teach children to regulate their emotions by helping them to calm down (e.g., by taking deep breaths and/or counting to ten) (*Stop*); replace “hard thoughts with cool thoughts” (coping statements, cognitive restructuring) to help them remain calm (e.g., “this is hard but I can do this”) (*Now and*); and generate effective solutions

which meet these three criteria: 1. make their problems smaller instead of bigger; 2. make them feel like a winner; and 3. not hurt anyone, anything or themselves (*Plan*).

Programs

The introduction of the first SNAP[®] program (SNAP[®] Under 12 Outreach Project, now known as SNAP[®] Boys) in Toronto in 1985 was designed specifically to address the gap in services when the age of criminal responsibility in Canada was raised from seven to 12 under the *Young Offenders Act* (YOA) in 1984. Prior to the YOA, the *Juvenile Delinquents Act* (JDA) enacted in 1908 prosecuted children as young as 7 years of age. The YOA placed these children under the responsibility of child welfare legislation versus criminal justice; this remained when Canada replaced the YOA with the Youth Criminal Justice Act (YCJA) in 2003.

SNAP[®] Boys was launched in partnership with Toronto Police Service through provincial funding from the former Ministry of Community and Social Services (today, Ministry of Children and Youth Services). The mandate of SNAP[®] Boys is to serve children under the age of 12 who are engaging in antisocial behaviors who do not legally fall under the purview of the YCJA. SNAP[®] Boys is noted as the most fully developed, longest sustained, empirically based multicomponent intervention specifically for “pre-offender” youth under the age of 12 (Howell 2001, 2003). Its sister program, the SNAP[®] Girls Connection (now known as SNAP[®] Girls), began in 1996 and is the first reported *gender-specific* intervention for girls under the age of 12 with disruptive behavior problems. Both programs are fully manualized and are in various stages of replication worldwide.

Presenting issues of the children admitted into the SNAP[®] programs typically include stealing, lying, mischief, vandalism, aggression, assault, bullying, and truancy. A significant number of these children also experience academic difficulties and comorbid mental health symptomology, such as depressive and anxious behaviors or ADHD (Pepler et al. 2004; Walsh et al. 2002).

Children admitted into the SNAP[®] clinical programs (SNAP[®] Boys and SNAP[®] Girls) have had police contact for their own misbehavior and/or have a score within the clinical range on standardized measures assessing externalizing behavioral issues (aggression, conduct, rule breaking). Primary referral sources include schools, police, child welfare, parents, and other mental health and medical professionals (see Fig. 2).

As noted in the SNAP[®] Logic Model (Fig. 2), assessment is informed by an ecosystemic approach (also a SNAP[®] principle, see Table 1) that takes into account interventions targeting the child, the family, the school, and the community. The Early Assessment Risk List (EARL-20B for boys or EARL-21G for girls), a structured clinical risk/need assessment device for use with aggressive and delinquent children, is also completed to provide a comprehensive framework for evaluating risk factors known to influence a child's propensity to engage in future antisocial behavior. Informed by the ecosystemic assessment, the risk assessment takes into account multi-informant perspectives (child, parent, teacher, and clinician), identifies the unique treatment needs of children and their families, and assists clinicians with treatment planning in order to mitigate these risks.

The SNAP[®] clinical programs (SNAP[®] Boys and SNAP[®] Girls) offer multifaceted services including core and adjunct components which are available to children and families based on their level of risk and need. In the SNAP[®] programs, components are goal oriented, skill focused, and developmentally responsive. Integrated into each component are key skill acquisition training techniques (i.e., role-play, modeling, self-talk) and generalization activities (home practice assignments) to transfer learning of the SNAP[®] technique and SNAP[®] parenting skills from the clinical environment to real-life settings. Following the ecosystemic and EARL assessments, a treatment plan is tailored to the child and family's strengths, risks, and needs. Children and families typically begin with completing the core components. Core components include:

SNAP[®] Children's Group – a gender-specific manualized core component that focuses on teaching children emotion regulation, self-control, and problem-solving skills with a special emphasis on challenging cognitive distortions/thinking errors. Examples of topics covered include introduction to SNAP[®], peer pressure, dealing with anger, and bullying. Children participate in a 13-week SNAP[®] group, occurring once a week for 1.5 h.

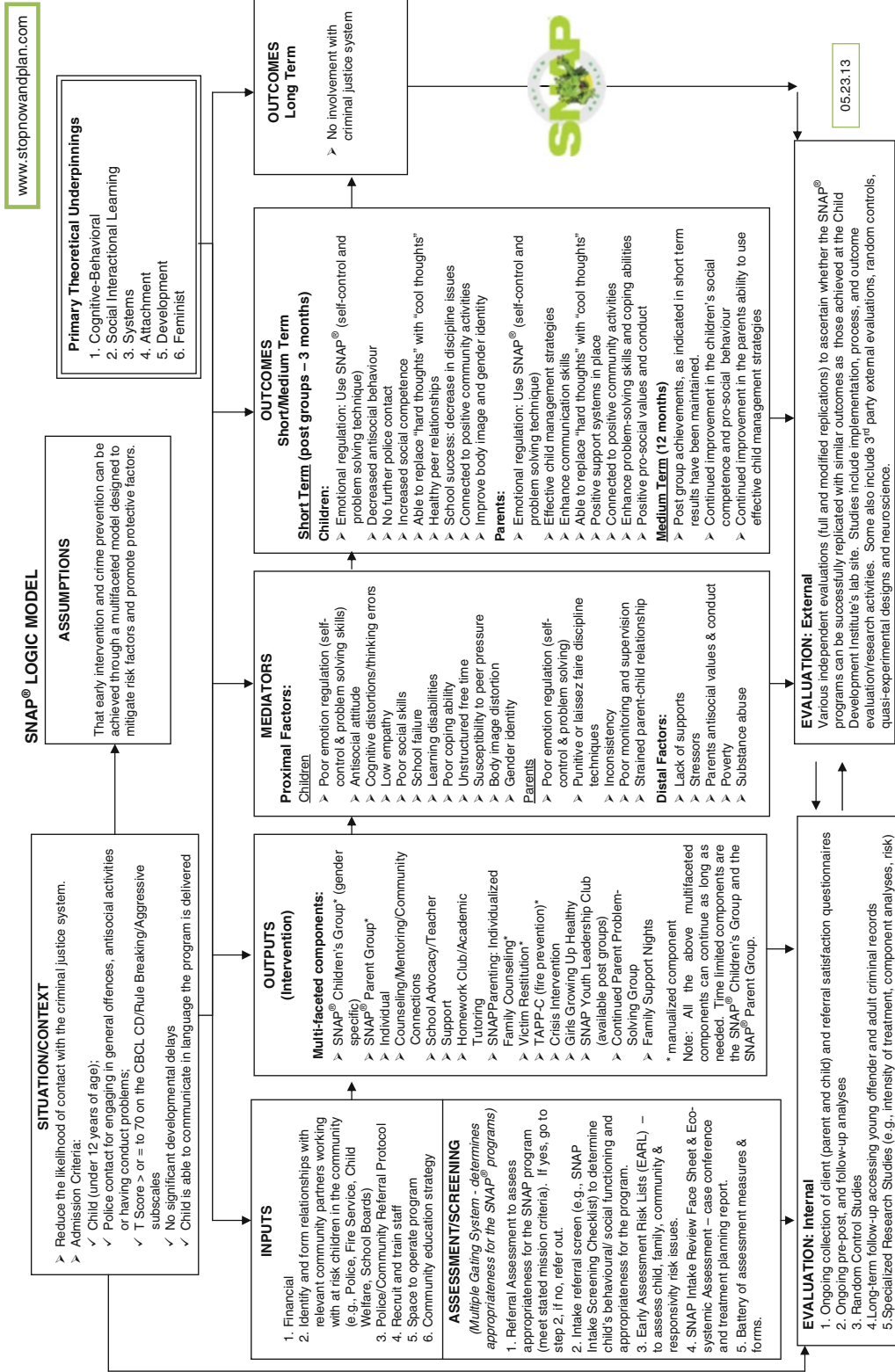
SNAP[®] Parent Group – a manualized core component that runs concurrent with the SNAP[®] children's groups. Parents learn emotion regulation, self-control, and problem-solving skills and effective child management strategies with a special emphasis on challenging cognitive distortions/thinking errors, reducing isolation, and enhancing parent-child relationships. Examples of topics covered include effective communication, positive reinforcement, limit setting and consequences, and family problem solving. Three of the group sessions are joint parent-child sessions where parents and children practice skills together.

Girls Growing Up Healthy (GGUH) – a manualized core component unique to SNAP[®] Girls, this mother-daughter group focuses on enhancing relationship capacity, healthy relationships, and physical and sexual health. Topics covered include preparing for puberty, deconstructing stereotypes, and planning for the future. Caregivers and daughters meet for 8 weeks, once a week for 2 h.

The following adjunct components are offered based on the child's and family's identified goals, strengths, risks, and continued treatment needs:

Individual Counseling/Mentoring/Community Connections – provides children with individualized support with a SNAP[®] worker to reinforce and enhance skills learned in the SNAP[®] children's group and address treatment targets (i.e., social skills, coping ability, cognitive restructuring) and goal attainment. Children can also be matched with volunteers who connect them with structured recreational activities in their community.

Stop Now and Plan Parenting (SNAPP): individualized family counseling – based on our SNAPP Manual, is offered to families who are



Stop Now and Plan (SNAP®) Model, Fig. 2 SNAP® Logic model



Stop Now and Plan (SNAP[®]) Model, Table 1 SNAP[®] principles and indicators

Principle	Indicator
Scientist-practitioner	An interactive science-practice paradigm is sustained by regular cross discipline workgroups to support the high-quality evolution of SNAP [®] program development, outcomes, and research
Client centered	Services are informed by client personal and cultural narratives and goals, ensuring client responsiveness through growth-oriented engagement and alliances
Gender sensitive	Specific gendered factors which account for differential development, learning styles, and trajectories of boys and girls with conduct problems are integrated throughout SNAP [®] programs
Ecosystemic	Each client's ecological system (individual, home, school, community) is assessed to identify and inform strengths, needs, and risk and protective factors and to match them with appropriate service components and treatment intensity
Strength and skill based	Specific, consistent use of positive reinforcement, as part of our evidence-based, cognitive-behavioral practice, promotes and strengthens individual capabilities and the acquisition of primary SNAP [®] skills: emotion regulation, self-control, and problem solving
Continuing services	Continuing needs and commitment to service are regularly and jointly assessed to support and ensure high-risk families, children, and youth are engaged in services
Collaborative	Effective collaborations with appropriate child- and family-oriented services are conducted to ensure service coordination and family support system development during and after SNAP [®] services
Community responsive	SNAP [®] programs are adapted to diverse, cultural, and socioeconomic factors that characterize communities in order to be responsive to social determinants of child and family mental health
Accountable service excellence	Combination of high-quality staff development activities that include consistent supervision, training, integrity, and the attainment of accountable standards assessed through a series of well-developed research, evaluation, fidelity, and quality assurance activities fosters overall service excellence

unable to attend the SNAPP parent group, who need additional assistance/practice with parenting skills, or who need parenting support to address barriers to skill acquisition (i.e., mental health, parent cognitive restructuring, attachment).

School advocacy/teacher support – ensures that children receive the best possible education meeting their individual behavioral and learning needs. Teachers of identified clients are contacted at the start of the program to introduce the program skills and offer behavior management support if needed. Parents are supported in advocating for their children within the school system environment.

Crisis intervention – a service available to assist parents and children involved in the SNAP[®] program in dealing with challenging situations as they arise and/or referral to appropriate crisis services.

TAPP-C (The Arson Prevention Program for Children) – offered to children with fire interest or fire setting as a presenting problem. It involves a fire interest assessment and recommendations, a home safety visit, and education regarding fire by Toronto Fire Services.

Victim restitution – activities that encourage children to apologize to their victim, redress behaviors, and begin to learn how to take responsibility for their actions.

Homework Club/academic tutoring – provides remedial sessions for children functioning below grade level. Weekly 1 hour tutoring sessions with teachers or specially trained volunteers are held in the child's home or community.

SNAP[®] Youth Leadership Club – a component offered in both the boys' and girls' programs for youth who have completed the core components of the SNAP[®] program but continue to be high risk. Staff provide group, individual, and family work to prepare at-risk youth for self-sufficiency, increase motivation for school involvement and success, improve their work-force career trajectories, and reduce their involvement with the law.

Parent problem-solving group – a 9-week group component for parents who have

completed the SNAP[®] parent group. The focus is on enhancement and refinement of family problem-solving skills in relation to ongoing issues the families are experiencing (i.e., keeping rewarding effective, difficulty with consequencing, school issues, media awareness) while providing continued support.

Long-term connections/continued care – families may continue to be involved in all components of SNAP[®] as long as there is a need and interest. In addition to previously listed components, this may also include activities such as participation as a peer or parent mentor.

Principles

Nine *principles* with specific indicators have been identified to describe the approach to service delivery and guide SNAP[®] programming. In addition, the *principles* ensure service and clinical excellence when organizations replicate SNAP[®] in their communities (sites). These are also used to assist in measuring implementation adherence, fidelity, and integrity of the SNAP[®] prevention and intervention programs (see Table 1).

SNAP[®] Theoretical Underpinnings

From the very beginning, the SNAP[®] model was built on well-known theoretical approaches showing promise in the early 1980s. These included social skills training, cognitive problem solving, self-control and anger management strategies, cognitive self-instruction, family management skills training, and parent training. As noted by Augimeri et al. (2011a), the model continued to evolve as SNAP[®] scientists and practitioners consulted with the Oregon Social Learning Center (OSLC) in Eugene, Oregon. These consultations helped to strengthen the SNAP[®] parenting component by adopting aspects of OSLC's Social Interactional Family Therapy's (now known as Social Interactional Learning) approach to working with families (Patterson et al. 2010). The SNAP[®] model programs have evolved to reflect the contributions of six core treatment theories, including Systems,

Behavioral, Attachment, Feminist and Developmental Theories (see Fig. 2). It is important to note that these theories are not viewed as stand-alone entities, but as interactive in their contributions to the foundation and ongoing development of the SNAP[®] model programs.

The SNAP[®] Logic Model (see Fig. 2) also illustrates that the primary targeted outcomes include improved overall child and family functioning with an emphasis on emotion regulation, self-control, and problem solving for both parent/caregiver and child. As Strayhorn (2002) indicates, "self-control difficulties are of central importance for many psychiatric disorders. . .[it] is also a crucial, and often missing, ingredient for success in most treatment programs" (p. 7).

Research

Research on the SNAP[®] programs has been an integral part of the model's ongoing development since its inception and continues on an ongoing basis. Scientists and clinicians work within a collaborative process (SNAP[®] *principle: scientist-practitioner*) to inform and update the theoretical approaches of the model, ongoing evaluation, and program development. Rigorous internal and external evaluations of SNAP[®] programs (e.g., process and outcome evaluation, quasi-experimental designs, random control trials, long-term follow-up – criminal record searches, cost-benefit analyses, third-party external evaluations, and neuroscience) have consistently demonstrated positive treatment effects over time: children improve significantly more than children receiving an attention-only group, delayed treatment, or an alternative treatment with notable effect sizes (moderate to large); treatment gains are maintained at 6, 12, and 18 months; parents report less stress in their interactions with their children and increased confidence in managing their children's behavior; children report improved quality of interaction with parents, less yelling, and more limit setting; children report more positive attitudes and less anxiety and demonstrate more pro-social skills with teachers, peers, and family members. Longitudinal research analysis showed that

91.8 % of the boys and 95 % of the girls had no history of criminal offences by age 14 and approximately 68 % of the children have not had a criminal record by age 19 (Pepler et al. 2010; Augimeri et al. 2007, 2011b). Brain imaging studies conducted by the Hospital for Sick Children in Toronto and the University of Toronto showed that children who responded positively to SNAP[®] treatment manifest changes in brain systems responsible for cognitive control and self-regulation, and a number of SNAP[®] families showed an ability to “repair” after engaging in a difficult parent-child interaction (see Granic et al. 2007; Lewis et al. 2008; Woltering et al. 2011).

Designations

As a result of these promising research findings, SNAP[®] has achieved the highest levels of recognition from independent reviewers who rate evidence-based programs. In 2012, the US Department of Justice, Office of Justice Programs, designated SNAP[®] as an “effective” crime prevention model (see <http://www.crimesolutions.gov/ProgramDetails.aspx?ID=231>). In 2011, Public Health Agency of Canada designated SNAP[®] as a Canadian best practice under their Preventing Violence Stream – Canadian Best Practice Portal (see <http://cbpp-pcpe.phac-aspc.gc.ca/intervention/707/view-eng.html>; <http://cbpp-pcpe.phac-aspc.gc.ca/intervention/706/view-eng.html>). In an excerpt outlining effective options for young children with conduct problems, Cipriani (2009) highlighted the SNAP[®] model as the “best example” of effective early intervention strategies and discussed the program and its successes at length. In 2008, the Canadian *National Crime Prevention Centre* designated SNAP[®] as a “Model Program” (see www.publicsafety.gc.ca/res/cp/res/2008-pcpp-eng.aspx), in 2012 it was designated as an “Effective Program” by the US Department of Justice’s OJJDP (see www.ojjdp.gov/mpg/mpgProgramDetails.aspx?ID=699) because of its robust treatment outcomes, and in 2006 it was given the highest effectiveness designation (Level 1) by the United States’ Whitehouse program *Helping America’s Youth*, now titled *FindYouthInfo.gov*; (see www.findyouthinfo.gov/ProgramDetails.aspx?pid=699).

[findyouthinfo.gov/ProgramDetails.aspx?pid=699](http://www.findyouthinfo.gov/ProgramDetails.aspx?pid=699)). At the beginning of 2000, a study group on very young offenders led by Drs. David Farrington and Rolf Loeber found the SNAP[®] program was the “most fully developed intervention to date for child delinquents” (Howell 2001, p. 312).

Current Issues/Controversies

SNAP[®] Model Replication and Implementation

For a discussion of the program’s replication standards and principles, licensing agreements, accreditation, and fidelity frameworks ensuring successful replication of the SNAP[®] model, see Augimeri et al. 2011b. In this discussion, having a team of dedicated scientists and practitioners was emphasized in order to create an effective, well-established, and recognized program, and it stressed that host organizations need commitment, support, and resources when incubating an evidence-based model program within a community setting.

Five important criteria were identified for the successful implementation and replication of SNAP[®]: (1) *Adherence* to the model is critical, (2) *restraint from making modifications* is essential, (3) *training and ongoing consultation* is mandatory, (4) *ongoing fidelity/integrity audits* are necessary to ensure the highest possible efficacy, and (5) *selecting the right staff* is paramount to program success.

SNAP[®] licensing agreements are established with the Child Development Institute. CDI’s experience in disseminating SNAP[®] is that with adequate training and support, this model can be successfully replicated and implemented with strong fidelity in a variety of settings. SNAP[®] fits in the classroom, in the clinician’s office, and at home. The program can be situated in a variety of diverse community settings and real-life community conditions. Currently, there are a number of successful SNAP[®] implementations in Canada, the United States, and Europe.

SNAP[®] Fidelity and Integrity Framework

For successful implementation and replication of evidence-based programs such as SNAP[®], training and ongoing consultation activities between the site and its replicators are considered paramount (Augimeri et al. 2011b). As the SNAP[®] model incorporates a complex therapeutic approach, it requires strict documentation of the services being delivered and records of any integrity activities conducted to ensure successful replication and outcomes. As a result, there is a need to identify all the intricate elements of the various treatment components within the SNAP[®] model (e.g., SNAP[®] core groups, individual counseling/mentoring, family counseling) in order to effectively monitor if the delivery of these key elements is done correctly and skillfully, when and where necessary. The SNAP[®] principle, *accountable service excellence*, highlights the requirement of fidelity practices that include case file audits, consultations, adherence to group manuals, and consistency of facilitation skills. It is essential that the integral pieces of SNAP[®] related to long-term positive outcomes (e.g., decreased criminal activity) are delivered with appropriate timing, skill, and adherence. Ultimately, SNAP[®] researchers and facilitators are concerned with delivering an effective program that adequately meets all objectives that were predefined with the clients.

Key Aspects to Ensure a Successful International Implementation

Even though successful SNAP[®] sites have been established worldwide, we continue to recognize that there are many obstacles to successful implementation. As noted earlier, the implementation of an evidence-based model can be challenging on its own. This is especially true when it is being adopted in another country or culture. As SNAP[®] implementations continue to reach communities worldwide, the onus is on SNAP[®] developers to explore creative methods for ensuring successful replications. There are several important factors (e.g., language, culture, travel) that may need to be considered when replicating a “foreign” intervention, even though its core strategies are proven to have universal applicability (e.g., cognitive-behavioral therapy).

Future Directions

A sixth implementation criterion that would greatly contribute to not only the successful implementation of the program, but most importantly, its sustainability, would be the *adoption of community teams for children under 12* (Augimeri et al. 2001a; Goldberg et al. 1999). *Community teams* would be comprised of representatives from child welfare, school personnel, and criminal justice systems such as the police, health, and children’s mental health. The *community team* would be responsible for setting up police-community referral protocols, managing and maintaining a centralized referral intake line, conducting comprehensive risk and needs assessments/screenings, and connecting children and families to the appropriate gender-sensitive services to at-risk children and their families. The creation of a government-supported *National Advisory Working Group for Children and Youth Involved in Offending Behaviour* would act as a knowledge-based resource center (similar to CDI’s CCCO) dedicated to knowledge transfer activities that would support the dissemination of current research and services tailored to the needs of high-risk children and their families. This national working group would act as an external, unbiased body responsible for monitoring fidelity of implementation and replication, thus ensuring integrity and accountability from those engaged in assessment and/or services.

Conclusion

After numerous decades of working with young children in conflict with the law, their families, and communities, CDI and SNAP[®] continue to support and advocate on behalf of these “forgotten children.” CDI and SNAP[®] researchers and clinicians remain committed to keeping such children out of the youth justice and adult criminal systems. These vulnerable children deserve our utmost attention and help to develop to their fullest potential.

Recommended Reading and References

- Augimeri LK, Koegl C (2012) Raising the bar: transforming knowledge to practice for children in conflict with the law. In: Loeber R, Welsh BC (eds), *The future of criminology*. Oxford University Press, New York, pp 204–210
- Augimeri LK, Koegl CJ, Goldberg K (2001a) Children under 12 years who commit offenses: Canadian legal and treatment approaches. In: Loeber R, Farrington DP (eds) *Child delinquents: development, intervention, and service needs*. Sage Publications, Inc., Thousand Oaks, pp 405–414
- Augimeri LK, Koegl CJ, Webster, Levene K (2001b) Early assessment risk list for boys: EARL-20B, version 2. Earls court Child and Family Centre, Toronto
- Augimeri LK, Farrington DP, Koegl CJ, Day DM (2007) The under 12 outreach project: effects of a community based program for children with conduct problems. *J Child Fam Stud* 16:799–807. doi:10.1007/s10826-006-9126-x, Published Online, January 10, 2007
- Augimeri LK, Enebrink P, Walsh M, Jiang D (2010) Gender-specific childhood risk assessment tools: early assessment risk lists for boys (EARL-20B) and girls (EARL-21G). In: Otto RK, Douglas KS (eds) *Handbook of violence risk assessment*. Routledge, Taylor & Francis, Oxford, pp 43–62
- Augimeri LK, Walsh MM, Liddon AD, Dassinger CR (2011a) From risk identification to risk management: a comprehensive strategy for young children engaged in antisocial behavior. In: Springer DW, Roberts A (eds) *Juvenile justice and delinquency*. Jones & Bartlett, United States, pp 117–140
- Augimeri LK, Walsh M, Slater N (2011b) Rolling out SNAP[®] an evidence-based intervention: a summary of implementation, evaluation and research. *Intern J Child Youth Fam Stud* 2.1:162–184
- Augimeri L, Walsh M, Woods S, Jiang D (2012) Risk assessment and clinical risk management for young antisocial children: the forgotten group. *Univ Psychol* 11(4):1147–1156
- Cipriani D (2009) Children's rights and the minimum age of criminal responsibility: a global perspective. Ashgate Publishing, Farnham
- Earls court Child and Family Centre (2002) SNAP[™] girls group manual: the girls club. Earls court Child and Family Centre, Toronto
- Earls court Child and Family Centre (1990b, 2001b) SNAP[™] parent group manual. Earls court Child and Family Centre, Toronto
- Earls court Child and Family Centre (1990a, 2001a) SNAP[™] children's group manual. Earls court Child and Family Centre, Toronto
- Goldberg K, Augimeri LK, Koegl CJ, Webster CD (1999) Canadian children under 12 committing offences: legal and treatment approaches. Earls court Child and Family Centre, Toronto
- Granic I, O'Hara A, Pepler D, Lewis M (2007) A dynamic system analysis of parent-child changes associated with successful "real-world" interventions for aggressive children. *J Abnorm Child Psychol* 35(5):845–857. doi:10.1007/s10802-007-9133-4, Printed On-Line
- Howell JC (2001) Juvenile justice programs and strategies. In: Loeber R, Farrington DP (eds) *Child delinquents: development, interventions, and service needs*. Sage Publication, Thousand Oaks, pp 305–322
- Howell JC (2003) Preventing & reducing juvenile delinquency: a comprehensive framework. Sage, Thousand Oaks
- Koegl CJ, Augimeri LK, Ferrante P, Walsh M, Slater N (2008) A Canadian programme for child delinquents. In: Loeber R, Slot NW, van der Laan P, Hoeve M (eds) *Tomorrow's criminals: the development of child delinquency and effective interventions*. Ashgate, Aldershot, pp 285–300
- Levene K (1998) SNAPP stop-now-and-plan-parenting: parenting children with behaviour problems. Earls court Child and Family Centre, Toronto
- Levene KS, Augimeri LK, Pepler DJ, Walsh MM, Koegl CJ, Webster CD (2001) Early assessment risk list for girls: EARL-21G, version 1, consultation edition. Child Development Institute, Toronto
- Lewis MD, Granic I, Lamm C, Zelazo PD, Stieben J, Todd RM, Moadab I, Pepler D (2008) Changes in the neural bases of emotion regulation associate with clinical improvement in children with behaviour problems. *Dev Psychopathol* 20:913–939
- Loeber R, Farrington DP (2001) *Child delinquents: development, intervention, and service needs*. Sage, Thousand Oaks
- Loeber R, Slot NW, van der Laan P, Hoeve M (2008) *Tomorrow's criminals: the development of child delinquency and effective interventions*. Ashgate Press, Burlington
- Moffitt TE, Arseneault L, Belsky D, Dickson N, Hancox RJ, Harrington H, Houts R, Poulton R, Roberts BW, Ross S, Sears MR, Thomson WM, Caspi A (2011) A gradient of childhood self-control predicts health, wealth, and public safety. *Proc Natl Acad Sci (PNAS)* 108(7):2693–2698
- Patterson GR, Forgatch MS, DeGarmo DS (2010) Cascading effects following intervention. *Dev Psychopathol* 22:949–970
- Pepler DJ, Walsh MM, Levene KS (2004) Interventions for aggressive girls: tailoring and measuring the fit. In: Moretti MM, Odgers CL, Jackson MA (eds) *Girls and aggression: contributing factors and intervening principles*. Kluwer Academic/Plenum, New York, pp 41–56, *Perspectives in Law and Psychology Series, Volume 19*
- Pepler DJ, Walsh MM, Yuile A, Levene K, Vaughan A, Webber J (2010) Bridging the gender gap: interventions with aggressive girls and their parents. *Prev Sci* 11:229–238
- Piquero AR, Jennings WG, Farrington DP (2010) Self-control interventions for children under age 10 for

- improving self-control and delinquency and problem behaviors. *Campbell Systematic Reviews* 2010 2. doi:10.4073/csr.2010.2
- Strayhorn JM (2002) Self-control: theory and research. *J Am Acad Child Adolesc Psychiatry* 41(1):7–16
- Walsh MM, Augimeri LK (2009) SNAP[®] fidelity rating index. Child Development Institute, Toronto
- Walsh M, Pepler D, Levene K (2002) A model intervention for girls with disruptive behavior problems: the earlscourt girls connection. *Can J Couns* 36:297–311
- Woltering S, Granic I, Lamm C, Lewis MD (2011) Changes in the neural correlates of emotion regulation with treatment. *Biol Psychiatry* 70(9):873–879

Strategies of Policing Terrorism

Simon Perry
Institute of Criminology, Faculty of Law, The Hebrew University, Jerusalem, Israel

Synonyms

[Counterterrorism strategies](#); [Internal security strategies](#)

Overview

In the last decade terrorism has become a prominent topic in many areas, including criminology. Since terrorism affects almost every aspect of life, it has become a central priority for the police in Western democracies. American police and even countries that had been prepared to fight terrorism prior to 9/11, such as Great Britain and Germany, began after that date to review their readiness and rethink the role of police in counterterrorism (Weisburd et al. 2009a; Bayley and Weisburd 2009; International Association of Chiefs of Police 2005; Bamford 2004). Howard (2004) argues that police departments should start thinking of themselves as proactive valuable assets in deterring, defeating, and recovering from terrorist attacks. Law enforcement, intelligence, and security agencies are expected to team up, join forces, and work

together with other organizations to uncover terrorist networks, foil terrorist attacks, respond to suspicious situations, and serve as first responders (Weisburd et al. 2009a). Yet there is a lack of evidence-based models for this new role of policing terrorism. A Campbell Collaboration (Lum et al. 2006) systematic review of strategies to combat terrorism could only identify seven studies that met minimal methodological requirements. None of these seven studies examined a police intervention. In fact, to date there are only a few descriptions of possible models for strategic and tactical activities of policing terror. Therefore, little is known about what the best antiterrorism strategies and tactics are. Furthermore little is known about how models can be systematically measured and assessed for their effectiveness (Weisburd et al. 2009a).

This entry will attempt to summarize recent developments in the field of counterterrorism. It will introduce and portray principal strategies, tactics, and practices that are presented in the literature and/or that have been widely adopted by practitioners in policing terrorism.

Why Police Bear Primary Responsibility for the Terrorism Threat in Democratic Countries?

Before discussing what these counterterrorism strategies, tactics, and practices are, four principal themes that naturally surface will be discussed: (a) the characteristics of the terrorism threat phenomenon, (b) the complexity of developing and evaluating a counterterrorism model, (c) the tension between preserving democratic principles and counterterrorism, and (d) why the police (according to Bayley and Weisburd 2009) bear primary responsibility for preserving public security in most countries.

The Terrorism Threat

In order to address the threat and develop an effective response to terrorism, one should first define this criminological phenomenon and its goals (Weisburd et al. 2009b). According to

Ganor (2009), even after the world recognized the scale of the terrorist threat after 9/11, there is still no agreement on a definition of terrorism. After studying various definitions of terrorism (Hoffman 2006; Ganor 2009; Martin 2009; Hasisi et al. 2009) in the context of seeking to develop an effective response to terrorism, one can conclude that the phenomenon of terrorism includes the following basic elements: (a) the creation of *fear and anxiety*; (b) doing so through the use, or threat of use, of *violence*; (c) non-state actors; (d) the goal of weakening the fabric and the *resilience of society*; (e) inflicting *random casualties* among the *general population*; (f) increasing the *frequency of attacks*; (g) targeting the *weakest and most vulnerable* members of society; (h) taking advantage of the *media*; and (i) all in order to pressure decision makers to *accept terrorist demands*.

The Complexity of Developing Strategic Policing-Terrorism Models

The lack of evidence-based models in the policing-terrorism field, as well as a lack of systematically evaluated strategic and tactical activities, is a result of two main reasons. First and foremost, law enforcement intelligence and security agencies are very reluctant to cooperate with such research, as they fear the possible ramifications including the exposure and compromise of counterterrorism methods, tools, sources, and tactics. Secondly, this type of research encounters difficulty in the measurement of success and the determination of cause and effect. These difficulties stem from the presence of other variables in the tested model, the effect of which is to make creating a control situation extremely difficult.

Preserving Democratic Principles and Human Rights in Counterterrorism

In counterterrorism the democratic liberal values of the society limit the capability of the state to make use of the full potential means and technological advantages that the state has. In its struggle against terrorism, the democratic state is obliged to select measures and utilize

capabilities that will cause minimum damage to human rights. Innes (2006) argues that collecting and using intelligence is part of the “dirty work” of a democracy. Usage of superior capabilities harming those who have no connection to terrorism, as well as harming fundamental moral principles that are essential for the democratic state, would constitute a victory for the terrorist organization. And such misuse could alienate parts of society, playing into the terrorists’ hands (Ganor 2009).

Bayley and Weisburd (2009) argue that legitimacy is the foundation of successful policing, whether related to terrorism or to crime in general. Even though protection from terrorism is a moral cause, it can lead to the violation of human rights, causing officers to approach individuals as suspects rather than as individuals who deserve service. Losing police legitimacy endangers the advantages of public cooperation that are greatly needed in the war on terror.

This is especially significant in minority communities that are linked ethnically or nationally to terrorist groups. For example, in the Israeli context, this applies to familial and national ties of Israeli Arabs to Palestinians in surrounding countries and in the Palestinian-controlled territories (Hasisi et al. 2009).

Why Should the Police Lead the Response to Terrorism?

There are five main reasons why the police force has been chosen to lead the response to terrorism in most countries:

1. Since *terrorists are criminals, and terrorism is a type of crime* in all essential characteristics (Clarke and Newman 2006), it is appropriate that, in democracies, the police force, which is best qualified to deal with criminals and crime, should lead the response to terrorism. **The criminal-justice model** for counterterrorism features terrorism as a crime and portrays the terrorist as violent criminal who should be arrested and punished according to the rule of law by the police and the criminal-justice system (Greene and Herzog 2009). According to Perliger, Hasisi, and Pedahzur (2009), most of the literature discusses on

democracy and counterterrorism believe that **the criminal-justice model** permits dealing with terrorism without seriously undermining the legal and moral foundations of the democratic system. Therefore, it is seen to be better for democratic countries to leave counterterrorism in the hands of the police which operate in the civilian arena. By contrast, **the war model** (Greene and Herzog 2009) features terrorism as an act of war that disputes and threatens the well-being of the state and the political system. Therefore, in this model, it is believed that the terrorist and the terrorist organization should be eliminated by the use of intelligence and military forces. However, employment of **the war model** paradigm would lead to situations in which military forces conduct combat warfare within their own territory, constituting a severe undermining of human rights and morality of the democratic state and its legal system.

2. *Many of the methods and means necessary to fight terrorism are used routinely by the police in their daily operations.* These include investigation and information and evidence collection; forensics (identification of weapons, explosives, victims, etc.); police-operated call centers which are first responders; and police liaison with the private sector, including the issuance of licensing to businesses. Moreover, many activities associated with the handling of terror attacks, whether before, during, or after such an attack, are an integral part of routine police activity (traffic control, managing crime scenes, and maintaining or restoring public order).
3. *The connections between terrorists and other criminals give the police an advantage.* Criminals facilitate terrorism with many necessary means such as weapons and explosives; documentation; vehicles; collecting, transferring, and laundering money; information; communications; technology; and even operational subcontracts for specific missions. Terrorist organizations, in order to finance their activities, have used classic organized-crime illegal activities such as money counterfeiting and the smuggling of drugs,

counterfeit goods, and taxable merchandise like cigarettes. This connection gives the police, who have a familiarity with the criminal world, an edge in counterterrorism.

4. *Police are responsive to irregularities in the environments in which they operate routinely.* During their everyday policing of crime and disorder, the police can be on the lookout for suspicious indicators connected to terrorism activities in their communities (Innes 2006). This puts the police in an exceptional position to collect information (that remote intelligence agents generally cannot obtain, since they are not connected to the community) and develop critical counterterrorism intelligence that can be disseminated in real time to other law enforcement agencies (Connors and Pellegrini 2005). Perliger, Hasisi, and Pedahzur (2009) argues that it is better to leave the struggle against terrorism in the hands of the police as part of a more comprehensive model – the criminal-justice model – which treats terrorism as violent criminal action in the civilian arena.
5. *Legal authority to perform policing procedures.* In many democracies only the police have legal authority to perform policing procedures that are required in counterterrorism. The military, for example, in the USA, Canada, UK, and Israel, is not legally empowered to engage in ordinary police functions within the borders of the country. This situation is unlike many European countries (e.g., France, Spain, and Italy) where the military play an active policing role (Weisburd et al. 2009a).

Defusing Terrorist Motivations Versus Preventing Terrorist Opportunities

Based on the “rational choice” theory (Clarke and Cornish 2001), crime is mainly a result of a *process* in which the individual evaluates *opportunities* weighing the anticipated benefits against the expected costs of the behavior, in the context of a certain time and place. Terrorism, like any other criminal behavior, is

an outcome of two necessary conditions: *motivation* to commit the terrorist behavior and *opportunity* to carry out a terror attack at a particular occasion (Clarke and Newman 2006). Therefore, any counterterrorism strategy should defuse *terrorist motivations* or prevent *terrorist opportunities* or do both.

The strategy that tries to prevent the terrorist from attacking by defusing the *terrorist motivation* is known as the **conciliatory model**. According to this model, terrorism is a political problem, and it should be given a political solution that will address the root cause of terrorism by the action of policymakers, brokers, and diplomats (Greene and Herzog 2009). However, Clarke and Newman (2006) argue that counterterrorism “must not rely on changing the heart and minds of terrorists. The motivation for terrorism results from long-term social, cultural and psychological pressures, which are difficult to alter” (Clarke and Newman 2006:11).

According to Weisburd and Waring (2001), the efficient approach for prevention of crime involves decreasing the *opportunities* (the core of the situational prevention approach presented by Felson and Clarke 1998) present in certain situations which encourage commission of a crime. Clarke and Newman (2006) argue also that in counterterrorism it is easier to reduce *opportunity* than to reduce terrorist *motivation*. This is especially true in the case of suicide bombers that are very hard to deter. Therefore, the efficient approach to prevent terrorism is to adopt strategies that attempt to reduce the opportunities to commit terror attacks. In any case, they claim, easy opportunities encourage terrorists to attack.

In short, defusing terrorist motivations is mainly not a law enforcement intelligence operational task and seems to be less efficient in preventing terrorism than the approach of reducing opportunities. **Therefore, most counterterrorism strategies concentrate on attempting to reduce the opportunities rather than reducing motivation.**

Reducing opportunities could be achieved by the *defensive model* that aims to protect

potential targets and victims from attacks, through “target hardening” (Clarke and Newman 2006), or by *proactively preventing operational capabilities* of terrorist organizations and individuals that pose the physical and psychological threat. According to Clarke and Newman (2006 p. 9), there are four “pillars of terrorism opportunity”: targets, weapons, tools, and facilitating conditions.

The Responsive Approach and the Proactive Approach

Law enforcement practices have two main models of operation: the responsive approach and the proactive approach. In the responsive approach, after a crime or specifically a terrorist act has occurred, law enforcement responds by gathering intelligence, identifying perpetrators, collecting evidence, and arresting suspects. This responsive approach calls for defensive methods that could perhaps foil an attack that has been launched, or at least minimize the damage. Therefore, in counterterrorism we also refer to it as a responsive defensive approach. This approach better protects vulnerable places and effectively restores order in the event of a terrorist attack (Weisburd et al. 2009b).

The proactive model attempts to foil the crime **before** it is launched. This approach has been used for decades by law enforcement against organized and other forms of crime, especially drug trafficking. In this approach law enforcement intelligence identifies potential criminals/terrorists, gathers intelligence and evidence, makes the arrest, and foils the crimes/attacks before they can occur. The main goal of this approach, which can be classified as “high policing,” utilizes covert intelligence gathering, surveillance, and foiling operational tools (Bayley and Weisburd 2009; Brodeur and Dupeyron 1993).

Terrorism seeks to create fear and anxiety, weaken the fabric and the resilience of society, and destabilize a social order, all by increasing the frequency of attacks and inflicting mass casualties. The objective of law

enforcement is therefore to reduce the operational capabilities of terrorist organizations through counterterrorism measures which can be characterized as proactive, preventive, and offensive in nature. This goal is a crucial part of every democratic country's duty to protect its citizens (Ganor 2009).

The proactive method is intended to deter, disrupt, and prevent terrorist activities (Weisburd et al. 2009b; Hasisi et al. 2009). This counterterrorism strategy damages terrorist organizations physically, psychologically, and economically and hurts their operational infrastructures. It pushes them into a defensive mode, where they spend a great deal of time and resources concealing their activity, which thus limits their effectiveness and reduces their ability to carry out attacks. According to Hasisi et al. (2009), this proactive strategy is becoming more complex as a result of the ever-changing nature of terrorism. In the contemporary age of globalization, terrorist groups tend to be stateless, making them a more elusive target that readily exploits open societies and advanced communication technologies, while evading surveillance by traditional security agencies.

The *responsive approach*, which is defensive in nature, and the *proactive approach*, which is preventive and offensive in nature, complement rather than contradict one other.

Since, unfortunately, it is unrealistic to expect to entirely prevent all terrorist attacks, a good responsive approach is important to have in place alongside an efficient proactive approach. This combined approach both apprehends terrorists and damages terrorist organizations. By combining both approaches, counterterrorism forces can prevent or at least significantly minimize terrorist *opportunities* to attack before the attacks have been launched, *responsively* foil the attacks once they have been launched, and effectively minimize the extent of the casualties and damage, as well as restore order and evacuate the scene during and after the attack. Such a comprehensive method of counterterrorism makes it possible for the general public to maintain their everyday routines and preserves their morale and sense of personal security.

In the United States the view of local law enforcement officers as primarily "first responders" is slowly changing, and they are becoming viewed also as "first preventers" of terrorism. In New York City, the NYPD adopted this concept of "prevention" that has been used so successfully against crime, and adapted it for the war on terror (Bratton and Kelling 2006). Connors and Pellegrini (2005) claim that local officials and police in the United States must be prepared to take the lead in the war on terror and not to wait for direction from federal agencies hundreds of miles away, if they want to prevent or recover from any future terrorist attacks.

According to Innes (2006), in the United Kingdom as early as the 1990s, a new intelligence-led policing improved police agencies' effectiveness by proactively identifying problems and then targeting individuals who were most likely to cause harm to the general public. This change was aimed at pursuing a more proactive mode of working based on the principles of risk management. Given that the police have limited resources and developing effective live informants ("HUMINT") is difficult, a network of community intelligence contacts was established to provide fairly effective surveillance over potentially dangerous groups and individuals. Expert community engagement units were set up to develop and maintain this "community intelligence feed" that would serve both anti-criminal and anti-terrorist purposes.

The Israel National Police (INP) has a long history of experience struggling with terrorism and is considered highly efficient and professional in its approach to counterterrorism (Weisburd et al. 2009b). The Israeli model for policing terrorism applies this principal strategy of combining the *proactive offensive* and *responsive defensive* approaches. It is implemented in three circles of activity: (a) *sources of terrorism circle* – proactive early prevention, interdiction, and treatment of the sources of terrorism in order to foil terrorist attacks before they are launched; (b) *the attack route* – response activities once the attack has

been launched, to foil terrorist attacks before they reach the target; and (c) *the terrorist targets* – defending and “hardening” potential targets before any attack and response activities at the scene during and after an attack.

In the Israeli antiterrorism model, there is a great deal of emphasis on the proactive model approach, since it enables capturing terrorists, decreases the frequency and severity of attacks, and results in fewer casualties. A proactive model is preferable since it has a greater counterterrorism effect, preventing terrorists from achieving their goals and making it possible for the general public to maintain their everyday routines.

Proactive Offensive Early Interception Counterterrorism

As discussed above, the offensive-proactive approach is the most effective counterterrorism strategy and is the key for preventing attacks before they are launched. This strategy identifies and responds to terrorist threats before they actualize, by developing intelligence, establishing operational capabilities, and uprooting terrorists and their infrastructure (Weisburd et al. 2009b). The proactive approach is based on three crucial competencies: producing *quality intelligence*, building *operational capability*, and creating a *hostile environment for terrorists’ operations*.

Effective counterterror thwarting operations rely on the ability to gather reliable information about the terrorists’ capabilities, intentions, and specific plans. Quality intelligence takes away from the terrorist the element of surprise, which is a critical advantage for a terrorist attack. Intelligence allows security forces to efficiently protect in advance targets that may be selected by terrorists. Furthermore, intelligence puts the element of surprise in the hands of law enforcement.

Once a terrorist attack has already been set in motion, security forces utilize intelligence both for proactive offensive thwarting operations and for responsive defense measures (Perlinger et al. 2009).

The Intelligence Process

Producing intelligence is a circular flow process which constantly aims to portray an accurate picture about the threats. The intelligence analyst defines (based on the initial information) the *information gaps*, which is the information he needs but does not have in order to portray a reliable *intelligence picture*. The *intelligence picture* is composed of all the information about subjects that are connected significantly to the different threats, which are called *topics of interest*, and all the information about individuals who are connected significantly and compose the threat; they are called *targets*.

In order to fill in the *intelligence gaps*, an *intelligence collection and coverage plan* is prepared. The plan includes three levels of coverage: the abovementioned *topic of interest* and *targets*, as well as the *territorial coverage*, which is the entire information about a specific area (neighborhood or town) including who is doing what, when, and where, and who knows about it. At the first stage, potential sources of information are identified in order to cover all three levels. The next stage is the preparation of the *recruitment plan*, which is the methods of recruiting and running live and technical sources. The various sources utilized for coverage are as follows: live informants (“HUMINT”) and technical sources (called “SIGINT”) such as wiretaps, surveillance activities, investigations and debriefings, archives, and databases; public open information such as the news media and the Internet; and fellow agencies from the same country or international. The role of the analysts is to prepare and present an *integrative intelligence picture*. Their job is an unending process, since they are the “nerve center” of information: they direct the collection and thwarting of intelligence according to the requirements of the command guidelines, they assist in the preparation of an operational perception, they support the development of “HUMINT” sources, they acquire and deploy intelligence capabilities and tools, and they administer the Intelligence Database.

Unlike military intelligence, which tends to rely on advanced technological monitoring and surveillance capabilities, HUMINT sources are considered more effective for penetrating terrorist groups. Police intelligence generally deals with criminal organizations which are small active groups from within the civilian setting and whose structure and character resemble terrorist groups (Perliger et al. 2009).

Intelligence Sharing and International Cooperation

In many cases the terrorism threat is both national and international. Therefore, cooperation and quality intelligence sharing are essential for law enforcement and intelligence organizations to prevent attacks. Kelling and Bratton (2006) emphasize the importance of collecting and sharing intelligence in the fight against terrorism.

The mutual interrelation between terrorist organizations and the classic hardcore criminal organizations enables the police to extend the well-established national and international cooperation on fighting organized crime to the foiling of terrorism.

Special Operations Capabilities

The mere collection of counterterrorism intelligence is not sufficient, since intelligence without the ability to reach a target or prevent an attack is of little value. These tasks are performed by special operations units which are trained to enter locations where terrorist actions are being planned and prepared. The goal is to foil the attacks before they are launched and to threaten the terrorists' own sense of security, keeping them busy and on the run. Such units should be able to conduct undercover operations in which they reach their target and arrange an arrest without being detected. The police should have specialized elite counterterrorism units, trained to handle very specific terrorist situations, such as releasing hostages and carrying out special operations using small disciplined teams highly trained in commando style military operations.

Bratton and Kelling (2006) argue that the police need special training and a new

mindset so they can proactively assault terrorism. Police departments from all over the world are already exchanging information with other countries as the war against terror is gradually becoming a global threat. The government of Israel, for example, has welcomed police forces from all over the United States for training and exchange visits.

Proactively Changing Terrorists' Operating Environment

Bratton and Kelling (2006) explain that the theory of "broken windows" is used to constantly create a hostile environment to potential criminals, building up the uncomfortable sentiment that they are the ones who are in danger. In the same way police officers should create a hostile environment, mainly through the potential terrorist support structures, in which terrorists will feel uncomfortable. Constantly changing the terrorists' operating environment by establishing "bottleneck passage" creates a hostile environment while taking away the element of surprise from the terrorist. Bottleneck passage, such as obstacles and barriers, compels the terrorist to take counteractions, involving other coconspirators by using communication channels. These communication channels leave "intelligence footprints," increasing the opportunities for intelligence collection, by both human and signal technical means.

Defensive Response

The *defensive response* to attacks (that were not foiled by the offensive-proactive activity) should be dealt with in two circles of activities: (a) response activities once the attack has been launched and it is on its way to the target and (b) response activities before, during, and after an attack has occurred on a terrorist targeted site.

Response Activities Foiling the Attack on Its Way to the Target

Once an attack is ongoing and the terrorist is on the route to the target, there are a series of possible tactics to delay the terrorist's

movement, including setting up road blocks, creating traffic jams, and closing specific public facilities or streets. There are two main goals for creating such obstacles. The first objective is to slow down the terrorist as much as possible, thus delaying the attack. This gives the police more time to bring special operations units to engage with the terrorist on route and at the same time to organize a better defense at potential terrorist targets. The second objective of the obstacles is (as mentioned above) to compel the terrorist to establish communication channels that increase the opportunities for “HUMINT” and “SIGINT” intelligence collection. The enhanced intelligence and better police deployment increases the possibility for interdiction before the attack takes place.

Once there is a specific threat that terrorists have got through the security and are out of control, the police should consider making an announcement through the media, informing the public to stay away from crowded places.

Defensive Activities at a Terrorist’s Potential Targeted Site Before an Attack

The number of potential target is endless, and police resources are limited. Therefore, the police need to conduct an efficient defensive effort, using risk assessment in order to build an effective protection plan. Vulnerability and risk analysis must then drive operational responses in order to create a truly effective policing apparatus against terrorism (Connors and Pellegrini 2005). The plan needs to prioritize the allocation of defensive tools in order to harden vulnerable potential targets. According to Davis et al. (2004) after 9/11 three-fourths of the police departments in the United States conducted risk assessment compared to only one-fourth that performed such analysis before 9/11. Private security firms are the police’s main partner in “target hardening”. In this partnership the police should train, provide necessary information, and supervise so that the private security firms can better protect their clients’ facilities. In Israel, all malls, shopping areas, restaurants, hospitals, office buildings, or any

other public facilities have private security guards who check customers entering the facility and conduct other security-related activities. The police, prior to approving the business licenses for a facility’s operation, review the business facilities and their security procedures. Additionally the police perform periodic security checks, in which facilities that do not hold to the security standards set by the police may be forced to shut down through issuance of a court order (Weisburd et al. 2009a).

Realistically, not all targets can be rendered attack proof; thus, the protection plan should aim to minimize the damage in case terrorists are able to attack. To reduce the number of victims and amount of damage in terrorist attacks that were not foiled, the police work with the private security agencies to prevent terrorists from accessing public facilities. The underlying principle here is that an explosion that occurs within a closed environment (especially if crowded) will result in a much more destructive outcome than if the same explosion took place outdoors, hopefully away from the large crowd.

An important element of target hardening is educating the public and improving routine security preparedness. The police should play a central role in educating the public from an early age to be aware of indications of possible terrorism events and to provide information by reporting their suspicions to the police. In Israel, for example, police visit elementary schools, teaching the children to be attentive towards suspicious people and objects and to report to grown-ups or, if possible, to a policeman. Despite the fact that most security calls to the police are false alarms, the police treat each and every call as if it were an actual explosive device or other security threat. It is important to do so because of the potential damage from every terrorist attack and to demonstrate that the police are responsive, so that the public will continue calling (Weisburd et al. 2009b). Such responsiveness increases public confidence in the police, when in the field of counterterrorism, they are viewed by the public as responding efficiently.

These approaches are also deployed in the United Kingdom, where the police invest a great deal of attention in developing public awareness of suspicious behavior (suspicious short-term tenants, suspicious people who have bought or rented a car, etc.). The police in the United Kingdom develop suitable reporting mechanisms, alongside creating working relationships with the business community to protect businesses from potential threats and to provide guidance on appropriate security measures (Howard 2004).

Defensive Activities on a Terrorist Targeted Site During and After an Attack

The police, being the first responders, are naturally expected to act in response to a terrorist attack in such a way as to effectively manage the crisis. The Israel National Police (INP), which has tremendous experience managing many scenes of terror attacks over the years in Israel, has developed effective practices and protocols. Through the whole process, the highest territorial police commander who is on the scene is responsible for all activities that take place throughout the procedure, until he has been released by his commander or when the scene is cleared. The overall responsibility is never divided and it is always clear. During the entire process, all organizations that are involved are subordinated to the police commander, including the medics, the firefighters, and even the employees of the local city council who later clean up the area. The first mission is to secure the scene from secondary explosive devices or additional terrorists, to prevent a tactic that has been used by terrorists to hit the first responders – the policemen and the medics. Therefore, the first on the scene are the bomb-squad technicians, who close down and search the inner scene. At this stage, along with the bomb-squad technicians, only the medics who treat and evacuate the most critically injured are allowed on the scene. Only after the inner scene has been secured are the rest of injured treated and evacuated. Forensic units follow to identify the dead and – with a special unit – evacuate the bodies; collect intelligence and try to identify the type

of explosive device or weapons used, for the purpose of linking it to a specific terrorist organization or laboratory; and collect evidence with the criminal investigators. At the same time traffic and patrol officers set up road blocks, closing down the outer ring and directing traffic, clearing the way for the ambulances, searching for accomplices who might have assisted the attacker and are trying to get away, and controlling crowds. Intelligence units collect information that may assist in identifying the source of the explosives and the people responsible

As mentioned, in order to defeat the goals of terrorism, the main object of the police in counterterrorism is to preserve and strengthen the resilience of the population, enabling the public to continue with their daily life routine. It is expected that a rapid clean-up of the terrorist scene will minimize the psychological effect of the attack. Therefore, timing is critical, both for psychological reasons and also for forensic reasons – to assure the collection of evidence before the scene is contaminated. In the Israeli model, all of these activities of clearing terrorist scenes are expected to be accomplished in no more than 4 h. Another important element in preserving public resilience and reducing distress and fear is keeping the important communication channels open to the public thru the media, in which the territorial commander (or deputy or spokesperson) reports calmly and informatively during and after the attack event to news agencies. The ongoing briefing prevents damaging rumors, giving the public the safe feeling that things are under control by providing such information as the description of the event, the areas or roads that have been shut down, and alternative routes (Weisburd et al. 2009a).

Organizational Elements in Policing Terrorism

A Clear Division of Authority and Responsibility Alongside Vital Cooperation and Partnership

Some countries have a highly centralized police organization at the national level, with a clearly

distinguished purpose and responsibility related to crime, terrorism, and public order. However, even in such a setting, the police force has partners in the counterterrorism effort. Furthermore, any police force is composed of different units, so that harmonized cooperation doesn't always come naturally. In the chaotic reality of a terrorist attack, it is fundamental to have an unambiguous division of authority and responsibility that clearly directs who is the one person in charge at a certain time and place and who bears responsibility and accountability. Vagueness about who is running the show leads to confusion, lack of necessary decision making, and failure that will end with unnecessary victims and destruction.

For example, Israel has only one national police force (Israel National Police, INP), and it is highly hierarchical and centralized, with unique centralized operational combat special units. Since intelligence gathering is also centralized, it makes the processes of detecting, identifying, deterring, and thwarting potential terrorist attacks more efficient (Greene and Herzog 2009). There are close and ongoing formal and informal contacts between the police and the Internal Security Agency (ISA) (Weisburd et al. 2009a). This intimate cooperation between the INP, which has overall responsibility for internal security, and the ISA, which is the main source of counterterrorism intelligence, was not a trivial matter to achieve. It required tremendous effort to build trusting relationships at all operational and command levels. Such a relationship makes it possible in a matter of minutes to turn critical ISA information into an INP foiling operation. More broadly, the close connection between the INP and the ISA facilitates the exchange of intelligence while enabling consistency in both offensive and defensive counterterrorist operations (Hasisi et al. 2009).

Countries that do not have a centralized police system have to synchronize their counterterrorism activity on the national level as well as the local level. For example, in the United States before 9/11, there was a lack of intelligence coordination that has been strongly

criticized (Weisburd et al. 2009b). Currently local, state, and federal law enforcement agencies exchange information on initiatives, as well as creating centralized special officers and units, including special response teams.

According to Bratton and Kelling (2006), intelligence-led policing is making its mark in the USA, and major efforts are being made to restructure police capabilities for an increasingly proactive intelligence gathering and analysis apparatus. Instead of relying on the federal government for intelligence, many state and local departments are now creating their own systems. Among other actions, they are assembling databases and sharing information. This very important development requires highly sophisticated coordination, especially in such a large country as the USA that has around 17,000 police organizations on the federal, state, and local levels. These organizations are adopting a set of national strategic guidelines that clearly define the division of authority and responsibility, setting vital cooperation and partnership procedures.

Recruitment, Training, Weapons, Equipment, and Relevant Exercises

Since counterterrorism is a major responsibility of the police, police officers need to be trained and equipped to confront the terrorism threat. Kelling and Bratton (2006) argue that counterterrorism has to be woven into the working protocols of every police department, so that it becomes part of the everyday thinking of the officers who are on the street.

All officers, including those whose main role within the police is *not* counterterrorism, should undergo basic counterterrorism training. The training should prepare officers for an unexpected encounter with a terrorist incident. It should focus on acquiring first response skills such as isolating the site of a terrorist attack and effectively handling the personal weapon. Fast response teams that would defuse terrorist situations and bring them to an end as soon as possible should be put together and trained. The goal of such teams is to be prepared for fast intervention to minimize the damage, and to

control and contain the scene until the attack is defused or the special counterterrorism unit takes over. Response time is crucial in containing a terrorist attack; therefore, the fast response teams should have sufficient training (such as urban warfare), relevant exercises, equipment, and suitable transportation, such as motorcycles (Weisburd et al. 2009a).

Shortcomings of Policing Terrorism

The leading role of police in counterterrorism elevates new problems and dilemmas for the police in democratic countries. On the one hand, counterterrorism is a natural extension of the classic police duties, as previously mentioned. On the other hand, it is clear that terrorism may call for new responsibilities that emphasize “high policing,” which is characterized by its focus on strategic issues at a macro level, rather than local crime and disorder problems (Bayley and Weisburd 2009; Weisburd et al. 2009b). High policing emphasizes controlling rather than servicing the public, a standpoint very different from the community policing ideas that have reinforced community-police relationships. Counterterrorism functions of the police are likely to clash with the goals of creating closer police-community relations, especially with minorities, since it is difficult to be “officer friendly” and at the same time collect intelligence on suspects that are part of the community or related to it (Weisburd et al. 2009b).

Summary

Counterterrorism is geared to making it possible for the general public to maintain their everyday routines, thus preserving their morale, feelings of personal security, and resilience. This entry has introduced, described, and explained the following *principal strategies, tactics, and practices* that are presented in the literature and that are adopted by practitioners in policing terrorism:

1. Most of the counterterrorism strategies concentrate on attempting to reduce the *opportunities* for terrorist attack, since it is easier than to reduce terrorist *motivation*.

2. Efficient counterterrorism combines *proactive offensive* and *responsive defensive* approaches in order to reduce terrorist *opportunities*.
3. *Proactive offensive* counterterrorism identifies and responds to terrorist threats before the terrorists have an *opportunity* to attack.
4. The first crucial component of the *proactive offensive* approach is *quality intelligence* that is an outcome of a circular flow process. This process aims to constantly portray an accurate picture about the threats, turning information into timely action-operational intelligence.
5. The second crucial component of the *proactive offense* is establishing *operational capabilities* that have accessibility to uproot terrorists and their infrastructure, pushing them into a defensive and ineffective mode by keeping them busy and on the run, thus foiling attacks before they are launched.
6. The third crucial component of the *proactive offense* is constantly changing terrorists’ operating environment by establishing bottleneck passage such as obstacles and barriers, causing an intelligence footprint. Bottleneck passage compels the terrorist to take counteractions, involving other coconspirators and using communication channels. These communication channels boost the opportunities for intelligence collection, both human and signal technical.
7. Since in many cases the terrorism threat is both national and international, cooperation and quality intelligence sharing are essential for law enforcement and intelligence organizations to prevent attacks.
8. The *responsive defensive* model enables the general public to maintain their everyday routines, strengthening their resilience and preserving their morale and feelings of personal security. The model includes “target hardening,” foiling launched attacks, and effectively restoring order during, and efficiently evacuating the scene rapidly after, the attack.

9. In order to foil launched attacks, there are a series of possible tactics to delay the terrorist's movement, compelling the terrorist to establish communication channels. This allows enhanced intelligence gathering, potentially allowing special operations units to engage with the terrorist on route, as well as better police defensive deployment at potential targets, increasing the possibility for interdiction before the attack takes place.
 10. In order to build an effective protection and "target hardening" plan, a vulnerability and risk analysis must be part of the police apparatus. Private security firms are the police force's main partner, to whom the police should provide necessary information and supervision so that the private security firms can better protect their clients' facilities and minimize the damage in case a terrorist was able to attack.
 11. The police should educate the public from an early age to be aware and to report any possible threat, which should always be treated by the police as if it was a genuine attack.
 12. It is fundamental to have clear division of authority and responsibility that clearly directs who is the one person in charge at a certain time and place in the event of an attack and who bears responsibility and accountability. Vagueness leads to confusion, lack of necessary decision making, and failure that will end with unnecessary victims and destruction.
 13. At an attack scene the first mission is to secure the scene from secondary explosive devices and to treat and evacuate only the critically injured. After the scene has been secured, the rest of the injured are treated and evacuated. To minimize the psychological effect of the attack, the forensic units follow, to effectively and rapidly identify and evacuate the dead, and collect evidence and intelligence, allowing a quick clean-up of the scene.
 14. The police commander needs to keep the important communication channels open to the public thru the media, reporting calmly and informatively during and after the attack in order to reduce distress and fear.
 15. Countries that do not have a centralized police system need to synchronize their counterterrorism activity on the national and local levels, by adopting a set of national strategic guidelines that clearly defines the division of authority and responsibility, setting vital cooperation and partnership procedures.
 16. All police officers should undergo basic counterterrorism training that should prepare officers for an unexpected encounter with a terrorist incident. Response time is crucial in containing a terrorist attack; therefore, the fast response teams should have sufficient training, relevant exercises, equipment, and suitable transportation.
- Now that the policing-terrorism models have been described, the key question that remains is "How effective are these *principal strategies, tactics, and best practices* of 'policing terror'?" As Weisburd, Jonathan, and Perry (2009a) argue, it is clearly time for scholars to begin to evaluate the effectiveness of police responses to terrorism.

Related Entries

- ▶ [Biometrics and Border Control Policing](#)
- ▶ [Democratic Policing](#)
- ▶ [Ecoterrorism](#)
- ▶ [Effectiveness of Situational Crime Prevention](#)
- ▶ [History of Criminal Investigation](#)
- ▶ [Homegrown Terrorism in the United States](#)
- ▶ [Incarcerating Terrorists](#)
- ▶ [Intelligence-Led Policing](#)
- ▶ [Methodological Issues in Evaluating Police Performance](#)
- ▶ [Policing Terrorism and Legitimacy](#)
- ▶ [Preventing of Terror at Shopping Malls](#)
- ▶ [Rational Choice Theory](#)
- ▶ [Risk Assessment, Classification, and Prediction](#)
- ▶ [Routine Activities Approach](#)
- ▶ [Situational Approaches to Terrorism](#)

- ▶ [Social Disorganization and Terrorism](#)
- ▶ [Terrorist Organizations](#)
- ▶ [Theories for Situational and Environmental Crime Prevention](#)
- ▶ [Use of Social Media in Policing](#)

Recommended Reading and References

- Bamford BWC (2004) The United Kingdom's war against terrorism. *Terrorism Polit Violence* 16:737–756
- Bayley D (2006) *Changing the guard: developing democratic police abroad*. Oxford University Press, New York
- Bayley D, Weisburd D (2009) Cops and spooks: the role of police in counterterrorism. In: Weisburd D, Feucht TE, Hakimi I, Mock LF, Perry S (eds) *To protect and to serve: policing in an age of terrorism*. Springer, New York, pp 81–99
- Bratton W, Kelling G (2006) Policing terrorism. *Civic Bull* 43:1–10
- Brodeur J, Dupeyron N (1993) Democracy and secrecy: the French intelligence community. In: Brodeur J, Gill P, Tollborg D (eds) *Democracy, law and society*. Ashgate, Aldershot, pp 19–23
- Clarke R, Cornish B (2001) Rational choice. In: Paternoster R, Bachman R (eds) *Explaining criminals and crime: essays in contemporary criminological theory*. Roxbury, LA, pp 23–42
- Clarke R, Newman G (2006) *Outsmarting the terrorists*. Praeger Security International, Westport, CT
- Connors TP, Pellegrini G (2005) *Hard won lessons: policing terrorism in the United States*. The Manhattan Institute, New York, Safe Cities Project
- Davis LM, Riley JK, Ridgeway G, Pace J, Cotton SK, Steinberg PS, Damphousse K, Smith BL (2004) *When terrorism hits home: how prepared are state and local law enforcement?* Rand Corporation, Santa Monica
- Felson M, Clarke R (1998) Opportunity makes the thief: practical theory for crime prevention. *Police Research Series*. Paper 98
- Freilich J, Newman G (2009) Introduction. In: Freilich J, Newman G (eds) *Reducing terrorism through situational crime prevention*. U.S. Criminal Justice Press, Monsey, pp 1–7
- Ganor B (2009) Trends in modern international terrorism. In: Weisburd D, Feucht TE, Hakimi I, Mock LF, Perry S (eds) *To protect and to serve: policing in the years of terrorism, and beyond*. Springer, New York, pp 11–42
- Greene JR, Herzog S (2009) The implications of terrorism on the formal and social organization of policing: some concerns and opportunities. In: Weisburd D, Feucht TE, Hakimi I, Mock LF, Perry S (eds) *To protect and to serve: policing in the years of terrorism, and beyond*. Springer, New York, pp 81–99
- Hasisi B (2006) Protest policing in deeply divided society: Israeli Arab protest under fire. Paper presented at the annual meeting of the American Society of Criminology, Toronto, Oct 2006
- Hasisi B, Alpert GP, Flynn D (2009) The impacts of policing terrorism on society: lessons from Israel and the U.S. In: Weisburd D, Feucht TE, Hakimi I, Felson Mock L, Perry S (eds) *To protect and to serve: policing in the years of terrorism, and beyond*. Springer, New York, pp 177–202
- Hoffman B (2006) *Inside terrorism*. Columbia University Press, New York
- Howard P (ed) (2004) *Hard won lessons: how police fight terrorism in the United Kingdom*. Manhattan Institute, New York
- Innes M (2006) Policing uncertainty: countering terror through community intelligence and democratic policing. *Ann Am Acad Polit Soc Sci* 605:222–241
- International Association of Chiefs of Police (2005) *Post 9-11 policing: the crime-control-homeland security paradigm—taking command of new realities*. International Association of Chiefs of Police, Alexandria
- Kelling GL, Bratton WJ (2006) *Policing terrorism*. Manhattan Institute for Policy Research, New York, Civic Bulletin 43
- Lum C, Kennedy LW, Sherley A (2006) Are counterterrorism strategies effective? The results of the Campbell systematic review on counter-terrorism evaluation research. *J Exp Criminol* 2:489–516
- Martin G (2009) *Understanding terrorism challenges, perspectives, and issues*. Sage, California
- Newman G (2009) Reducing terrorist opportunities: a framework for foreign policy. In: Freilich J, Newman G (eds) *Reducing terrorism through situational crime prevention*. Criminal Justice Press, USA, pp 33–59
- Newman G, Clarke R (2008) *Policing terrorism: an executive's guide*. U.S. Department of Justice, Washington, DC, Office of Community Oriented Policing Service (COPS)
- Perliger A, Hasisi B, Pedahzur A (2009) Policing terrorism in Israel. *Crim Justice Behav* 36:1279–1304
- Sherman L, Gottfredson D, MacKenzie D, Eck J, Reuter P, Bushway S (1997) *Preventing crime: what works, what doesn't, and what's promising*. University of Maryland, College Park
- Weisburd D, Waring E, Chayet E (2001) *White-collar crime and criminal careers*. Cambridge University Press, Cambridge
- Weisburd D, Jonathan T, Perry S (2009a) The Israeli model for policing terrorism: goals, strategies, and open questions. *Crim Justice Behav* 36:1259
- Weisburd D, Feucht T, Hakimi I, Mock L, Perry S (2009) Introduction. In: *To protect and to serve: policing in the years of terrorism, and beyond*. Springer, New York, pp 1–9
- Weisburd D, Hasisi B, Jonathan T, Aviv G (2010) Terrorism threats and police performance: a study of Israeli communities. *Br J Criminol* 50:725–747

Street Drug Market

- ▶ [Spatial Perspectives on Illegal Drug Markets](#)

Stress

- ▶ [Officer Safety, Health, and Wellness](#)

Structured Sentencing Outside the United States

Julian V. Roberts
Faculty of Law, Centre for Criminology,
University of Oxford, Oxford, UK

Overview

How much discretion should judges have at sentencing? If they have too much, disparity of outcome will be the inevitable consequence. On the other hand, removing their discretion entirely will result in injustice. The solution to this problem lies in the concept of sentencing guidelines which in theory at least provide direction to courts but without unduly restricting their discretion. The US-style grid-based guidelines are only one form of sentencing guidelines; other models are possible. This entry reviews the experience in jurisdictions other than the USA. Several countries have attempted to structure judicial discretion at sentencing in different ways. These other forms of guidance represent alternative approaches to promoting consistency at sentencing. After noting some developments in some other countries, this entry focuses on the experience in England and Wales which is the only common law jurisdiction outside the United States to have developed and adopted a comprehensive sentencing guidelines scheme.

Introduction

When most judges or other criminal justice professionals think about sentencing guidelines, the systems found across the United States come readily to mind. Sentencing guidelines have been evolving in that country for over 30 years now, at both the state and federal levels. Most states have a formal sentencing guidelines scheme to assist judges at sentencing. The best known guidelines model involves a two-dimensional sentencing grid – much like a mileage chart which shows the distance between two cities. Under a sentencing grid, the two dimensions are crime seriousness and criminal history. In order to determine the sentence that should be imposed, a court selects the appropriate level of seriousness and the appropriate criminal history category. Where the crime seriousness row and the criminal history column intersect, there is a grid cell containing a relatively narrow range of sentence length. Sentencing grids of this kind are found in a number of states including Minnesota (e.g., Minnesota Sentencing Guidelines Commission 2010a; Frase 2005, 2009).

If the guidelines are presumptively binding (some states use advisory guidelines, and the US federal guidelines are now considered advisory following several judgments from the US Supreme Court), the court must impose a sentence within the range found in the guidelines grid. A court wishing to impose a more or less severe sanction than that which is prescribed by the guidelines must first find “substantial and compelling” reasons to justify what then becomes a “departure” sentence, namely, one outside the official guideline ranges. Departure rates vary across the US systems. Statistics from Minnesota reveal that in 2009, 25 % of all felony offenders received a sentence different from that prescribed by the guidelines, while a further 14 % of all sentences of custody were outside the guidelines sentence length limits (Minnesota Sentencing Guidelines Commission 2010b, p. 26). The total departure rate for 2009 was therefore 38 %. Departure rates in other states are lower. In Pennsylvania, only 10 % of sentences imposed in 2008

fell outside the guidelines; over the period 1985–2008, the departure rate averaged 11 % (Pennsylvania Commission on Sentencing 2009, p. 46). In Oregon, less than one-fifth of sentences were outside the guidelines (Oregon Criminal Sentencing Commission 2003, p. 13).

Structuring judicial discretion at sentencing is one of the most significant challenges for a legislature. If they prescribe specific sentences – such as mandatory terms of custody – courts are prevented from doing justice by reflecting the individual circumstances of specific offenders. For example, legislating a mandatory sentence of imprisonment for all convictions of robbery means that offenders of different levels of culpability will receive the same sentence – a clear injustice. On the other hand, if legislatures leave the courts to regulate themselves at sentencing, outcomes may be too variable, leading to sentencing disparity. In short, it is a fine balance between offering too much and too little structure.

Sentencing Reform in Other Jurisdictions: Guidance by Words Alone

Sentencing guidelines do not have to be numerical in nature, providing a specific range of sentence for each crime. A number of Scandinavian countries have developed what may be termed “guidance by words” (see also Ashworth (2009) for discussion of techniques to reduce disparity through increased guidance). This approach to structured sentencing involves the legislature placing relatively detailed guidance in a sentencing law. For example, the Swedish Penal Code identifies proportionality as the primary rationale for sentencing and requires courts to assess the seriousness of the crime in order to determine sentence. A number of mitigating and aggravating factors are also specified in the Swedish sentencing law in order to guide judges in the determination of sentence. Finally, the law also contains guidance for courts with respect to the choices they should make between different sentencing options (for further information, see von Hirsch and Jareborg 2009). The advantage of

the “guidance by words” approach is that it leaves courts with considerable flexibility to determine an appropriate and proportionate sentence. On the other hand, this may result in much greater disparity than would be the case in a jurisdiction such as Minnesota where judges have to follow detailed and prescriptive sentencing guidelines.

Sentencing Structures: Commissions and Councils

How would a country go about creating a sentencing guidelines scheme? The first step in any move towards structuring judicial discretion involves the creation of an independent authority to develop and issue sentencing guidelines. All US guidelines schemes emerge from a sentencing commission, such as the Minnesota Sentencing Guidelines Commission or the US Sentencing Commission at the federal level. In other countries, these bodies are usually called Sentencing Councils, and there is significant variation in their structures and functions. The Sentencing Council of England and Wales is headed by the Lord Chief Justice and is tasked with devising and disseminating guidelines as well as a range of other functions (see Roberts 2011, 2012). On the other hand, sentencing councils in Australia such as the Sentencing Advisory Council in New South Wales are, as the name implies, advisory in nature. These councils do not issue sentencing guidelines per se but rather provide advice and conduct research upon a wide range of sentencing matters.

All sentencing councils are involved in public legal education of one kind or another. This may mean publishing reports to help the public understand the sentencing process better, or it may mean releasing comprehensive sentencing statistics. For example, some guidelines authorities publish periodic Sentencing Bulletins which summarize sentencing trends for specific offenses (see <http://sentencingcouncil.vic.gov.au/page/about-us/council>). The public typically rely on news media accounts of sentencing decisions, and these generally focus on unusual or exceptionally lenient sentences – those which

are newsworthy in some respect. It is important therefore for a guidelines authority to dispel public misperceptions of sentencing.

Progress towards developing sentencing guidelines around the world has been fitful and slow. The Law Commission of New Zealand developed a comprehensive and principled set of guidelines, but the legislature in that jurisdiction has yet to proclaim the necessary legislation to permit implementation (see Young and Browning 2008). The New Zealand scheme involved a comprehensive guideline for each offense; the guideline contained categories of crime seriousness, each with an associated range of sentence. A sentencing court would match the case appearing for sentencing to the guideline category using information in the guideline. The system was more flexible than the US-based schemes. The New Zealand proposed guidelines also included “generic” advice – guidelines which apply to more than a single offense. For example, the guidelines provide guidance on considering the impact of the crime upon the victim and also the way in which courts should approach the sentencing of multiple crimes on the same occasion.

Other jurisdictions – including Scotland, Western Australia, Northern Ireland, and Israel – have explored the use of guidelines for sentencers but so far have not actually adopted a formal scheme. The Sentencing Commission for Scotland recommended creation of an Advisory Panel on Sentencing to assist the introduction of sentencing guidelines, but this has yet to become operational (see Hutton and Tata 2010). Following recommendations from a Sentencing Working Group (2010), Northern Ireland held a consultation on the possible options for a form of sentencing guidelines (Criminal Policy Unit 2010).

In January 2012, the Israeli Parliament (Knesset) approved a sentencing law. This law adopted parts of a Bill which provide for “guidance by words” but without establishing the guidelines authority which would have been empowered to develop and issue guidelines scheme involving “starting point sentences” (see Gazal-Ayal and Kannai 2010). Under the new

legislation, courts are required to devise their own proportionate sentence range for the case being sentenced and to provide reasons if they impose a sentence outside this range. South Korea has also launched a guidelines scheme (see Park 2009). Several jurisdictions (including New South Wales and the state of Victoria in Australia) have created advisory bodies which disseminate information about sentencing but which do not actually disseminate guidelines (see Abadee (2008) and more generally, Freiberg and Gelb 2008).

Countries Without Guidelines

Finally, some countries – Canada, South Africa, Ireland, and India, for example – have resisted all appeals for greater structure at sentencing (e.g., Terblanche 2003; Doob 2011; Roberts et al. 2011). Although scholars and practitioners in those countries have long advocated creation of some kind of guidelines scheme, legislatures in these countries have so far rejected calls to introduce sentencing guidelines. The consequence is that judges in these jurisdictions continue to impose sentence much as they have for decades, with the only guidance coming from the appellate courts. This approach to sentencing may be termed “judicial self-regulation” (see Ashworth 2009). The limitation of this approach is that higher courts hear only a small proportion of cases on appeal, which means that the opportunities for guidance are limited. When a court of appeal does hear a sentence appeal, it does not always give general guidance.

Guidelines Structures

If appropriately constructed, and not subject to political interference, sentencing guidelines remain the best hope for constraining prison populations and achieving principled sentencing (see Stemen and Rengifo (2011); von Hirsch et al. (2009)), but the question remains: what form of guidelines is appropriate for any given

jurisdiction? The guidelines movement remains strong across the USA, but despite its high profile, the model employed in states such as Minnesota has not proven a popular penal export. Canada was the first country to reject this approach to structured sentencing.

In 1984, the Canadian government created a term-limited Sentencing Commission which visited several American states (including Minnesota and Pennsylvania) and concluded that two-dimensional grids held no promise for sentencing in Canada (see Canadian Sentencing Commission 1987). Canada's flirtation with guidelines ended in the 1990s when its Parliament approved a modest sentencing reform package and declined to create a permanent sentencing commission or to create guidelines of an even advisory nature (see Roberts and Cole 1999; Doob 2011). A generation later the Sentencing Commission Working Group in England and Wales visited the home of numerical guidelines and drew the same conclusion (Sentencing Commission Working Group 2008). Western Australia considered adopting a two-dimensional sentencing grid in 1999 but also ultimately abandoned the idea.

The proliferation of two-dimensional sentencing grids across the USA since the 1970s may paradoxically have undermined the appeal of all presumptively binding guidelines. Sentencing guidelines of any kind are often regarded by judges as harbingers of grids and as being antithetical to sentencing as a "human process" (see Hogarth 1971). Calls for the introduction of any kind of sentencing guidelines system are perceived as an attempt to move towards the ultimate goal of a grid. Indeed, opposition in Canada (and England and Wales) to sentencing guideline schemes of all stripes was fueled by predictions that any move towards structuring judicial discretion would culminate in the imposition of a rigid two-dimensional grid. In England and Wales, despite considerable judicial and professional resistance to the concept of guidance derived from a source other than the Court of Appeal, guidelines have slowly emerged over the past decade. Definitive guidelines now exist for most high-frequency offenses.

Sentencing Guidelines in England and Wales

Distinguishing Guideline Schemes in the USA and England

The English guidelines offer a different but equally viable model to follow, and there are lessons for other jurisdictions seeking to impose greater structure at sentencing without adopting a US-style two-dimensional grid. Guidelines have been developing slowly in England and Wales since creation of the Sentencing Advisory Panel in 1998. In addition, the Court of Appeal has been issuing guideline judgments periodically over the years. Courts in England and Wales have therefore a tradition of guidance on which to draw. However, with the creation of the Sentencing Council in 2010 (see Roberts 2011), the guidelines entered a new area, and the new Council began issuing guidelines more frequently and in a new format.

How do the English sentencing guidelines differ from the US-based schemes? The principal difference is that guidelines in England and Wales promote consistency by requiring sentencers to proceed through a series of steps. This represents a different approach to promoting consistency in sentencing. The effectiveness of the English approach has not been evaluated, as the Council responsible for issuing the guidelines has yet to publish the data which will shed light on any changes in consistency of sentencing outcomes. The second principal difference between the English and the US guidelines is that Sentencing Commissions across the USA devise and issue a complete package of guidelines encompassing all offenses – usually, as noted, within a single guidelines grid. The English Sentencing Council issues guidelines sequentially for each offense category (such as all assault offenses, or drugs crimes). This means that it takes longer for guidelines to be available for all offenses. (All English guidelines are available at <http://sentencingcouncil.judiciary.gov.uk/>).

Step-by-Step Sentencing Methodology

How do the English guidelines work? The English guidelines structure contains a series of

Offence: Assault occasioning actual bodily harm

Step 1: Use the factors provided in the guideline which comprise the principal elements of the offence to determine the category which is appropriate:

- Category 1: Greater harm *and* high culpability;
- Category 2: Greater harm *and* lower culpability *or* lesser harm *and* higher culpability;
- Category 3: Lesser harm *and* lower culpability.

Step 2: Use the *starting point* sentence from the appropriate *offence category* to generate a provisional sentence within the *category range*. The guideline contains a list of additional aggravating and mitigating factors which should result in upward or downward adjustment from the *starting point* sentence.

Offence Category	Starting Point Sentence	Category Sentence Range
1.	18 months custody	12 months – 3 years' custody
2.	6 months custody	Community order – 12 months custody
3.	Community order	A fine – community order

Step 3: Consider if any reduction in the provisional sentence should be made to reflect assistance offered or provided to the prosecution.

Step 4: Consider the level of reduction appropriate to reflect a guilty plea.

Step 5: Consider whether the offender meets dangerousness criteria necessary for imposition of an indeterminate or extended sentence.

Step 6: If sentencing for more than one offence apply the totality principle to ensure that the total sentence is just and proportionate to the total offending behaviour.

Step 7: Consider whether to make a compensation order and/or other orders.

Step 8: Give reasons for and explain the effect of the sentence on the offender.

Step 9: Consider whether to give credit for time on remand or bail.

(For further information, see Sentencing Council (2012); Roberts and Rafferty (2011)).

Structured Sentencing Outside the United States, Fig. 1 Example of sentencing guidelines format in England and Wales

nine steps, of which the first two are the most critical. (See Fig. 1 for a summary of the steps). The idea is that if all courts follow the same step-by-step procedure, sentencing decisions across courts should be more consistent.

For the purposes of illustration, let us examine the guideline for *Assault occasioning actual bodily harm* (see Sentencing Council of England and Wales 2011). As with many offenses for which a definitive guideline has been issued, ABH has been stratified into three categories of seriousness. The guideline provides a separate range of sentence and starting point sentence for each category. Step 1 of the guidelines methodology requires the sentencing court to match the

case appearing for sentencing to one of the three categories of seriousness. The three categories have been created to reflect gradations in harm and culpability, with the most serious category (1) requiring greater harm and enhanced culpability. Category 2 is appropriate if either greater harm or higher culpability is present, while Category 3 involves lesser harm and a lower level of culpability.

Determining the Offense Category

Step 1 of the guideline identifies an *exhaustive* list of sentencing factors which should be used to determine which of the three categories of seriousness is most appropriate for the particular

offender appearing for sentencing. These factors constitute what the guideline describes as the “principal factual elements” of the offense. Their importance is reflected in the fact that determination of the category range is the step which will have greatest influence on severity of sentence. For example, if the court chooses the lowest category, the most severe sentence it should impose is a community-based (noncustodial) sentence. In contrast, the next category up has a maximum sentence of 51 weeks imprisonment. Having determined the relevant category range, a court should use the corresponding starting point sentence to move towards a sentence which will then be shaped by the remaining steps in the guideline.

Step Two: Shaping the Provisional Sentence

Step 2 requires a court to “fine-tune” the sentence by considering circumstances which provide “the context of the offense and the offender.” The guidelines provide a non-exhaustive list of sentencing factors for courts to consider at this step of the methodology. The aggravating factors include committing the offense while on bail or license, while under the influence of drugs, or with abuse of trust. The guideline factors which reduce seriousness (and which result in less severe sentence) include an absence of prior convictions and the fact that the crime was an isolated incident. A diverse collection of factors is cited as relevant to personal mitigation, including remorse, the fact that the offender was a sole or primary carer and “good character and/or exemplary conduct.”

Other Steps Towards the Final Sentence

Leaving Step 2 leads a court to the remaining seven steps of the guidelines methodology, which may be briefly summarized. Step 3 directs courts to reduce sentence in cases where the offender has provided or offered to provide assistance to the prosecution or police. Step 4 involves the reduction for a guilty plea. Offenders who plead guilty receive a reduction of up to one-third off their custodial sentence, depending upon how early they enter their plea. The remaining steps to be followed relate to other

sentencing provisions such as the consideration of any time that the offender spent in remand while awaiting trial. (This time is taken off any custodial sentence ultimately imposed). The guideline steps also require the court to give reasons for the sentence imposed and to explain the effect of the sentence for the benefit of the offender. Finally, courts give reasons for the sentence and, in particular, if they have imposed a sentence which falls outside the guidelines range.

Legal Status of English Guidelines

As noted, in many US states, the sentencing guidelines are binding on courts: a judge may depart from the guideline range if he or she finds “substantial and compelling” reasons why the guideline sentence is inappropriate. The English guidelines are less restrictive, although they are formally binding on courts. When sentencing an offender in England and Wales, a sentencing court “must follow” any relevant sentencing guidelines – unless it would be contrary to the interests of justice to do so (see discussion in Ashworth 2010). However, courts may impose a sentence within a wide range and still remain compliant with the guidelines. In this sense, although they are required to follow the guidelines, courts retain considerable freedom to sentence offenders: first, they have considerable discretion within the guidelines, and second, they may depart from the guidelines if they believe it is necessary in the interests of justice.

Conclusion

What have we learned about the experience with structuring sentencers’ discretion outside the United States? A number of lessons can be drawn. First, the Scandinavian model suggests that numerical guidelines are not necessarily the only model to follow. It is possible to offer guidance to courts without prescribing specific sentencing ranges in terms of numbers of months or years. Second, the English guidelines demonstrate that there is a middle ground lying between the relatively tight sentencing guidelines found

across the United States and the looser systems of “guidance by words” found in countries like Sweden and Finland (see von Hirsch et al. 2009, Chapter 6). Along with the New Zealand proposals, the English guidelines offer a system which is numerical, prescriptive, and yet quite flexible in application. These two systems offer a plausible improvement upon the highly discretionary sentencing arrangements found in countries like Canada, South Africa, and India. Third, judicial acceptance of greater structure (and reduced discretion) is far more likely when the judiciary is heavily implicated in the development and evolution of the guidelines. The statutory bodies responsible for the guidelines in England and Wales have generally been dominated by the judiciary. The Canadian Sentencing Commission proposals failed, in part because judges perceived the guideline scheme to be a bureaucratic scheme created by academics.

Fourth, there may be an advantage to the gradual evolution of the guidelines. The English guidelines have been criticized for being slow to develop – with guidelines for specific offenses issued periodically over the years rather than in one step as was the case in the United States or the proposals advanced in New Zealand. There was no “big bang” to the English experience in the sense of a comprehensive guidelines package covering all offenses. In retrospect, this potential weakness of the guidelines may paradoxically have ensured their survival and development. They have evolved incrementally – from their modest origins in 1999 (see Ashworth and Wasik 2010) through to the much more comprehensive and detailed scheme of 2011. Judges who are traditionally resistant to any attempts to curb their discretion may be more likely to accept guidance when it comes in this format.

The ultimate question to pose, however, is the following: Are the guidelines proposed or implemented in other countries better or worse, more or less effective than those developed across the United States? Unfortunately the absence of truly comparative research makes it impossible to resolve the issue one way or another. In addition, the non-US-based guidelines such as those in England have yet to be

evaluated. For example, we do not yet know how much the English guidelines have improved consistency in that country. At the very least, however, the experience in that country demonstrates that it is possible to introduce detailed and prescriptive sentencing guidelines even in a common law jurisdiction which, in the 1980s and 1990s, was committed to the traditional model of privileging judicial discretion (see Ashworth 2010).

Related Entries

- ▶ [Mitigation and Aggravation at Sentencing](#)
- ▶ [Sentencing as a Cultural Practice](#)

Recommended Reading and References

- Abadee A (2008) The New South Wales sentencing council. In: Freiberg A, Gelb K (eds) *Penal populism, sentencing councils and sentencing policy*. Willan, Cullompton
- Ashworth A (2009) Techniques for reducing sentence disparity. In: von Hirsch A, Ashworth A, Roberts JV (eds) *Principled sentencing. Readings on theory and policy*, 3rd edn. Hart Publishing, Oxford
- Ashworth A (2010) *Sentencing and criminal justice*, 5th edn. Cambridge University Press, Cambridge
- Ashworth A, Wasik M (2010) Ten years of the sentencing advisory panel. In: *Sentencing advisory panel annual report*. Available at: www.sentencing.council.org
- Canadian Sentencing Commission (1987) *Sentencing reform: a Canadian approach*. Supply and Services Canada, Ottawa
- Criminal Policy Unit (2010) *Consultation on a sentencing guidelines mechanism*. Northern Ireland Department of Justice, Belfast
- Doob AN (2011) The unfinished work of the Canadian Sentencing Commission. *Can J Criminol Crim Justice* 55:279–298
- Frase R (2005) Sentencing guidelines in Minnesota, 1978–2003. In: Tonry M (ed) *Crime and justice*. University of Chicago Press, Chicago
- Frase R (2009) Sentencing policy under the Minnesota sentencing guidelines. In: von Hirsch A, Ashworth A, Roberts JV (eds) *Principled sentencing. Readings on theory and policy*, 3rd edn. Hart Publishing, Oxford
- Freiberg A, Gelb K (eds) (2008) *Penal populism, sentencing councils and sentencing policy*. Willan, Cullompton
- Gazal-Ayal O, Kannai R (2010) Determination of starting sentences in Israel – system and application. *Fed Sentencing Report* 22:232–242

- Hogarth J (1971) *Sentencing as a human process*. University of Toronto Press, Toronto
- Hutton N, Tata C (2010) A sentencing exception? Changing sentencing policy in Scotland. *Fed Sentencing Report* 22:272–278
- Minnesota Sentencing Guidelines Commission (2010a) *Minnesota sentencing guidelines and commentary*. Revised, Aug 2010. MSGC, St. Paul
- Minnesota Sentencing Guidelines Commission (2010b) *Sentencing practices. Annual summary statistics for felony offenders sentenced in 2009*. MSGC, St. Paul
- Oregon Criminal Sentencing Commission (2003) *Sentencing practices, 2001*. State of Oregon, Criminal Sentencing Commission, Salem
- Pennsylvania Commission on Sentencing (2009) *Sentencing in Pennsylvania. Annual report 2008*. Available at: <http://pcs.la.psu.edu/>
- Park H (2009) The basic features of the First Korean sentencing guidelines. Korean Ministry of Justice, Seoul
- Roberts JV (2011) Sentencing guidelines and judicial discretion: evolution of the duty of courts to comply in England and Wales. *Br J Criminol* 51:997–1013
- Roberts JV (2012) Structured sentencing: lessons from England and Wales for common law jurisdictions. *Punishment and society. Int J Penol* 14(3):267–288
- Roberts JV, Azmeh U, Tripathi K (2011) Structured sentencing in England and Wales: lessons for India? *Natl Law School India Rev* 23:27–46
- Roberts JV, Cole D (eds) (1999) *Making sense of sentencing*. University of Toronto Press, Toronto
- Roberts JV, Rafferty A (2011) Structured sentencing in England and Wales: exploring the new guideline format. *Crim Law Rev* 9:680–689
- Sentencing Commission Working Group (2008) *Sentencing guidelines in England and Wales: an evolutionary approach*. SCWG, London
- Sentencing Council of England and Wales (2011) *Assault. Definitive guideline*. Available at: <http://sentencingcouncil.judiciary.gov.uk/>
- Sentencing Working Group of Northern Ireland (2010) *Monitoring and developing sentencing guidance in Northern Ireland*. Northern Ireland Department of Justice, Belfast
- Stemen D, Rengifo A (2011) Policies of imprisonment: the impact of structured sentencing and determinate sentencing on state incarceration rates, 1978–2004. *Justice Q* 28:174–201
- Terblanche S (2003) Sentencing guidelines for South Africa: lessons from elsewhere. *S Afr Law J* 120:858–882
- von Hirsch A, Ashworth A, Roberts JV (eds) (2009) *Principled sentencing. Readings on theory and policy*, 3rd edn. Hart Publishing, Oxford
- von Hirsch A, Jareborg N (2009) The Swedish sentencing law. In: von Hirsch A, Ashworth A, Roberts JV (eds) *Principled sentencing. Readings on theory and policy*, 3rd edn. Hart Publishing, Oxford
- Young W, Browning C (2008) New Zealand's sentencing council. *Crim Law Rev*, April 2008, pp 287–298

Subsidized Housing

- ▶ [Public Housing and Crime Patterns](#)

Suicide

- ▶ [Officer Safety, Health, and Wellness](#)

Suicide and Prisons

Gregory F. Kellermeier¹ and Hal S. Wortzel²

¹Denver Health Medical Center and Division of Forensic Psychiatry, Department of Psychiatry, University of Colorado Denver School of Medicine, Denver, CO, USA

²Division of Forensic Psychiatry, Department of Psychiatry, University of Colorado Denver School of Medicine and VISN–19 MIRECC, Denver Veterans Affairs Medical Center, Denver, CO, USA

Overview

An increasing body of literature reflects the heightened risk for suicide faced by justice-involved persons, whether they are in jails, in prisons, or in the process of transitioning from a correctional facility back into the community. Suicide and suicide prevention have garnered increased attention, efforts, and resources within correctional facilities. Despite this increased awareness of suicide and heightened efforts to prevent suicide in correctional settings, suicide rates among incarcerated persons in the United States have continued to exceed that of the general population (albeit less so than in previous decades). Pertinent case law clearly creates potential liability for both correctional and health-care staff in the wake of a death by suicide. This entry will review applicable case law; explore statistical trends, risk factors, and characteristics of suicides in correctional settings;

review methods for prevention of suicide supported by the literature; discuss select special populations of incarcerated individuals at risk; and explore the suicide risk surrounding the transition from correctional settings back into the community. This entry will validate the importance of the issue of suicide in correctional settings and highlight areas where further research and progress are needed.

Case Law History

The potential for legal liability resulting from suicide in correctional settings originates in two seminal cases: *Estelle v. Gamble* (1976) and *Bowring V. Godwin* (1976). In the *Estelle* case, the plaintiff filed suit claiming that the medical director and two correctional officers of a state correctional department violated the Eighth Amendment (cruel and unusual punishment) by not providing appropriate medical care for the inmate's back pain. Upon appeal, the US Supreme Court established the legal standard of "deliberate indifference" and by using this standard held that the inmate's medical care did not violate the 8th Amendment. Deliberate indifference requires more than mere negligence. Negligence typically involves the failure to exercise the care toward others which a reasonable or prudent person would do in similar circumstances, with the resulting harm generally being accidental. A reckless culpable mental state typically involves gross negligence, meaning that the individual could foresee the potential for harm, but was heedless of the consequences. The deliberate indifference standard is more akin to recklessness, necessitating that in order to be deemed liable, prison staff must have the subjective knowledge of risk for serious harm and then disregard that risk.

In the *Bowring* case, an inmate in a state correctional facility claimed Eighth Amendment (cruel and unusual punishment) and Fourteenth Amendment (due process) violations in lower court due to being denied parole, in part as a result of a psychological opinion that he would not be a successful parolee. The plaintiff

argued that he was not provided the psychiatric or psychological diagnosis or treatment needed to address the psychological concerns of the parole board. The Fourth Circuit Court of Appeals held that psychiatric illnesses are medical in nature, thereby extending the treatment obligations incumbent upon correctional facilities to mental illness. The combination of these two seminal cases sets the stage for subsequent decisions finding liability in failure to prevent inmate suicide.

Concerning suicide specifically, circuit-level court findings in *Partridge v. Two Unknown Police Officers* (1986) and *Colburn v. Upper Darby Township* (1991) establish inmate suicide as a potential source of liability for correctional facilities (Scott 2010). The *Partridge* Fifth Circuit court determined that psychiatric illnesses potentially resulting in deaths by suicide constitute a "serious medical need." In *Colburn v. Upper Darby Township* (1991), the Third Circuit court established that correctional facilities have an obligation to train staff members to identify potentially suicidal inmates and to intervene in an effort to prevent suicide deaths. The court determined that the standard for correctional staff's competency in identifying potentially suicidal inmates would be comparable to that of a layperson's and that correctional staff would be held to a standard of "reckless indifference" when legally evaluating efforts to intervene and prevent an inmate death by suicide. The Third Circuit court confirmed the lower court's analysis, indicating that a plaintiff must prove not only that the deceased inmate had a "particular vulnerability to suicide" but that custody staff knew or should have known about the vulnerability and that the custody staff acted with "reckless indifference" in relation to the individual's particular vulnerability. Existing case law clearly sets a bar for correctional facilities' role in suicide prevention, though sets that bar relatively low.

Suicide Rates in Correctional Settings

Historically, suicide rates have been exceedingly high within correctional settings relative to the general population. In 1983, the suicide rate for

jails was 129 per 100,000 inmates, and suicide accounted for 56 % of all jail deaths (Noonan 2010). Fortunately, suicide rates in both jails and state prisons have decreased since the 1980s. From 2000 to 2007, the suicide rate within jails decreased by approximately 25 %, from 48 per 100,000 to 36 per 100,000. Despite this relative decline in suicide rates, suicide remained the single leading cause of preventable death in jails. When jail suicide deaths rates from 2000 to 2007 are adjusted for demographic factors such as age, race, and sex, the rates still remain high relative to those observed in the general population in the United States (14 per 100,000). Overall, the adjusted rate for suicide in jail is 59 per 100,000 for males and 27 per 100,000 for females. Comparable rates of suicide in the US general population are 21 per 100,000 for males and 5 per 100,000 for females. It is worth noting that suicide rates within jails may vary substantially depending on the size of (and presumably resources available within) the correctional facility. Large jails had a lower suicide rate (27 per 100,000) relative to small jails (167 per 100,000), with small jails being defined as those housing 50 or fewer inmates. Noonan theorizes that a lack of suicide prevention measures, including appropriate mental health services and more inmate turnover in smaller jails, may account for this difference (2010). State prisons have also demonstrated a fortunate downward trend in suicide rates over the past few decades, though they still exceed that of the general population. The rate of suicide was 34 per 100,000 in 1980 (Mumola 2005) and decreased to a rate of 16 per 100,000 for the years 2000–2007. This rate has held fairly stable since the 1990s (Noonan 2010). Federal prison suicide rates are reported at 10 to 17 per 100,000 (White et al. 2002).

A poignant criticism has been raised regarding the method employed in measuring the rate of suicide in jails. These reported rates are based upon the total national average daily inmate population rather than the number of admissions to the specific incarceration facilities (Hayes 2010). If the rates were calculated based upon the number of admissions rather than the average daily population, calculated suicide rates would be

lower. Calculating suicide rates based on the total number of admissions would arguably be more accurate because this method takes into account the high turnover of jails and thus more fully captures the at-risk population. Larger jails release approximately 46 % of inmates within the first 2 days of confinement and release 80 % of their population within an inmate's first month of incarceration. Calculations based on admissions would thus more fully take into account all the unique individuals incarcerated. Others argue that community suicide rates are not calculated on total "admissions" to the population, and moreover, the majority of admissions are brief incarcerations and thus should not be included in the statistical calculation. But, as will be discussed below, many suicides occur very early after incarceration, illustrating the potential importance of capturing all admissions when calculating rates.

Incarceration within correctional facilities does seem to carry increased risk for suicide. Why has the rate of suicide been high in correctional settings, and furthermore, why has the rate decreased since the early 1990s? Various investigations and expert opinions have proposed several possible explanations.

Prevalence of Mental Illness in Correctional Settings

One significant risk factor for suicide, applicable to individuals across settings (whether they are in jails, prisons, or the general community), is a diagnosis of mental illness. In one study of individuals who died by suicide in jails between 2005 and 2006, 38 % suffered from mental illness, 20 % had a history of being prescribed psychiatric medications, and 34 % had a history of self-directed violence, including suicide attempts and nonsuicidal self-directed violence (Hayes 2010). Patterson reported that 73 % of the California prison inmates who died by suicide in a 6-year period had a history of mental health treatment and 62 % had a history of suicidal thoughts or behavior (2008). Fazel conducted a systematic review of suicides in prison and

found that history of attempted suicide, current psychiatric diagnosis, and current prescription for psychotropic medication were risk factors for suicide in prison (2008).

Rates of suicide in correctional settings are almost certainly related to the prevalence of mental illness. For jails and prisons, the number of admitted individuals suffering mental illness has increased over the last half of the twentieth century. Since the 1960s, the number of long-term care mental health institutions (such as state mental health hospitals) has decreased under the deinstitutionalization movement (Lamb and Weinberger 2005). Deinstitutionalization evolved from legal and ethical concepts favoring treatment of mentally ill individuals in the least restrictive environment possible. Under this theory, individuals would be released from highly restrictive psychiatric facilities and would instead be treated and supported in community-based mental health centers. Unfortunately, many experts agree that the deinstitutionalization movement has largely failed due to the inadequate creation and funding of needed community mental health centers, such that many severely, chronically mentally ill individuals are not receiving the treatment and support needed to sustain tenure in the community (Scott 2010). In the absence of such support, many of these mentally ill persons go on to engage in behaviors that eventually lead to incarceration. Hence, the failed deinstitutionalization movement has, over time, resulted in a shift in location for the severely mentally ill, from mental health facilities to correctional settings. In 1955, there were 559,000 psychiatric hospital beds for a total United States population of 165 million, compared to 59,403 psychiatric hospital beds for a total United States population of 273 million in 2000 (Lamb and Weinberger 2005). The notion that individuals suffering from a mental illness have shifted institutional settings, from mental health hospitals to correctional facilities, has been termed the criminalization of the mentally ill (Scott 2010). Other proffered explanations for this shift include lack of resources for community mental health care, peace officers assuming the role of triaging individuals to hospitals versus jail, and more

restrictive involuntary psychiatric hospitalization laws (Lamb and Weinberger 2005).

Simultaneous to the deinstitutionalization of the mentally ill, the United States began its incarceration binge which has resulted in an increased rate and prevalence of incarceration. The total number of inmates in the United States in 1980 was 501,886, compared to over 2.2 million in 2009 (Glaze 2010). As a result, the total number of mentally ill incarcerated individuals increased with the incarceration binge. A study by Bradley-Engen supports this additional explanation (2010). The study reported the change in the total number of individuals admitted to Washington State prisons from 1998 to 2006 and the relative proportion of inmates diagnosed with serious mental illness and co-occurring substance use at time of admission. The results of the study suggest that the increase in the diagnosis of serious mental illnesses and co-occurring substance use disorders is related to the increase in total admissions to prisons during the study time period and is not due to an increase in the percentage of admissions involving serious mental illness. In other words, this study suggests that at least some of the increase of the total number of incarcerated mentally ill is a result of the overall higher number of individuals being incarcerated. Collectively, it seems that an unfortunate combination of factors, chiefly involving the failed deinstitutionalization movement and a criminal justice system featuring a strong predilection toward incarceration, has resulted in growing populations of mentally ill persons behind bars in the United States.

Estimates vary for prevalence rates of mental illness in jails and prisons depending on the method of the study. The Bureau of Justice Statistics reported that 56 % of state prisoners, 45 % of federal prisoners, and 64 % of jail inmates suffered from some type of mental health problem (James et al. 2006). Bradley-Engen reported that past studies suggest that the prevalence of serious mental illness in prisons ranges from 16 % to 24 %, compared to 5–7 % in the general community (2010). A recent study suggests that the rate of serious mental illness in jail is 14.5 % for men and 31.0 % for women (Steadman et al. 2009).

Risk Factors for Suicide Pertinent to Correctional Settings

Besides housing an influx of many individuals that might have otherwise been institutionalized in mental health facilities during an earlier era, the correctional setting tends to contain populations with high rates of other important risk factors for suicide. Many of the risk factors associated with inmates' suicide deaths match risk factors associated with suicide deaths in the community: male gender, Caucasian race, substance abuse, psychiatric treatment history, and prior suicide attempts (Anasseril 2006; Patterson and Hughes 2008; Hayes 2010). Fazel reported that being a male and being Caucasian were risk factors for suicide in prison; he also found that being African American was inversely related to suicide (2008). Noonan reported that the suicide rate for African Americans was lower compared to Hispanic and Caucasian inmates (2010). However, it is vital to appreciate that demographic risk factors of this kind are of very limited utility in determining risk at the single inmate level, keeping in mind that favorable demographic factors identified via population-based research do not keep individuals safe; inmates of any race or gender may of course present at heightened risk for suicide.

Complicating risk assessment in correctional settings is the fact that so many of the identified risk factors for suicide occur at a very high prevalence, higher than the prevalence in the general community. For example, inmates in correctional settings are predominately male (Glaze 2010). The male gender is dramatically overrepresented in correctional settings, is overrepresented in suicide deaths in the general community, and also represents approximately 93 % of jail suicides (based on figures from 2005 to 2006) (Hayes 2010). In federal prisons, despite having a 7 % female population, from 1993 to 1997, 100 % of suicides were completed by males (White et al. 2002). Noonan reports that white males in jail have a rate of suicide 5.7 times greater than white males in the general population (2010).

Substance use disorders also have a high prevalence in correctional settings. As a conservative

estimate, 50 % of incarcerated individuals have substance use disorders, with jail inmates having a higher prevalence rate than prison inmates (Scott 2010). Hayes found that of the individuals that died by suicide in jail between 2005 and 2006, 47 % had a history of substance abuse; 20 % were intoxicated at the time of the suicide (2010). In New York state prisons between 1993 and 2001, 95 % of inmates who died by suicide had a history of substance abuse (Way et al. 2005).

Many incarcerated individuals have experienced violent trauma during their lifetime (Wolff et al. 2010). As cited by Wolff, at least 50 % of incarcerated women have reported victimization in their lifetime. Additional profile characteristics for individuals that committed suicide in prisons are low education level and limited social supports in the community (White et al. 2002; Anasseril 2006). The frequency of trauma, victimization, and limited support among correctional populations seems particularly relevant to suicide risk when one considers the interpersonal-psychological theory of suicide risk offered by Thomas Joiner (2005). Under this model, the convergence of acquired ability (relating to habituation to pain and suffering along with perceived burdensomeness and/or a sense of failed belongingness results in a heightened risk for death by suicide). The above-referenced research indicates that inmates are frequently themselves victims of trauma, potentially contributing to habituation to pain and suffering, and thus acquired ability to engage in suicidal behaviors. Additionally, limited psychosocial support and the impact incarceration has on families likely engender feelings of burdensomeness and failed belonging. It would seem that the convergence described by Joiner is probably a common occurrence among inmates and may in part explain the observed elevation in suicide rates.

The high rates of violent perpetration and aggression among correctional populations are a consideration when thinking about suicide risk and suicide rates. Violent offenders have been found to face increased suicide rates across a number of studies, with the Bureau of Justice statistics reporting a suicide rate nearly three

times that of nonviolent offenders in jail, and violent offenders in prison being more than twice as likely to die by suicide. Hayes reported that 43 % of inmates that died of suicide had been arrested on personal and/or violent charges, including murder, negligent manslaughter, armed robbery, rape, sexual assault, indecent assault, child abuse, domestic violence, assault, battery, aggravated assault, and kidnapping (2010). Noonan also found that violent offenders had the highest suicide rate, with public-order offenders having the second highest rate of suicide, compared to individuals with property and drug offenses (2010). Inmates incarcerated for the alleged offenses of homicide, kidnapping, and rape had the highest rate of suicide (Noonan 2010). Notably, aggression features prominently in a theoretical model for suicide risk offered by Kerr et al. (2007). Under that model, aggression may lead to suicide attempts via various routes: aggression may lead to depression and then to suicidal ideation; aggressive acts may yield negative consequences leading to worsening depression and suicidal ideation; or aggression may more directly be turned inwardly resulting in self-directed violence and suicide attempts.

The incarcerated individual's coping styles, changing emotional states, or emotional traits have also been associated with suicide risk. Individuals that have a maladaptive style of interacting with the environment and coping with stress have increased risk for suicide and self-directed violence. Specifically, rigidity and dichotomous thinking, poor problem-solving ability, avoidant behavior, and poor positive reinterpretation are personality styles that lend themselves to suicidal behavior (Polluck and Williams 2006). Emotional states or traits involving anxiety, depression, aggression/agitation, hostility, fear, and impulsivity are characteristic of individuals who have suicidal behavior (Polluck and Williams 2006; Anasseril 2006; Way et al. 2005). An individual's self-report of feeling hopeless is arguably a highly significant risk factor for suicide in most if not all populations, including incarcerated populations (Anasseril 2006). Recent "behavior changes" of incarcerated individuals from baseline observations also

has been noted as a characteristic profile of individuals who are at risk for suicide in prison (Way et al. 2005). For example, prior to some suicide deaths, staff members report that the inmate had not been acting in his or her usual manner.

Psychosocial stressors unique to incarceration and the criminal justice process may also contribute to increased suicide risk. For example, many jail and prison suicides occur during the initial incarceration phase or after transfer to a new facility (Anasseril 2006; Hayes 2010). After arrest and initial incarceration in jail, an individual has gone from living in a free society to living in isolation, or possibly in a cell under 23 h lockdown where freedoms are severely curtailed and most aspects of daily life kept under strict control (Patterson and Hughes 2008; Hayes 2010). Many recent detainees are unclear of the specific consequences of their legal charges or have not spoken with an attorney, are intoxicated or experiencing withdrawal from drugs or alcohol, and/or have not yet communicated with their primary supports in the community. Some detainees are experiencing guilt or shame concerning their charge and/or have just experienced a personal/family crisis that is related to their charge (Hayes 2010). Although the above-described stressors are challenging for all inmates, individuals from higher social economic backgrounds incarcerated for the first time seem to be particularly susceptible to such stressors and at higher risk for suicide (Anasseril 2006).

For longer stays in jail or prisons, stressors from both within the correctional setting or crises occurring outside the correctional setting can contribute to suicide risk. Specifically, death of a loved one, relationship problems/losses, financial losses, new charges or setbacks in court, gang problems, victimization, bullying, recent experience or risk of sexual assault, and recent institutional disciplinary action are often antecedents of inmate deaths by suicide (Way et al. 2005; Anasseril 2006). A life sentence is considered a risk factor for suicide in prisons (Fazel et al. 2008). Death row inmates have the highest suicide risk (Anasseril 2006). Although counterintuitive to some, marriage was found to be a risk factor for suicide in prisons and during periods of

longer incarceration in jails (Fazel et al. 2008; Hayes 2010). It is interesting to note that jail and prison suicides were not associated with certain holidays but evenly distributed through the year (White et al. 2002; Anasseril 2006; Hayes 2010).

Despite curtails in freedom and relatively limited access to means (compared to life in the community), the correctional environment does offer sufficient opportunity for inmates that are having suicidal thoughts to engage in lethal acts of self-directed violence (Hayes 2010). A one-person cell, time spent alone in a cell, and access to lethal means (such as anchor points for hanging) provide the necessary means for an inmate to die by suicide (Patterson and Hughes 2008; Hayes 2010). Fazel found that single-cell occupation is a risk factor for suicide (2008). The most common method of suicide in both jails and prisons is hanging, featuring in over 80 % of suicides (Patterson and Hughes 2008; Hayes 2010). Other methods employed include overdose, self-inflicted laceration resulting in exsanguination, self-strangulation, swallowing objects, electrocution, poisoning, self-immolation, self-inflicted gunshot wound, and intentional fall (Felthous 2011). All these techniques are more easily accomplished when an inmate is isolated.

Suicide Prevention Programming

The primary method to mitigate suicide risk and implement protective factors is to create an effective comprehensive suicide prevention program in individual correctional facilities (Hayes 2010). Effective suicide prevention programs have been credited with reduction of suicides in correctional settings. National organizations and credentialing bodies such as the American Psychiatric Association, the National Commission on Correctional Health Care, and the American Correctional Association all recommend that correctional facilities have suicide prevention programs and articulate specific standards for these programs. Hayes recommends the following eight essential components for a comprehensive suicide prevention program: staff training, intake screening/assessment, communication, housing, levels of

observation, intervention, reporting, and follow-up and morbidity-mortality review (Hayes 2010). A brief summary highlighting the major aspects of Hayes' recommendations for a comprehensive suicide prevention program follows.

Hayes states that all custody, mental health, and medical staff should initially have 8 h of training and 2 h of training annually (Hayes 2010). Training custody staff is crucial due to the large amount of time they spend with inmates, resulting in their significant role in identifying potentially suicidal inmates. Hayes asserts that staff training should include, at a minimum, "reasons why correctional environments are conducive to suicidal behavior, staff attitudes about suicide, potential predisposing factors to suicide, high-risk suicide periods, warning signs and symptoms, identification of suicide risk despite the denial of risk, liability issues, critical incident stress debriefing, [review of] recent suicides and/or serious suicide attempts within the facility or agency, and components of the facility's or agency's suicide prevention policy" (Hayes 2010).

Hayes recommends that individuals receive screening and assessment for suicide risk upon intake to a correctional facility (Hayes 2010). Hayes asserts that screening should include questions concerning suicidal thoughts/plans (including such thoughts during past incarcerations), past suicidal attempts, history of mental health treatment, recent stressors and losses, suicide behavior by family or close friends, and reports from arresting officers concerning suicide risk (Hayes 2010). As in other settings, suicide risk assessment is a process, not an event; suicide screening does not end after intake, but continues throughout the individual's incarceration. Inmates' self-reporting should be considered in concert with behavioral observations and other collateral information in determining the level of suicide risk.

Communication between arresting officer, correctional staff, mental health staff, and medical staff is essential to reduce suicide risk (Hayes 2010). Protocols for appropriate suicide precautions should be in place to enable staff to take appropriate action if an inmate is considered to be at risk for suicide in order to keep the inmate safe. Hayes suggests that custody staff should not be

underestimated in their importance in identifying potentially suicidal inmates and communicating this concern to mental health staff (2010). In additions, barriers need to be reduced in order for inmates to obtain mental health care.

Appropriate housing of inmates with significant suicide risk is an important component of a proper suicide prevention program (Hayes 2010). Inmates with clinically determined differing levels of suicide risk should be housed in the least restrictive environment needed to provide the appropriate level of safety, such as medical/mental health units or general population housing where staff can readily interact with them. Only individuals with the highest acuity of suicide risk should be housed in a locked-down environment, and only when the appropriate suicide precautions are provided. Although segregation may be the safest option for individuals at imminent suicide risk, isolation may also be detrimental to the inmate's emotional well-being and seldom addresses the factors driving the crisis. Careful consideration regarding the risks and benefits of segregation is warranted. Cells used for individuals at acute risk for suicide should be free of methods which might be employed to act upon suicidal thoughts, such as anchors or hanging devices in the cell or on the inmate. Level of observation of a suicidal inmate in a particular housing setting depends on the particulars of the individual's suicide risk (Hayes 2010). Observation should be, at a minimum, staggered in 15-min intervals and, at least for inmates judged to be at imminent risk, should consist of continuous in-person observation. Cameras in cells can help monitor inmates at acute risk for suicide but do not replace the need for continuous in-person observation. Inmates flagged for suicide precautions should be closely followed for evaluation and treatment by mental health staff.

All staff with regular contact with inmates should be provided with first aid, cardiopulmonary resuscitation, and automated external defibrillator training to be prepared to intervene after a suicide attempt (Hayes 2010). Staff should be prepared to quickly triage the emergency situation, contact appropriate medical staff, and

initiate the appropriate above treatment(s). Staff should not assume that the individual is deceased. If a serious suicide attempt or completed suicide has occurred, family and appropriate authorities should be notified immediately. Morbidity-mortality reviews and psychological autopsies are recommended. Within 24–72 h, Hayes recommends that staff involved in the incident be offered critical incident stress debriefing by a trained professional (Hayes 2010).

Correctional facility culture and environment seem to be related to suicide risk. Liebling argues that inmates that are already vulnerable to suicide and self-directed violence face even higher suicide risk if the institution creates feelings of lack of safety, lack of respect/fairness, alienation, and frustration (2006). In fact, she argues that institutions that provide individual support create a protective factor. She recommends that institutions focus on improving both correctional custody and health-care staff culture, specifically avoiding environments that use excessive authority, have a lack of personalization of the individual inmate, and avoid addressing or listening to inmate complaints. An international study analyzed possible interventions or factors that reduced the incidence of suicide in prisons. "Purposeful activity" (as defined as classes or activities for inmates, such as education, substance abuse treatment, work, and family visits) was significantly associated with lower suicide rates (Leese et al. 2006). In a study conducted in one jail, protective factors included reasons for living and social connectedness.

Special Populations

Special populations mandate specific considerations when evaluating the risk for suicide and self-directed violence within correctional settings. A few examples follow.

Adolescents

In the United States general population, suicide is one of the leading causes of death for adolescents. Suicide is the fourth leading cause of death of individuals between the ages of 10 and 14, while suicide is the third leading cause of death for

teenagers between the ages of 15 and 24. The Centers for Disease Control and Prevention reported that the suicide rate in 2007 for adolescents between the ages of 15 and 19 is 6.8 per 100,000. According to Roberts, a leading cause of death in juvenile correctional settings is suicide, with incarcerated juveniles having a risk for suicide approximately four times greater than juveniles in the general population (2006). As with incarcerated adults, confined juveniles have a high rate of mental illness compared to the general population. Hayes cited multiple studies which indicate prevalence rates of mental illness in the confined juvenile population to be well above 50 %, with many of the juveniles having multiple psychiatric diagnoses, including substance use disorders, conduct disorder, and history of childhood abuse (2009). Penn reported that 12.4 % of 289 adolescents recently admitted to a New England correctional facility had a previous history of attempting suicide (2003).

Characteristics of juveniles who have died by suicide while in confinement are similar to characteristics of adult victims of suicide in correctional settings. A review by Hayes of 79 juvenile suicides between 1995 and 1999 documented many characteristics similar to adult suicide victims in correctional settings, as well as some demographics unique to this population (2009). For example, juvenile suicide victims were 68 % white and 79.7 % male. Also similar to incarcerated adults, 73.4 % had a history of substance abuse, 74.3 % had a history of mental illness, and 69.6 % had history of prior suicidal behavior. The method of suicide, similar for adult suicide deaths, was hanging by 98.7 %. As with adult correctional facilities, the physical plant characteristics provided opportunities for suicide behavior with 50 % being in room confinement and 74.7 % housed in single-occupancy rooms. In addition, the data provided specifics for juvenile suicide victims, including a mean age of 15.7 with 70 % between ages of 15 and 17, 69.6 % confined for nonviolent offenses, and 78.5 % with history of previous offenses. Approximately 40 % of juvenile suicide victims confined in detention centers had died by suicide during the

first 72 h of confinement. Other confinement settings including training school/secure facility, reception/diagnostic center, and residential treatment center had more diverse time frames of juvenile suicide deaths. Hayes argued that although adolescent suicide has gained national attention, suicide of juveniles in confinement has gained much less consideration. He recommended more research into investigating the possible precipitating factors of juvenile suicide in confinement (Hayes 2009).

Women

The total percentage of incarcerated women in the United States has increased by 14 % from 1990 to 2009 (Glaze 2010). Steadman estimated that the rate of serious mental illness for female jail inmates is 31 % (2009). The Bureau of Justice Statistics reported that the rate of suicide for female inmates in both jails and prisons between 2000 and 2002 (32 per 100,000 and 10 per 100,000, respectively) is lower than that for males (Mumola 2005). Recent rates using data from 2000 to 2007 reflect a decrease in the rate of suicide deaths for female jail inmates (28 per 100,000) (Noonan 2010). However, the suicide rate still remains five times higher than the general population when these rates are demographically adjusted (Noonan 2010). Noonan also reports that incarcerated Hispanic females have a rate of suicide 10.5 times greater than the rate of suicide for Hispanic women in the general population (2010). Incarceration does appear to substantially increase the risk for suicide among women.

Research on incarcerated female suicides has been sparse. Clements-Nolle reported that one risk factor for suicide attempts in prison is childhood trauma (2009). Marzano found that in English and Welsh prisons, there was a higher prevalence of current depression, two or more psychiatric diagnoses, history of inpatient psychiatric hospitalization, and history of suicide attempt in women with near-lethal self-directed violence compared to a control group (2010). Fazel hypothesized that substance use disorder and loss of contact with dependent children may contribute to incarcerated female suicide risk (2009).

Veterans

Another special population is veterans. Estimates suggest that nearly 10 % of incarcerated persons are veterans, although such estimates were derived from data that probably fails to fully reflect the impact of current conflicts in the Middle East. Clinical experience among those familiar with the population of returning veterans suggests that such numbers may be on the rise (Mumola 2000; Noonan 2004). The Veterans Administration now officially recognizes that veterans are arrested for a variety of offenses, some of which may be related to extended periods of battle readiness and combat exposure during multiple deployments and to maladaptive coping with the return to civilian life. The VA's Uniform Mental Health Services Package now calls not only for assistance for veterans reentering the community from state and federal prisons but also for outreach efforts to veterans who are interfacing with jails, courts, and law enforcement and for education to these agencies regarding mental health problems relevant to veteran populations, such as posttraumatic stress disorder (PTSD) and traumatic brain injury (TBI). Unfortunately, relatively little is known about the subpopulation of incarcerated veterans, including their suicide rates. It has been argued that suicide risk factors faced by the general veteran population may interact with risk factors specific to incarceration or justice involvement, potentially placing the incarcerated veteran at especially high risk for suicide. Additional research is needed to better characterize this potentially high-risk population to determine rates and opportunities for intervention (Wortzel et al. 2009).

Risk After Release

The suicide risk associated with incarceration unfortunately does not end upon release. Rather, research suggests that the time of release represents yet another difficult transition period that carries increased risk for mortality in general and increased risk for suicide specifically. A study by Pratt examined suicide following prison incarceration in England and Wales (2006). Compared to

the rate of suicide of 18 per 100,000 in the general population, the rate for recently released inmates was 348 per 100,000, with 21 % of the suicides occurring within 28 days of release. Suicide risk was also found to be high for individuals after release from New South Wales, Australia, prisons (Kariminia et al. 2007). The suicide rate for individuals within 2 weeks of release from prison was found to be 3.87 times higher when compared with the suicide rate for individuals 6 months after release. Binswanger conducted a retrospective cohort study of inmates released from the Washington State Department of Corrections (DOC) and reported an adjusted risk of death among former inmates 3.5 times higher than that faced by other state residents (2007). More specifically, former inmates had a 3.4 relative risk for suicide, a 10.4 relative risk for homicide, and a 12.2 relative risk for accidental overdose. Meaningful efforts to mitigate suicide risk among recently incarcerated individuals will necessarily involve linkage to appropriate resources in the community.

Conclusions

The ongoing high risk of suicide death in correctional settings compared to the general population has resulted in more attention by the academic community, the courts, and correctional settings. This attention has led to a better understanding of the risks that exist in these settings and pragmatic solutions to decrease suicide risks. More recent studies have suggested that the rate of suicide in correctional settings has decreased over time. However, the decreased rates continue to remain higher than the general population rates, suggesting that efforts need to continue to find solutions to further reduce correctional suicide rates. The academic examination of suicide death and self-directed violence in correctional settings involving adolescents, women, and veterans is relatively new and is an area requiring further study. The high suicide risk/rate of recently released inmates from correctional settings is an area that is ripe for further study. Future studies could help provide

direction in formulating a plan to target and reduce the high suicide risk/rate in this subpopulation. Finally, this entry demonstrates the responsibility that correctional institutions have in further developing policies and procedures that aim to reduce the risks of inmate suicide.

Related Entries

- ▶ [Drug Abuse and Alcohol Dependence Among Inmates](#)
- ▶ [History of Substance Abuse Treatment](#)
- ▶ [Mental Health Courts](#)
- ▶ [Mortality After Release from Prison](#)
- ▶ [Surveys on Violence Against Women](#)
- ▶ [Treating Mentally Ill Offenders](#)

Recommended Reading and References

- Anasseri DE (2006) Preventing suicide in prison: a collaborative responsibility of administrative, custodial, and clinical staff. *J Am Acad Psychiatry Law* 34(2):165–175
- Anasseri DE, Fleming J (2006) Suicides in a state correctional system, 1992–2002: a review. *J Correct Health Care* 12(1):24–35
- Binswanger IA, Stern MF, Deyo RA, Heagerty PJ, Cheadle A, Elmore JG et al (2007) Release from prison – a high risk of death for former inmates. *N Engl J Med* 356(2):157–165. doi:10.1056/NEJMsa064115 [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]
- Bradley-Engen MS, Cuddeback GS, Gayman MD, Morrissey JP, Mancuso D (2010) Trends in state prison admission of offenders with serious mental illness. *Psychiatr Serv* 61(12):1263–1265. doi:10.1176/appi.ps.61.12.1263 [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]
- Clements-Nolle K, Wolden M, Bargmann-Losche J (2009) Childhood trauma and risk for past and future suicide attempts among women in prison. *Womens Health Issues* 19(3):185–192. doi:10.1016/j.whi.2009.02.002 [Research Support, Non-U.S. Gov't]
- Fazel S, Benning R (2009) Suicides in female prisoners in England and Wales, 1978–2004. *Br J Psychiatry* 194(2):183–184. doi:10.1192/bjp.bp.107.046490 [Research Support, Non-U.S. Gov't]
- Fazel S, Cartwright J, Norman-Nott A, Hawton K (2008) Suicide in prisoners: a systematic review of risk factors. *J Clin Psychiatry* 69(11):1721–1731 [Meta-Analysis Review]
- Felthous AR (2011) Suicide behind bars: trends, inconsistencies, and practical implications. *J Forensic Sci* 56(6):1541–1555. doi:10.1111/j.1556-4029.2011.01858
- Glaze LE (2010) Correctional population in the United States, 2009. U.S. Department of Justice Statistics
- Hayes LM (2009) Juvenile suicide in confinement – findings from the first national survey. *Suicide Life Threat Behav* 39(4):353–363. doi:10.1521/suli.2009.39.4.353 [Multicenter Study Research Support, U.S. Gov't, Non-P.H.S.]
- Hayes LM, National Institute of Corrections (U.S.), National Center on Institutions and Alternatives (U.S.) (2010) National study of jail suicide: 20 years later. U.S. Dept. of Justice, National Institute of Corrections, Washington, DC
- James DJ, Glaze LE, United States Bureau of Justice Statistics (2006) Mental health problems of prison and jail inmates. U.S. Dept. of Justice, Office of Justice Programs, Bureau of Justice Statistics, Washington, DC
- Joiner TE (2005) Why people die by suicide. Harvard University Press, Cambridge
- Kariminia A, Law MG, Butler TG, Levy MH, Corben SP, Kaldor JM et al (2007) Suicide risk among recently released prisoners in New South Wales, Australia. *Med J Aust* 187(7):387–390 [Research Support, Non-U.S. Gov't]
- Kerr DC, Washburn JJ, Feingold A, Kramer AC, Ivey AZ, King CA (2007) Sequelae of aggression in acutely suicidal adolescents. [Research Support, N.I.H., Extramural]
- Lamb HR, Weinberger LE (2005) The shift of psychiatric inpatient care from hospitals to jails and prisons. *J Am Acad Psychiatry Law* 33(4):529–534
- Leese M, Thomas S, Snow L (2006) An ecological study of factors associated with rates of self-inflicted death in prisons in England and Wales. *Int J Law Psychiatry* 29(5):355–360. doi:10.1016/j.ijlp.2005.10.004
- Liebling A (2006) The role of the prison environment in prison suicide and prisoner distress. In: Dear GE (ed) Preventing suicide and other self-harm in prison. Edith Cowan University, Australia
- Marzano L, Fazel S, Rivlin A, Hawton K (2010) Psychiatric disorders in women prisoners who have engaged in near-lethal self-harm: case-control study. *Br J Psychiatry* 197(3):219–226. doi:10.1192/bjp.bp.109.075424 [Research Support, Non-U.S. Gov't]
- Mumola C (2000) Bureau of Justice Statistics Special Report, veterans in prison or jail. (178888). Retrieved from <http://www.ojp.usdoj.gov/bjs/pub/pdf/vpj.pdf>
- Mumola CJ, United States Bureau of Justice Statistics (2005) Suicide and homicide in state prisons and local jails. U.S. Dept. of Justice, Office of Justice Programs, Washington, DC
- Noonan M (2004). Veterans in state and federal prison. (<http://www.ojp.usdoj.gov/bjs/abstract/vsfp04.htm>)

- Noonan M, United States Bureau of Justice Statistics (2010) Mortality in local jails, 2000–2007. U.S. Dept. of Justice, Office of Justice Programs, Bureau of Justice Statistics, Washington, DC
- Patterson R, Hughes F (2008) Review of completed suicide in the California department of corrections and rehabilitation, 1994–2004. *Psychiatr Serv* 59(6):676–682
- Penn JV, Esposito CL, Schaeffer LE, Fritz GK, Spirito A (2003) Suicide attempts and self-mutilative behavior in a juvenile correctional facility. *J Am Acad Child Adolesc Psychiatry* 42(7):762–769. doi:10.1097/01.CHI.0000046869.56865.46 [Research Support, Non-U.S. Gov't]
- Polluck LR, Williams MG (2006) Psychological perspectives on intentional self-harm. In: Dear GE (ed) *Preventing suicide and other self-harm in prison*. Edith Cowan University, Australia
- Pratt D, Piper M, Appleby L, Webb R, Shaw J (2006) Suicide in recently released prisoners: a population-based cohort study. *Lancet* 368(9530):119–123. doi:10.1016/S0140-6736(06)69002-8 [Research Support, Non-U.S. Gov't]
- Roberts AR, Bender K (2006) Juvenile offender suicide: prevalence, risk factors, assessment, and crisis intervention protocols. *Int J Emerg Ment Health* 8(4):255–265
- Scott CL (2010) Legal issues regarding the provision of care in a correctional setting. In: Scott CL (ed) *Handbook of correctional mental health* (2nd edn). American Psychiatric Publishing, Inc., Washington, DC, London
- Steadman HJ, Osher FC, Robbins PC, Case B, Samuels S (2009) Prevalence of serious mental illness among jail inmates. *Psychiatr Serv* 60(6):761–765. doi:10.1176/appi.ps.60.6.761 [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]
- Way BB, Miraglia R, Sawyer DA, Beer R, Eddy J (2005) Factors related to suicide in New York state prisons. *Int J Law Psychiatry* 28(3):207–221. doi:10.1016/j.ijlp.2004.09.003
- White TW, Schimmel DJ, Frickey R (2002) A comprehensive analysis of suicide in federal prisons: a fifteen-year review. *J Correct Health Care* 9(3):321–345
- Wolff P, Nancy L, Shi J, MS (2010) Trauma and incarcerated persons. In: Scott CL (ed) *Handbook of correctional mental health* (2nd edn). American Psychiatric Publishing, Inc., Washington, DC, London
- Wortzel HS, Binswanger IA, Anderson CA, Adler LE (2009) Suicide among incarcerated veterans. *J Am Acad Psychiatry Law* 37(1):82–91 [Comment Review]

Surveillance Technology and Policing

Gavin J.D. Smith

Senior Lecturer in Sociology, The Australian National University, Canberra, ACT, Australia

Overview

Policing is in a state of transformation. Traditional paradigms founded on reactive and responsive principles are being superseded by contemporary risk-based models, proactive and preemptive approaches that encompass an array of stakeholders and sophisticated technologies. The “new policing” is underpinned and driven by intelligence-led doctrines, performance indicator metrics, and desires to better identify crime patterns. Although many relational factors account for these changes in policing ideology and praxis, the integration of – and reliance on – surveillance technologies and associated “visibility” rationalities has been particularly pivotal. Yet the widespread availability of surveillance technologies is simultaneously diversifying policing practices and responsibilities. Civil/commercial organizations and citizen/consumers are, for example, increasingly participating in the inspection and assessment of behavior, albeit for often differing purposes. The perpetual interweaving of visibility-inducing systems into the cultural fabric of everyday living and concomitant heterogeneity of actors who now engage in regulatory activities has generated two major effects: (a) policing opportunities have increased; and (b) supervisory webs have become progressively intricate and multi-dimensional in semblance. “Policing” is now a pluralized social practice that extends far beyond “the police.” Indeed, the “policing creep” emerging from intensifying visibilities generates for criminologists important theoretical, empirical and ethical questions, some of which inform the substantive focus and thematic arrangement of this entry.

In order to arrive at an understanding of the surveillance technology-policing nexus, this entry commences with an overview of the

Supervision

► Desistance and Supervision

emerging techniques of visibility now infusing the social orders of privileged global provinces. This will enable us to better grasp the assortment of observational devices being routinely deployed, the heterogeneous nature of operational procedures, and the differing populations now engaged in the policing of subjects. It will also permit us to comprehend some emergent social effects. A case study analyzing closed-circuit television (CCTV hereafter) camera system growth in the global north will then show empirically how surveillance technologies can modify and transform existing policing structures. This section draws further attention to the evolving labor force now performing, in this case, paid policing duties.

This entry, therefore, is not a story about “surveillance and the police” per se. Rather, it is an account describing how the widespread distribution of surveillance technologies ruptures previous modes of social control and instigates novel policing configurations. Criminologists are only beginning to comprehend the social implications attendant on these changes.

Surveillance Technologies and Rationalities

The rapid and extensive diffusion of surveillance technologies into the architecture of contemporary living has, in recent decades, been a particularly salient development in social governance. Everyday actions are targeted by an expanding arsenal of monitoring devices. Indeed, it has become well-nigh impossible to complete commonplace tasks in privileged wordly regions without encountering a diverse set of data capturing “probes.” CCTV camera systems, airport body scanners, DNA swabs, X-ray scanners, and Web browsing cookies, for example, perpetually extract personal information “particles” from the data trails that accompany our daily activities. Tracking and transmission technologies such as radio-frequency identification (RFID) microchips and Global Positioning Systems (GPS) are now embedded in all

manner of mundane objects: credit cards, cell phones, and smart passports being obvious illustrations.

Despite embodying differences in terms of technical components and physical dimensions, there are general similarities between and among surveillance technologies on an operational front. Most function to procure, transport, and then ultimately deposit data emissions derived from embodied acts – that is, from walking down a street, making an online purchase, and sending a SMS text message – into a complex cybernetic matrix commonly known as “the database.” These vast informatic chambers decontextualize behaviors both by systematically removing them from their circumstantial origin and by rendering behavioral expressions into a series of discrete data flows. A person’s consumption profile, for instance, can be inferred from analyzing her or his credit card statements. Buying habits, in this example, are deemed to reveal social status and thus customer “worth.” Rather than take into consideration the infinite causal variables influencing the decision to purchase a product, these systems focus attention instead upon tangible outcomes and seek to aggregate data from multiple sources so that “tendencies” can be ascertained. Similar to how social research data is arranged, personal information is recurrently coded, collated, mined, or “recontextualized” by a set of surveillance workers – be they human (e.g., a CCTV operator) or nonhuman (e.g., a computer code) – in accordance with an assortment of organizational imperatives. This allows for historical activities to be logged, social relations to be mapped, actions to be compared, and causal patterns to be identified. Based on such information, risk/return evaluations of specific phenomena can be undertaken and decisions reached, for example, whether or not to invest capital in a specific enterprise, whether or not to trust a job applicant, and whether or not to proffer credit to a prospective customer. Direct human mediation in data analysis and processing is declining as systems of surveillance develop in capacity and sophistication. Computerized software is increasingly favored to autonomously dispose the information collected.

Surveillance technologies are a rationalized means to enhance visibility, to sharpen decision-making, and to more effectively manage resources, helping their overseers determine risk propensities, investment potentials, trustworthiness, and suspiciousness. They expose and scrutinize discrete social processes (such as disease intensity, voting patterns, or market relations) and help identify probable correlations between and among variables. They permit interest groups to accumulate intimate knowledge about specific phenomena (e.g., population health in specific regions) and to fashion corresponding interventions from an informed perspective (e.g., the proposed location of a health clinic). They provide a calculative lens through which all manner of truth claims can be assessed and adjudicated (e.g., an asylum seeker application or a credit request). As such, surveillance operates at the very heart of organizational protocol, specifically as it (a) enables social complexity in all its manifestations to be captured and (b) facilitates complexity governance and uncertainty reduction. Indeed the visibility – that is, the capacity to see and be seen – produced by systems of surveillance is fundamentally transforming the conditions of contemporary social life and how we relate to one another. Desires to know, order, govern, and entertain are exposing people and processes in much finer detail, and this situation is both diversifying and intensifying formal and informal policing arrangements.

Surveillance technologies reveal previously concealed facets of life and circulate mediated personal intimacies to an audience that increasingly is pluralized, remote and anonymous. An example of this process can be discerned from examining Web 2.0 engagement. Social media sites like Facebook proactively encourage users both to self-disclose/self-promote and to peruse the exteriorized musings (“newsfeeds”) of fellow participants. User curiosity and voluntarism is exploited and manipulated in strategic ways, system programmers systematically designing and assembling an array of infrastructural mechanisms to sustain interest levels – and thus asset sustainability and value. Incentivizing architectures ensure that

acts of revelation are celebrated and rewarded and that tracking the lives of others is experienced as an addictive and pleasurable pastime. The pellucid nature of these online registries permits emotive posts and graphic images to be accessed and deciphered by an army of unseen inspectors (e.g., peers, employers, marketers, and law enforcement agencies), the biographical information deposited revealing vital clues about a particular person’s ideological values, consumer preferences, and psychological disposition. This knowledge can then be utilized by the various spectators to construct character profiles and to rationalize interventions, specifically procedures aimed at modulating and engineering conduct in some desired way (Trottier 2012).

We can perceive from this illustration the contemporary heterogeneity of policing, in terms of contributors and purposes. The operational reach and remit of policing continues to broaden as surveillance technologies extend purviews deeper into the backstage regions of personal life. Indeed, it is helpful to conceive people as “cyborg” entities, comprised of physical biomatter and informatic particles: the latter material being willingly emitted – and coercively extracted – from our embodied actions in the course of everyday living. The point to note is that personal data has become a highly valued unit of currency for problem analysis and decision-making, its textual content shaping significantly how particular individuals and social groups are perceived and subsequently treated by differing organizational systems. Rather than trusting a person’s narrative or status claim at face value, interlocutors increasingly prefer to discern legitimacy and authenticity from the information encased within a set of corresponding archival records or what I term, “surveillance texts.” Although often derived from certified or valid documents – for example, medical consultation notes, criminal records, bank statements, and tax returns – employers, law enforcers, and insurance actuaries also obtain knowledge fragments from unofficial and aspatial communicational mediums like Twitter and Facebook, and from personal email correspondence. Each source comprises discrete “event descriptors” that,

when collated, depict previous conduct (e.g., consumption preferences) and circumstances (e.g., health conditions). It is believed that analysis of multiple surveillance texts will reveal (a) a detailed picture of a person's disposition and (b) a more accurate representation of her or his social positioning and future trajectory. This permits the adjudicating authority to effectively bypass or short-circuit what is perceived as the contrived nature of self-presentations, unveiling instead preconscious forms of cognition (Andrejevic 2010). In a world increasingly defined by information glut and uncertainty (a situation ironically accentuated by the intensification of visibility), surveillance systems act as critical arbitration tools, a means for truth establishment and narrative verification, a means for individuating and stratifying populations, and a means for regulating access to resources.

Surveillance systems mostly operate in hegemonic ways, functioning to maintain the interests of a privileged elite and thus a particular fiscal status quo. Yet the gradual diffusion and sedimentation of surveillance should not be understood as some ideological conspiracy that has been engineered by a state-corporate cartel. In contrast, a surveillance system or device intrinsically is neither benign nor malign, the sociocultural relations in which the apparatus is entangled effectively shaping its everyday orientation (see Smith 2013). Indeed, state authorities and administrators depend upon an assortment of surveillance machineries and practices to ensure the deliverance of democratic commitments and ideals: that voting rights are upheld, that law and order is maintained, that taxation thresholds are accurately set, and that designated populations are granted access to welfare services and benefits. Surveillance technologies, for example, are regularly utilized to assess and manage a nation's health. Disease epidemiologists use biostatistics to identify and gauge immanent biosecurity threats, while medical practitioners use hi-tech screening equipment to better envision patients' bodies so that malignant cells can be located and tissue injuries reliably evaluated. Law enforcers employ a range of surveillance techniques to investigate

suspects, to acquire evidence and to direct operations. Moreover, surveillance is exploited strategically as both a thematic currency and a textual commodity by an assortment of cultural producers, be they formal commercial organizations or civic communities. People are regularly exposed to surveillance symbolism and output in the course of consuming a range of media broadcasts. Surveillance texts often depict graphic content the witnessing of which can be both shocking and entertaining. As a distinctive subject-matter, surveillance is artfully appropriated by groups possessing a diversity of political agendas, artistic aesthetics and economic motivations. The ensuing struggle over surveillance representations in popular culture encourages audiences to approach surveillance technologies with an attitudinal cynicism and ambivalence stimuli and representations. Surveillance is also deployed to increase the speed and simplicity of transactions between individuals and organizations and to enhance the flexibility, convenience, and efficiency of these encounters. In addition, surveillance is fast becoming a medium through which powerful groups can be made more accountable for their actions. Footage captured on a G-20 protestor's personal video camera, for instance, revealed a civilian being unlawfully assaulted by a police officer (see <http://www.guardian.co.uk/uk/2009/apr/07/ian-tomlinson-g20-death-video>, accessed 18/5/2012).

Despite notable exceptions – where surveillance technologies are harnessed to serve the interests of vulnerable individuals and periphery collectives – data-hungry social and commercial institutions are the overwhelming winners of surveillance ubiquity. This is because capturing and analysis technologies afford these agencies statistical knowledge on social trends and attitudes, allowing evidence-based responses – in terms of policy or commodity design – to be instrumentally engineered and actualized. Surveillance infrastructures and circuitries greatly abet elite groups, to the extent that they help to legitimize their domination, accentuate their expertise, foster their definitional capacities, and protect their capital, interests and resources.

Indeed, a core operational objective is the differential classification and division of groups in accordance with their perceived position in the social hierarchy, their projected risk level, and their assigned market value (Gandy 1993; 2009). Thus, in an economic order defined by mass consumption and in a political order characterized by omnipresent security concerns, population distrust, and risk-sensitive policies, it is people residing in disadvantaged locales without the means to consume, or those whose ethnicity, piousness, skin color, and occupation contravene a dominant (normative) security discourse, who are recurrently subject to disproportionate levels of scrutiny (Norris and Armstrong 1999; Lyon 2003). Procurement of a stigmatizing or discrediting label can mean that an individual or collective is denied access to certain spaces and/or refused basic rights, resources and privileges (Monahan 2008; Zureik et al. 2011). Thus, because surveillance systems embody the prejudicial essentialism of their controllers, they tend to produce exclusionary outcomes and perpetrate social injustices.

The Diversification of Policing Responsibility

A major outcome of neoliberal policy has been an aspiration by elite groups to instill a “responsibility ethic” and “cosmopolitan sensibility” within the psyches of the population: an expectation that citizens play an active role in managing risk. This expectation primarily relates to personal health and safety practices but also to deportment in the economic and ecological spheres. In the law and order field, adoption of a “responsibilization” rationality (see O’Malley 1996) by state authorities has profoundly transformed the meaning of “criminality” and “citizenship.” What historically had been a matter for formalized policing agencies has now become a community issue demanding multiagency interventions; policymakers eager to enlist in quotidian supervision practices both civil society actors and the private sector. Criminality itself has been discursively reconstituted as a set of hierarchical

thresholds, ranging from lower level civil transgressions to higher order terrorist/biosecurity risks. This stratification – and concomitant governmental desire to mobilize as “crime managers” various groups – has produced a distinctive set of stakeholders (e.g., citizens, proprietors, and council workers) who are now equipped with specialized powers – and ascribed specific obligations – for addressing particular expressions of disobedience (Rose 2000; Garland 2001).

The emergence of civil “on-the-spot” penalty notifications, laws persecuting parents for child truancy, community “wardens,” neighborhood watch programs, and dispersal/antisocial behavior orders are notable empirical instances of a general trend to prescribe codes of etiquette and to assign any violation – or manifestation of deviance – with a particular regulatory custodian (e.g., a parent, a schoolteacher, a social worker, or a transport marshal). Indeed, the diversification of actors now responsible for everyday surveillance and social ordering marks a distinct paradigm transition in policing ideology and application, from a traditional model which was largely “reactive and localized/centralized,” that is, post facto and exclusive, to one which is now “proactive and dispersed/pluralized,” that is, anticipatory and inclusive (McLaughlin 2007).

Such changes in policing praxis correlate with the neoliberal state’s remit to buttress declining security budgets/provision and to “govern at a distance” (Miller and Rose 1990) with a corresponding engagement (or exploitation) of community groups and corporate agencies in the collective “fight against crime” (Garland 2001). They also accord with an allied desire to sponsor the implementation of hi-tech surveillance infrastructures which operate according to intelligence-led axioms and preemptive principles (see Elmer and Opel 2008). The Texas Virtual Border Watch program is an apposite illustration (<http://blueservo.com/vcw.php>, accessed 30/5/2012). This online initiative effectively transforms citizens into custodial guardians of the United States–Mexico border. It encourages participants, through financial incentives, to identify and report suspicious mobilities and to adopt a nationalistic sensibility: a cognitive disposition

that is hostile to the context prompting a decision to enter illegally a foreign sovereign territory. An uncompassionate and uncompromising mentality orientated to informing and supervising is, in this process, deemed to be a desirable civil ethic (Koskela 2011). These policy transformations (e.g., enlisting communities in addressing lower order criminality) simultaneously grant formal policing organizations special license to concentrate finite resources on tackling serious and organized crime and an assortment of high-level national security threats. Moreover, it is in this transitional climate that policing authorities have increasingly adopted surveillance technologies as ways of managing shrinking quotas and the complexities associated with transnational law enforcement in the information age.

The rolling out of large-scale socio-technical assemblages such as DNA databanks and CCTV camera networks demonstrates how responsibilities for managing crime and disorder are transferring from a dependence on theological precepts, authoritarian physicality, and human informants to a faith in scientific expertise and technical solutions. The fetishization or overprivileging of scientific discourse and technical systems by governmental actors, while partly operational, is in large part symbolic and in response to a perceived “crisis” in traditional policing approaches, brought on by austerity measures, rising crime fears, and high profile institutional failings/operational incompetence (McLaughlin 2008). Utilizing sophisticated technological infrastructures and computer-powered calculative metrics permits policing agencies to identify criminogenic patterns and to anticipate prospective security threats (Ericson and Haggerty 1997). It is now possible to collate previously discrete data fragments and create “criminality constellations.” The Metropolitan Police’s crime map (<http://maps.met.police.uk/>, accessed 20/5/2012) is an apt example. Based on the quantitative aggregation of crime reports and statistics, policing authorities are able to spatially and temporally envision felony rates and correlations within the urban metropolis. This ensures that any proposed crime prevention initiative or awareness-raising campaign has an

empirical foundation and is measurable. It also means that resources can be rationally deployed in strategic and targeted ways and decision-making protocol as it relates to criminal investigations can be externally validated and legitimized. It additionally permits bureaucratic officials and police managers to set performance targets and impose evaluative frameworks.

Evidently, these applications have multiple users and serve numerous purposes. Insurance companies, for example, are keen to exploit crime data to calibrate premiums, privileged citizens to distinguish (un)desirable locations for settlement, and marketers to direct advertising toward amenable communities (and away from others). The employment of actuarial techniques by formalized policing entities means that, based on available indicators, persons and groups acquire a risk rating (Feeley and Simon 1994). This computer-generated calculation effectively comes to determine the type and intensity of scrutiny that is then actioned. Gradual erosion of the “innocent until proven guilty” legal criterion has been an outcome of these developments and wider anti-social behavior and anti-terror policies. Increasingly, policing logic has more unity with a “guilty until proven innocent” aphorism, as shown in recent UK legislation endowing extraordinary preventive and regulatory powers to police forces. These laws grant police officers special license to “randomly” stop and search individuals, to disperse groups of young people, to preemptively detain terror “suspects” without charge, and to demand that protestors register their personal details before attending demonstrations. They also permit law enforcers to acquire sensitive personal data derived from the activities of law-abiding persons (<http://www.bbc.co.uk/news/technology-18102793>, accessed 20/5/2012) and to ascribe “suspect status” to an individual on the basis of her or his communicative position within a criminalized informatic network (<http://news.bbc.co.uk/2/hi/technology/8245056.stm>, accessed 24/5/2012). Similarly, landmark statutes such as the USA Patriot Act (2001) bestow exceptional powers to national security agencies in terms of intelligence

gathering and information sharing protocol, and facilitate the introduction of ever more intrusive surveillance measures. These bills tend to be passed with limited public consultation or debate in moments of intense emotionality, for example, in the immediate aftermath of a terrorist act or in advance of mega events like the Olympic Games or a G20 convention, when anger and fear – rather than critical rationalism and reflexivity – orientate the collective consciousness and the political will.

Surveillance Technologies and Surveillance Work

The previous sections described how and why governmental desires and reality-capturing devices have extended policing practices into the circuitries of everyday living. CCTV is an insightful example of how policing praxis has co-evolved with technological supplementation, specifically in terms of the emerging complexity of social relations and operational remits that now constitute such regulatory systems. There are additional reasons for considering CCTV. Focusing on the exponential growth of camera networks in the global north illuminates the complex politics and rationalities underpinning surveillance diffusion. Moreover, considering novel policing modalities such as CCTV focuses attention on the unique type of labor – or “surveillance work” – that the camera operators routinely practice. These “order custodians” are subsidized to remotely encounter, capture, and “manage” mediated scenes of criminality and disorder, transgressive spectacles to which such workers are repeatedly exposed but over which they have no direct control. Reflecting on how the introduction of CCTV cameras transforms policing practices in both intentional (official) and unintentional (informal) ways provides clues for envisaging the likely impact of additional surveillance technologies on conventional policing protocol.

CCTV provision has dramatically increased in recent years. This is particularly true of the UK, a world leader in terms of its camera

coverage. While a complex politics, at local, regional, and national levels, surrounds the augmentation process (see Fussey 2007; Norris and Armstrong 1999), the 1980s economic recession was a decisive factor in the state’s decision to legislatively promote and financially sponsor a series of installation programs. Rising crime rates associated with growing unemployment, a declining welfare system and resource inequities, and a corresponding amplification of the issue by media reportage impelled a public perception that policing agencies were failing in their bid to control criminality. Successfully redressing this situation became a central political issue, Home Office officials desperate to introduce a progressive “silver bullet” solution that would enhance the policing service and “treat” in an effective manner the physical symptoms of structural disparities (note similarities between this scenario and the one depicted in science fiction movie, *Robocop* (1987)). The intelligence battles of the Cold War period meant that an array of advanced militarized surveillance technologies were available for civil market appropriation. CCTV was one such innovation (Norris and Armstrong 1999).

Preliminary evaluations of pilot projects in several jurisdictions showed that public space camera networks could reduce particular acts of disobedience by around 40%. Although this research was methodologically problematic and predominantly administered by self-interested practitioners, the extraordinary statistics persuaded ruling technocrats that CCTV camera networks were worthy of significant financial investment (Gill and Spriggs 2005). Cities were encouraged to compete with one another for allocated monies totalling millions of pounds, and camera systems rapidly became akin to a “fifth utility” service (Graham 2001).

The favouring of CCTV systems by urban authorities was also the outcome of neoliberal “urban renaissance” policies which advocated restructuring of the metropolis around commercial imperatives and resiliency ideals (Atkinson and Helms 2007; Coaffee et al. 2009). This principally involved spatial revitalization, market liberalization, and the assembly of vibrant,

leisure economies. Manipulating consumer perception became a central directive of the project, and CCTV cameras were deployed to help promote an impression of safety that would facilitate increased consumption and a “feel good factor”.

A further factor fostering CCTV’s wider cultural appeal was the global circulation of images showing Merseyside toddler, Jamie Bulger, being abducted by two ten-year-old boys in a Bootle shopping mall (Norris and Armstrong 1999). Although the cameras did nothing to prevent the child’s violent death, the moral outrage and anxiety associated with this act as symbolized by the grainy pictures was, for different reasons, exploited by an array of state, media, and industry actors. Notwithstanding an assortment of competing rationales and rationalities, each agency had a vested interest in sponsoring CCTV deployment. The government had an interest in endorsing system efficacy to accrue public support and thus political capital/leverage. The police had an interest in supporting CCTV implementation as cameras could bolster conviction rates, sharpen operational activities and enhance investigative practices. The media had an interest in acquiring “authentic” camera footage to both disturb and amuse a scopophilic audience, that is, an audience obsessed with the spectacle. The security industry had an interest in promoting CCTV’s effectiveness to generate increased demand for the technologies and for training programs.

The symbolism embodied in these initiatives is worth reflecting upon. The cameras act as materialized signifiers of sovereign authority and symbolize a profound distrust of the public by ruling groups. They cultivate conformity to the consumer order and, in the process, reify dominant definitions of appropriate urban conduct. They become the instruments through which discrediting labels can be ascribed to and enacted upon particular stigmatized groups (Norris and Armstrong 1999). They operate as architectural barriers for rationally thinking criminals to circumnavigate, obstacles hindering the ease with which an act of criminality can be perpetrated in a particular locale. Camera

networks are aesthetic additions to the urban landscape and represent a city’s prosperity and progressiveness (Graham 2001). They become part of the political discourse on risk and security, and function as discursive horizons on which an array of propagandistic vistas can be fashioned. Such partisan conceptions tend to overemphasize the technology’s disciplinary influence and to fetishize its role in crime reduction and public safety. CCTV camera initiatives are also a tokenistic reminder, at least in the public imaginary, that the state is actively engaged in addressing both the fear and actuality of criminality.

The above analysis demonstrates that surveillance technologies are much more than mere technical-material structures. It shows, in contrast, that they are also social enterprises. Although surveillance assemblages emerge as a result of political economic rationalities and comprise technical components, their everyday operational vitality and verve is entirely dependent on a complex chain of social actors and relations. Considering the role of workers in the daily production of surveillance is, therefore, an important part of the story. This focus permits us to discern the exploitative dimensions of policing reality, the artful practices comprising the operational lifeworld and the changing configuration of policing praxis.

Surveillance workers tend to occupy a privileged symbolic position, wielding discretionary interpretive powers and an anonymous and asymmetrical gaze. Recent research on the topic of surveillance work(ers) has illustrated (a) the extent to which such laborers impose their own definitional criteria upon that which is observed and (b) the consequences for those who are involuntarily exposed. Norris and Armstrong’s (1999) ethnographic study of several CCTV control rooms in the UK is an illustrative example. The two scholars sought to establish the interpretive frameworks or “working rules” – effectively, subjective schemata – operators drew upon to determine camera positioning. The researchers show that suspicion is not an innate or discernible behavioral quality but is rather the outcome of social construction

processes. It was racist, sexist, fascist, and classist ideologies that largely determined where and at whom cameras were pointed, the operators predominantly associating criminality with young males, minority groups, and underprivileged populations, specifically those not explicitly engaged in legitimized forms of consumption. These findings emphasize the categorical or labeling power surveillance workers exercise both in the arrangement of social reality and in concomitant decision-making protocol.

Another way of comprehending the social dimensions of surveillance work has been developed by Smith (see 2007, 2009, 2013). He found camera operators to be active consumers of social reality, attaching anecdotal stories to the mediated events encountered. The surveillance workers' musings comprised a mix of imagination and memory, and this creative interchange between the fictional and the real helped foster a set of ritualized occupational customs that made the trying work of watching a more pedagogically meaningful and pleasurable activity. It is easy to imagine and perceive "the gaze" projected by surveillance workers as inevitably embodying disciplinary properties and to assume watching practices as invariably empowering, rewarding and seductive. Yet surveillance work is ambiguous. It comprises important affective and exploitative components. Camera operators, for example, are required to search for and identify events threatening a desired normative order and, as such, recurrently confront graphic scenes depicting social disorder and human suffering. They must dispassionately film people attempting suicide, injecting drugs, committing robbery, and being viciously assaulted. As a result of the labor that they practice and the traumatic spectacles to which they are habitually exposed, camera operators come to perceive the social world around them as chaotic, precarious, and dangerous. Being routinely subjected to mediated scenes of distress distills in such laborers a distinctive set of pathological beliefs about human conduct and relations, such values reorientating their everyday engagements with the world (Smith 2013).

A sequence of coping strategies are drawn upon by camera operators to manage their general estrangement both from the commodity that they are trying to produce (i.e., social order and its inherent fragility) and from the manufacturing process more generally (i.e., system technologies and occupational conditions). Cameras, for example, are used informally by operators in an attempt to escape, albeit temporarily, from the stresses and tensions associated with witnessing behavioral transgressions and social disarray. Nature becomes an attractive focal point. Its structural stability and intrinsic grace, as represented in dew-covered spider webs, cumulus clouds, and majestic sunrises, helps operators reestablish a sense of trust and coherency in a world that so often appears fractured and anarchic. Thus, the same technologies fuelling the operators' general malaise with, and alienation from, social reality, also facilitate transcendental pursuits, in this case, identification of the natural order's aesthetical mechanics, internal consistency and harmonic patterns. Camera operators also partake in individualized and collective forms of emotional management. The application of black humor, bullying, and sensory detachment – and exploiting camera optics for voyeuristic ends, for instance – allows potentially pathological (and health degenerative) feelings to be channeled and mitigated (and to be projected onto external others).

Surveillance workers are situated within and stratified according to wider occupational hierarchies. They are increasingly embroiled in struggles over camera positioning/control and continually receive quasi indictments requiring them to publicly defend their productions in law courts (Smith 2009). Moreover, research by Gould (2004) and by Newburn and Hayman (2002) has shown how the introduction of CCTV into the public sphere and the custody suite has increased the visibility of policing practices and placed novel accountability pressures on law enforcement officers – and generated corresponding occupational dilemmas for camera operators. Despite the fact that CCTV networks are supposed to support the

police in their pursuit of occupational goals, it is often the case that such systems challenge traditional structures in the police's occupational culture, particularly their monopoly on the legitimate use of violence. The routine exposure of policing practices has been further facilitated by the mass diffusion of portable capturing devices and by global circulations of information, especially as facilitated by Internet platforms such as Facebook and YouTube.

Future Directions

This entry has considered how emerging cultures of visibility transform the concept and practice of policing. In particular, the proliferation of surveillance devices into organizational circuitry and cultural canals has been shown to diversify the nature and meaning of policing. Institutions are becoming as much "the watched" as "the watchers." The same can be said of individuals. People embrace surveillance platforms to derive knowledge about social phenomena, to speed up their worldly transactions, and to be entertained. As such, persons increasingly find themselves experiencing surveillance regimes as both exposed "subjects" and contributory "agents." It would seem that, for many, scopophilia has become a seductive practice and voluntary disclosure a desirable state of being.

A plurality of conditions account for these transformations, and more research need examine the political economies, rationalities, and imaginaries underpinning the design and construction of surveillance technologies and attendant product marketing. Research on how different social groups understand, negotiate, and partake in surveillance activities is also required. Further criminological research, therefore, must excavate the nuances associated with surveillance "work," specifically how the introduction of computerized algorithms to autonomously determine patterns of causality in social reality – and to dispose specific interventions – demands a critical rethinking of how contemporary policing is increasingly

performed. Today's world is defined by the visibility of particular social flows and the corresponding invisibility of general surveillance processes. The latent nature of computer processing and the overvaluing of technology means that power and domination are becoming harder processes to empirically locate, discern, follow, and – importantly – hold to account. If criminology is to remain analytically relevant to the rapidly evolving structures of social regulation, it must innovate methodologically and epistemologically, particularly appropriating knowledge and tools from quantum physics, mathematics and computing sciences. Policing now operates through informatic circuits, and it is to a better understanding of these cybernetic territories – as distinctive "fields of power" – that criminologists increasingly need focus their energies.

Related Entries

- ▶ [Biometrics and Border Control Policing](#)
- ▶ [CCTV and Crime Prevention](#)
- ▶ [Crime Mapping](#)
- ▶ [Crime Science](#)
- ▶ [Criminal Profiling](#)
- ▶ [DNA Profiling](#)
- ▶ [DNA Technology and Police Investigations](#)
- ▶ [Electronic Monitoring](#)
- ▶ [Fingerprint Identification](#)
- ▶ [Forensic Science and Criminal Investigation](#)
- ▶ [History of Technology in Policing](#)
- ▶ [Identification Technologies in Policing and Proof](#)
- ▶ [Information Technology and Police Work](#)
- ▶ [Innovation and Crime Prevention](#)
- ▶ [Intelligence-Led Policing](#)
- ▶ [Prediction and Crime Clusters](#)
- ▶ [Predictive Policing](#)
- ▶ [Risk Assessment, Classification, and Prediction](#)
- ▶ [Telemetric Policing](#)
- ▶ [Use of Social Media in Policing](#)
- ▶ [Video Technology and Police Interrogation](#)
- ▶ [Visualizing Data: A Brief History](#)

Recommended Reading and References

- Andrejevic M (2010) Thin-sliced thoughts and theory's ends. *Media Tropes* 2(2):45–64
- Atkinson R, Helms G (eds) (2007) *Securing an urban renaissance: Crime, community and British urban policy*. The Policy Press, Bristol
- Coaffee J, Murakami-Wood D, Rogers P (2009) *The everyday resilience of the city: how cities respond to terrorism and disaster*. Palgrave Macmillan, London
- Elmer G, Opel A (2008) *Preempting dissent: the politics of an inevitable future*. Arbeiter Ring Press, Winnipeg
- Ericson RV, Haggerty KD (1997) *Policing the risk society*. University of Toronto Press, Toronto
- Feeley M, Simon J (1994) Actuarial justice: power/knowledge in contemporary criminal justice. In: Nelkin D (ed) *The future of criminology*. Sage, London
- Fussey P (2007) An interrupted transmission? Processes of CCTV implementation and the impact of human agency. *Surveill Soc* 4(3):229–256
- Gandy O (1993) *The Panoptic sort: A political economy of personal information*. Westview, Boulder
- Gandy O (2009) *Coming to terms with chance: Engaging rational discrimination and cumulative disadvantage*. Ashgate, Farnham
- Garland D (2001) *The culture of control: Crime and social order in contemporary society*. Oxford University Press, Oxford
- Gill, M, and A. Spriggs (2005) 'Assessing the impact of CCTV', Home Office Research Study 292, London: Home Office Research, Development and Statistics Directorate
- Goold B (2004) *CCTV and policing: public area surveillance and police practices in Britain*. Oxford University Press, Oxford
- Graham S (2001) CCTV: the stealthy emergence of a fifth utility? *Plan Theory Prac* 3(2):237–41
- Koskela H (2011) Hijackers and humble servants: individuals as camwitnesses in contemporary controlwork. *Theor Criminol* 15(3):269–282
- Lyon D (ed) (2003) *Surveillance as social sorting: privacy, risk and digital discrimination*. Routledge, London
- McLaughlin E (2007) *The new policing*. Sage, London
- McLaughlin E (2008) Hitting the panic button: policing/'mugging'/media/crisis. *Crime, Media, Cult* 4(1):145–154
- Miller P, Rose N (1990) Governing economic life. *Econ Soc* 19(1):1–31
- Monahan T (2008) Editorial: surveillance and inequality. *Surveill Soc* 5(3):217–226
- Newburn T, Hayman S (2002) *Policing, surveillance and social control: CCTV and police monitoring of suspects*. Willan Publishing, Cullompton
- Norris C, Armstrong G (1999) *The maximum surveillance society: the rise of CCTV*. Berg, Oxford
- O'Malley P (1996) Risk and responsibility. In: Barry A, Osborne T, Rose N (eds) *Foucault and political reason*. UCL Press, London
- Rose N (2000) Government and control. *Br J Criminol* 40(2):321–339
- Smith GJD (2007) Exploring relations between watchers and watched in control(led) systems: strategies and tactics. *Surveill Soc* 4(4):280–313
- Smith GJD (2009) Empowered watchers or disempowered workers? The ambiguities of power within technologies of security. In: Aas KF, Gundhus HO, Lomell HM (eds) *Technologies of insecurity: the surveillance of everyday life*. Routledge–Cavendish, London
- Smith GJD (2013) *Opening the black box: surveillance in everyday life*. Routledge, London
- Trottier D (2012) *Social media as surveillance: rethinking visibility in a converging world*. Ashgate Publishing, Farnham
- Zureik E, Lyon D, Abu-Laban Y (eds) (2011) *Surveillance and control in Israel/Palestine: population, territory and power*. Taylor–Francis, Abingdon

Surveys on Violence Against Women

Kristiina Kangaspunta¹ and Ineke Haen Marshall²

¹United Nations Office on Drugs and Crime (UNODC), Vienna, Austria

²Department of Sociology and Anthropology and School of Criminology and Criminal Justice, Northeastern University, Boston, MA, USA

Synonyms

[Domestic violence](#); [Gender-based violence](#)

Overview

The general consensus to measure violence against women is that population-based stand-alone surveys are the instruments of choice for collecting statistics on the victimization of women while recognizing that a well-designed module within a general or other purpose survey would be an appropriate tool as well. Several countries have measured violence against women using different types of survey methods. This entry examines different ways of measuring violence against women giving examples of national and multi-country surveys.

Introduction

The international community has been concerned with violence against women (VAW) or gender-based violence for a number of years. The first United Nations resolution on “Abuses against women and children” was adopted in 1982 calling upon the Member States to “take immediate and energetic steps to combat these social evils.” The Convention on the Elimination of All Forms of Discrimination against Women was adopted by the United Nations General Assembly in 1979. The Committee on the Elimination of Discrimination against Women (CEDAW) monitoring the national implementation of the Convention defines gender-based violence as violence that is directed against a woman because she is a woman or that affects women disproportionately. It includes acts that inflict physical, mental, or sexual harm or suffering, threats of such acts, coercion, and other deprivations of liberty.

In 1993, the United Nations General Assembly adopted the Declaration on the Elimination of Violence against Women which also introduced the first internationally agreed upon definition on violence against women:

For the purposes of this Declaration, the term “violence against women” means any act of gender-based violence that results in, or is likely to result in, physical, sexual or psychological harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or in private life. (United Nations 1993)

Soon after the adoption of the Declaration in 1993, the lack of reliable data on gender-based violence was recognized. The United Nations Beijing Platform of Action in 1995 recommended the collection of prevalence data to better understand different forms of violence against women. The recommendation has been repeated several times in different international forums.

The hidden nature of gender-based violence makes it difficult to measure its prevalence. Police records reflect no more than the tip of the iceberg. Victimization surveys have been seen as key to reveal the experiences of the victims of violence against women which could not be

captured in the official records. The first large scale victimization surveys focusing on gender-based violence were carried out in the Netherlands and Canada. The prevalence rates for sexual violence and other violence were generally seen as shocking and caused disbelief. Soon after, other countries started to carry out surveys with similar results. Most violence against women surveys have been done at the national level. Comparative research involving several countries has proven to be problematic because of differences in cultural and legal definitions and data collection methods. However, there are now a growing number of comparative studies on different aspects of violence against women even though the number of countries participating in these surveys has been rather limited.

Measuring Violence Against Women

Indicators of Violence

Much of the research on violence against women has focused on domestic violence and intimate partner violence (IPV), but gender-based violence is a much wider issue. The United Nations, true to its commitment to human rights, has played an important role in broadening the concept. Gender-based violence can take many forms, and it can take place within or outside the family; it can occur in the workplace, the doctor’s office, in state institutions, on the street, or other public places. The victims may be children, adults, or the aged, and the offenders can be total strangers or those that are known to the victim. Violence can take the form of harassment, emotional and psychological violence, economic abuse or exploitation, sexual and physical assault, or murder. Even though many forms of violence against women are criminalized, in many parts of the world women still experience violence outside the reach of the law. It is virtually impossible to capture this very broad range of harmful practices by any standardized measurement process (Kangaspunta and Marshall 2012).

Gender-based violence is a universal phenomenon which takes different forms in different cultural, social, and economic contexts. Our

knowledge about the different forms of violence against women is spread very unevenly across the globe, the most developed countries often having the most sophisticated and elaborate survey data and record-keeping systems. This situation is likely to improve in the future, since important progress is being made on the development of global statistical indicators of violence against women.

In 2006, the United Nations adopted a resolution on “Intensification of efforts to eliminate all forms of violence against women” requesting the Member States to ensure systematic collection and analysis of data. Also, the Statistical Commission was requested to develop a set of possible indicators on violence against women in order to assist States in assessing the scope, prevalence, and incidence of violence against women.

Following the request of the resolution, the United Nations Statistical Commission adopted a list of indicators on violence against women which are based primarily on two criteria, that is, on the availability of data at the national level and the seriousness of the violence itself. The below list is suggested to be a starting point for initiating further work on identifying the most appropriate measurements:

1. Total and age-specific rate of women subjected to physical violence in the last 12 months by severity of violence, relationship to the perpetrator(s), and frequency
2. Total and age-specific rate of women subjected to physical violence during lifetime by severity of violence, relationship to the perpetrator(s), and frequency
3. Total and age-specific rate of women subjected to sexual violence in the last 12 months by relationship to the perpetrator (s) and frequency
4. Total and age-specific rate of women subjected to sexual violence during lifetime by relationship to the perpetrator(s) and frequency
5. Total and age-specific rate of women subjected to sexual or physical violence by current or former intimate partner in the last 12 months by frequency
6. Total and age-specific rate of women subjected to sexual or physical violence by current or former intimate partner during lifetime by frequency

In addition, harmful practices including female genital mutilation/cutting and early marriage were considered for inclusion in the indicators list. Also, a number of other manifestations of violence against women were identified that need to be assessed as possible topics for measurement, such as psychological and economic violence, stalking, physical and sexual violence in childhood, forced marriage, discrimination and economic violence at work, trafficking in women, impact of incidence of sexual violence against women on sexually transmitted diseases and HIV/AIDS, assessing risk factors, assessing the extent to which women recognize the suffered violence as a crime, and the percentage of hidden violence unreported to the authorities, or, indeed, even within the community.

It should be pointed out that currently only impressionistic and qualitative information is available for many countries. It is often collected by non-governmental organizations or, for those countries which lack accessible academic research on the topic, provided by the news media. In addition, in some countries we can glean information about gender-based violence through a review article by a local scholar (e.g., Simister et al. 2010; Wasileski and Miller 2010).

One of the most difficult tasks in measuring violence against women is to operationalize the definition of violence against women. During the 1990s, there was a considerable controversy over the concepts and terms used in different studies. “Conflict tactics” was used to describe how couples try to settle their differences. Some other studies used rather “violence” and “force” to describe the incidences of violence against women. Also, the question on whether the act or the impact of the act should be measured created prolonged discussion (Walby and Myhill 2001).

Currently, it is widely accepted that surveys should use specific behavioral measures of a range of types of violent acts rather than general terms such as violence or assault. This can at least partly prevent different interpretations of

violence among the respondents who might have individual and culture-bound understandings of what constitutes violence. For example, the International Violence against Women Survey used a list of concrete violent acts in its questionnaire including acts like throwing, hitting, pushing, grabbing, twisting arm, pulling hair, slapping, kicking, biting, strangling, suffocating, burning, scalding, and using a knife or gun. Also different acts of sexual violence were mentioned as well as the impact of violence and the steps victims take to obtain help (Johnson et al. 2008).

Measuring Prevalence

Official, administrative records, in particular police and court statistics, have as said limited capacity for measuring the prevalence and incidence of violence against women. Lack of victim willingness to report the incident to the police and reluctance by the police and court officials to record such reports in an adequate manner have historically been cited as impediments to using official data to measure violence and crime. Police statistics may be seen as measures of the criminal justice system's response to crime, rather than as measures of the crime itself. Police statistics about violent crimes against women such as sexual assault and rape, physical assault, and partner violence can be seen as indicators of social sensitivity and changing norms with regard to acceptable levels of interpersonal violence (see, for example, Mucchielli 2010). A number of countries now have specific legislation to record intimate partner violence or the more inclusive concept of domestic violence. However, these statistics primarily measure the social response of victims, bystanders, service providers, and the police to victimization (Kangaspunta and Marshall 2012).

The general consensus to measure gender-based violence is that population-based stand-alone surveys are the instruments of choice for collecting statistics on violence against women while recognizing that a well-designed module within a general or other purpose survey would be an appropriate tool as well. Several countries have measured violence against women using different types of survey methods. Targeted

surveys usually cover all forms of violence against women including physical violence, sexual violence, intimate partner violence, and threat of violence. In some cases, also harassment and stalking have been studied.

Some countries use a specific module on partner violence which is added to general victimization surveys or to health surveys to assess gender-based violence. Also, a self-completion module about personal physical and sexual victimization has been added to the general victimization surveys. In some countries, only some forms of violence against women have been studied, usually addressing partner violence and sexual violence.

Surveys have also been targeted in different ways in different countries. In some countries, only women were included in the survey; in other countries also men could participate. Also age of the respondents differ: in most surveys the minimum age has been 18, while the maximum age has varied. The UN Statistical Commission recommends the surveys to be started from age 15.

The mode of interviewing has varied in different countries. Violence against women surveys have been conducted by interviewing respondents face-to-face or over the phone; also mailed questionnaires, self-completion on a computer, or combined methods have been used. While some research argues that the use of different methods might have an impact on the data collection, some others have not found much evidence for this. Face-to-face interviewing can build up more rapport between the interviewer and the respondent making it easier to speak about victimization experiences. On the other hand, the anonymity of telephone-based interviews and self-completed interviews may increase the likelihood to reveal sensitive information. In any case, the enquiry mode may have an impact on the response rate. In general, mailed questionnaires tend to have the lowest response rate compare to the other methods (Walby and Myhill 2001). However, the impact of the enquiry method might be different in different countries depending on cultural and other factors. For example, the first violence against women survey in Finland was conducted using a mailed questionnaire with a relatively

high response rate of 70 % (Heiskanen and Piispa 1998, p. 8).

Violence against women data drawn from general victimization surveys should be interpreted with caution for many reasons. General surveys might not be able to include specific-enough questions on experiences of violence, and they may lack sensitive question wording. Also lack of selection of and training for interviewers needed to sensitively conduct interviews or the use of male interviewers may have an impact on the willingness to disclose painful memories of violence. Lack of efforts to facilitate victims to speak about their experiences such as ensuring privacy as well as the broad scope of crime victimization surveys not allowing adequately to address sensitive issues might prevent women to disclose their victimization experiences (Johnson et al. 2008, p. 12).

Also, data collection methods based on face-to-face or telephone interviews might lower the reporting of these crimes. When comparing different methods used in the British Crime Survey, it was noticed that prevalence rates for domestic violence from the self-completion method are around five times higher than rates obtained from face-to-face interviews (Flatley et al. 2010, p. 58).

The International Crime Victim Survey (ICVS) which has been carried out since 1989 includes questions on sexual violence which could be relevant in measuring violence against women. However, there are only small changes in the average sexual assault against women 1-year prevalence rate which varies between 0.6 % and 1 % in different sweeps. Measuring violence against women based on ICVS is not easy, and the figures should be interpreted with caution as mentioned above. Relatively low numbers of female respondents as well as low figures on sexual violence experiences make it difficult to ascertain statistically significant differences. Some findings also suggest that rates of sexual offences are less stable over the years than those of other types of crime. This may indicate that media campaigns or other measures can temporarily raise awareness about sexual violence which in turn increases the reporting of these incidents (van Dijk et al. 2007, p. 77). Even so,

national victimization surveys (with larger samples) when used with due caution, may be a useful source for estimates of the rate of victimization of women.

National Violence Against Women Surveys

Targeted Stand-Alone Surveys

Targeted stand-alone surveys have been carried out in many countries. Examples of the first such surveys can be found in Australia, Finland, and the USA, besides Canada and the Netherlands. The first Women's Safety Survey was carried out in Australia in 1996 to provide information on women's safety at home and in the community and to depict the nature and extent of violence against women in Australia (Mc Lennan 1996). The survey was repeated in 2005 as part of the Personal Safety Survey in Australia (ABS 2005) updating the information of the first survey concerning women's experiences of violence (Phillips and Park 2006). In Australia both male and female violence against women was studied, and the focus was on the experiences during the previous 12 months or during the relationship. In 1996, around 7 % of the respondents had experienced violence during a 12-month period; the corresponding figure in 2005 was around 6 % (ABS 2005).

In Finland, the first national survey on men's violence against women and its consequences was carried out in 1997 (Heiskanen and Piispa 1998). Also in Finland, the survey was repeated, and the second report on Violence against Women in Finland was published in 2006 presenting information about the prevalence, patterns, and trends of violence committed by men against women in 2005. Also, figures on fear of violence and how victims have sought and received help from different agencies were presented. The results showed that in 1997, 40 % of women aged 18–74 had at least once been exposed to men's violence or threats since the age of 15. In comparison, the percentage was 44 in 2005. In general, there were only small changes between 1997 and 2005 in women's

violence experiences, the most significant being the decrease of intimate partner violence and the increase of work-related violence (Piispa et al. 2006).

The National Violence Against Women (NVAW) Survey was carried out in 1995 and 1996 in the USA, and it sampled both women's and men's experiences with violent victimization. Respondents to the survey were asked about physical assault, forcible rape, and stalking they experienced at any time in their life by any type of perpetrator. Like the other surveys already mentioned, the US Survey showed that violence against women was widespread: nearly 60 % of surveyed women said they were physically assaulted either as a child or as an adult. The survey also showed that many American women were raped at an early age. Out of all women surveyed who said they had been the victim of a completed or attempted rape at some time in their life (18 %), 54 % were younger than 18 years when they experienced their first attempted or completed rape. The survey showed that violence against women in the USA was primarily intimate partner violence. Twenty-two percent of all surveyed women and 64 % of those who reported being raped, physically assaulted, and/or stalked since age 18 were victimized by a current or former husband, cohabiting partner, boyfriend, or date. Only 7 % of all men and 16 % of victimized men had similar experience of intimate partner violence (Tjaden and Thoennes 2000, pp. iii–iv).

NVAW was a onetime survey which has not been compared with any information collected after the survey. However, the Bureau of Justice Statistics' National Crime Victimization Survey (NCVS) has included a special module to their general victimization survey from 1993 to 2008. The findings cover nonfatal and fatal violent crimes committed against females including nonfatal intimate partner violence (IPV), fatal IPV, rape and sexual assault, and stalking. The victimization survey includes women and girls age 12 or older (Catalano 2009).

One of the first violence against women surveys was carried out in Ghana in 1997 where 3,047 women and men aged between 15 and

72 years were interviewed throughout the country. The high number of languages made the survey particularly challenging in Ghana since the country has more than 27 local languages. All the terms used in the study were dually translated into eight local languages representing the main ethnic groups. Another challenge was the access to the people especially women, because in many cases women need permission from husbands/household heads for interviews. Households also needed to confer with chiefs and elders (Ardayfio-Schandorf 2005).

The violent acts identified in the study included wife beating, rape, defilement, widowhood rites, forced marriages, and female circumcision. Around 23 % of the female respondents reported that they have been beaten by their husbands or boyfriends. Six percent of the female respondents stated that they had been defiled; 78 % of the perpetrators were either close relations, acquaintances, or family friends. Eight percent of women said they had been raped; 59 % of the respondents said they never reported these cases to anybody. On forced marriage, 22 % of the married females stated that their parents decided for them, and 30 % of the females reported that their parents and other closer relatives chose their partners for them. In addition, 12 % of the female respondents indicated that they have been circumcised. Respondents were also asked about widowhood rites. Thirty-one percent of widowed respondents said they were being asked to marry their dead husband's brother. Other widowhood rites which the respondents mentioned include shaving of hair, ritual bath, being confined to a room for days, and wearing of rope around the neck (Ardayfio-Schandorf 2005).

The latest targeted stand-alone surveys include the Nationwide Survey on Domestic Violence Against Women in Armenia, 2008–2009 and the National Research on Domestic Violence Against Women in Georgia in 2009, as part of the UNFPA project on Combating Gender Based Violence in the South Caucasus which was launched in April of 2008 (UNFPA 2008). In the Netherlands large scale surveys of violence against women have been conducted in

1988, 1997, and 2010 (Van der Veen and Bogaerts 2010).

General Victim Surveys with Special Module

Canada was the first country to carry out a series of large scale-targeted survey to measure violent victimization of women aged 18 and above. The survey found that 51 % of women had experienced violence at some point of life out of which 39 % was sexual violence, 34 % physical violence, and 29 % intimate partner violence (Statistics Canada 1993). The survey established a baseline to measure violence against women in Canada, and it was used to include a module on spousal violence in the Statistics Canada's General Social Survey (GSS) on victimization which is carried out every 5 years which collects information of Canadians aged 15 years and older; a spousal violence module has been used in 1999, 2004, and 2009 surveys. Sexual assault on the other hand is covered as one of the eight crime types addressed routinely in the general survey (Johnson 2006).

In Denmark, the survey information on violence against women is collected through health-related surveys. The 2000 and 2005 national health interview surveys included a self-administered questionnaire on violence against women including questions on specific forms of violence (Helweg-Larsen and Frederiksen 2008).

For France, there are two different types of surveys available. First, there is the general victimization survey in the Paris region conducted every 2 years between 2001 and 2009. It is not a national population-based survey, but it does cover a large population (representative sample of 10,500 households). This survey includes questions about sexual violence, sexual assault by an intimate, insults, and threats. Estimates of two-year victimization rates are possible. Since 2006, a more focused survey "Cadre de vie et securite" (CVS) is conducted annually by the Observatoire national de la delinquance (OND). This national survey has a representative sample of about 17,000 households. This survey uses a confidential self-completion module on personal violence asking respondents between ages 18 and 75 about annual victimization by physical

or sexual violence, either by members of their household or by non-family members (INHESJ 2010).

Since 2004/2005, the yearly British Crime Survey (BSC) has included a self-completion module asking respondents aged between 16 and 59 about their experiences of domestic abuse, sexual assault, and stalking (Smith et al. 2011, p. 69). With this method the sensitive nature of the victimization related to violence against women is taken into consideration, and reporting of these crimes is evidenced to be higher (Flatley et al. 2010, pp. 56, 58).

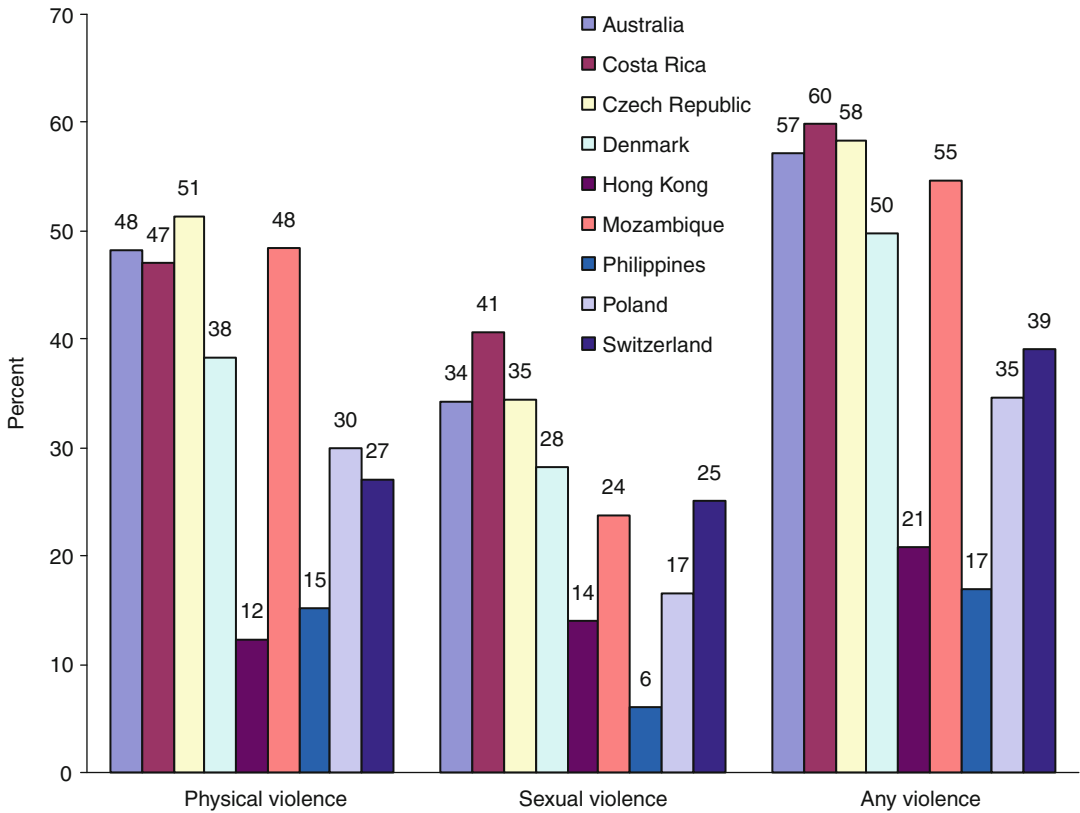
Comparing Victimization

Multi-Country Violence Against Women Surveys

As shown above, because of several differences in national surveys, they cannot be used to compare the level of victimization across different countries. For that reason, international surveys were developed where victimization data were collected in several countries using the same methodology in order to allow comparison between different countries. However, these surveys also have challenges particularly with the translation of questionnaires, interpretation of concepts, structural differences of societies, and other cultural issues.

The International Violence Against Women Survey (IVAWS) was specifically designed to target violence by males against women, assessing the level of victimization of women in a number of countries worldwide. The IVAWS relies largely on the methodology of the International Crime Victim Survey (ICVS). Between 2003 and 2005, standardized surveys were carried out in Australia, Costa Rica, Czech Republic, Denmark, Greece, Hong Kong, Italy, Mozambique, the Philippines, Poland, and Switzerland. The main findings of the IVAWS are:

- Between 35 % and 60 % of women in the surveyed countries have experienced violence by a man during their lifetime.
- Between 22 % and 40 % have experienced intimate partner violence during their lifetime.



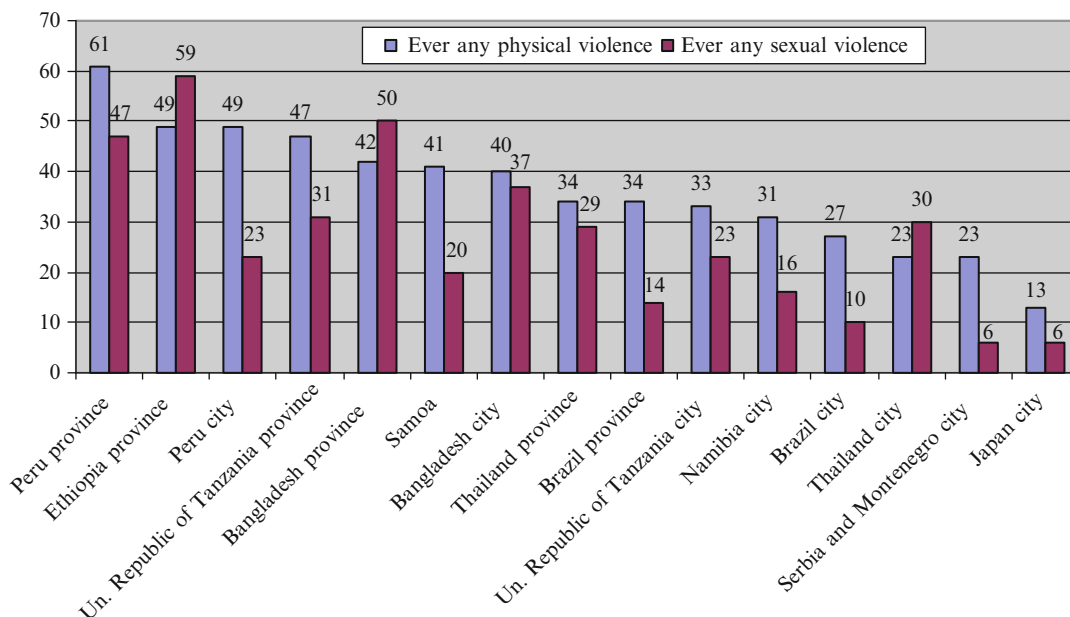
Surveys on Violence Against Women, Fig. 1 Percentage of lifetime physical, sexual, or any violence among adult women (Source: Johnson et al. 2008, 39)

- Less than one third of women reported their experience of violence to the police; women are more likely to report stranger violence than intimate partner violence.
- About one fourth of victimized women did not talk to anyone about their experiences (Johnson et al. 2008) (Fig. 1).

The WHO Multi-country Study on Women’s Health and Domestic Violence against Women collected data from over 24,000 women between 2000 and 2003 in the following ten countries: Bangladesh, Brazil, Ethiopia, Japan, Peru, Namibia, Samoa, Serbia and Montenegro, Thailand, and the United Republic of Tanzania. The study estimated the prevalence of physical, sexual, and emotional violence against women, with particular emphasis on violence by intimate partners. The study consisted of standardized

population-based household surveys which were carried out in different locations. In five countries (Bangladesh, Brazil, Peru, Thailand, and the United Republic of Tanzania), surveys were conducted in the capital or a large city and one province or region, usually with urban and rural populations. One rural setting was used in Ethiopia, and a single large city was used in Japan, Namibia, and Serbia and Montenegro. In Samoa, the whole country was sampled. The results showed that among women aged 15–49 years:

- Between 15 % (Japan) and 70 % (Ethiopia and Peru) of women reported physical and/or sexual violence by an intimate partner.
- Between 0.3 % and 11.5 % of women reported experiencing sexual violence by a non-partner.
- The first sexual experience for many women was reported as forced – 24 % in rural Peru,



Surveys on Violence Against Women, Fig. 2 Percentage of lifetime physical and sexual violence by an intimate partner among ever-partnered women (Source: WHO 2005, 29)

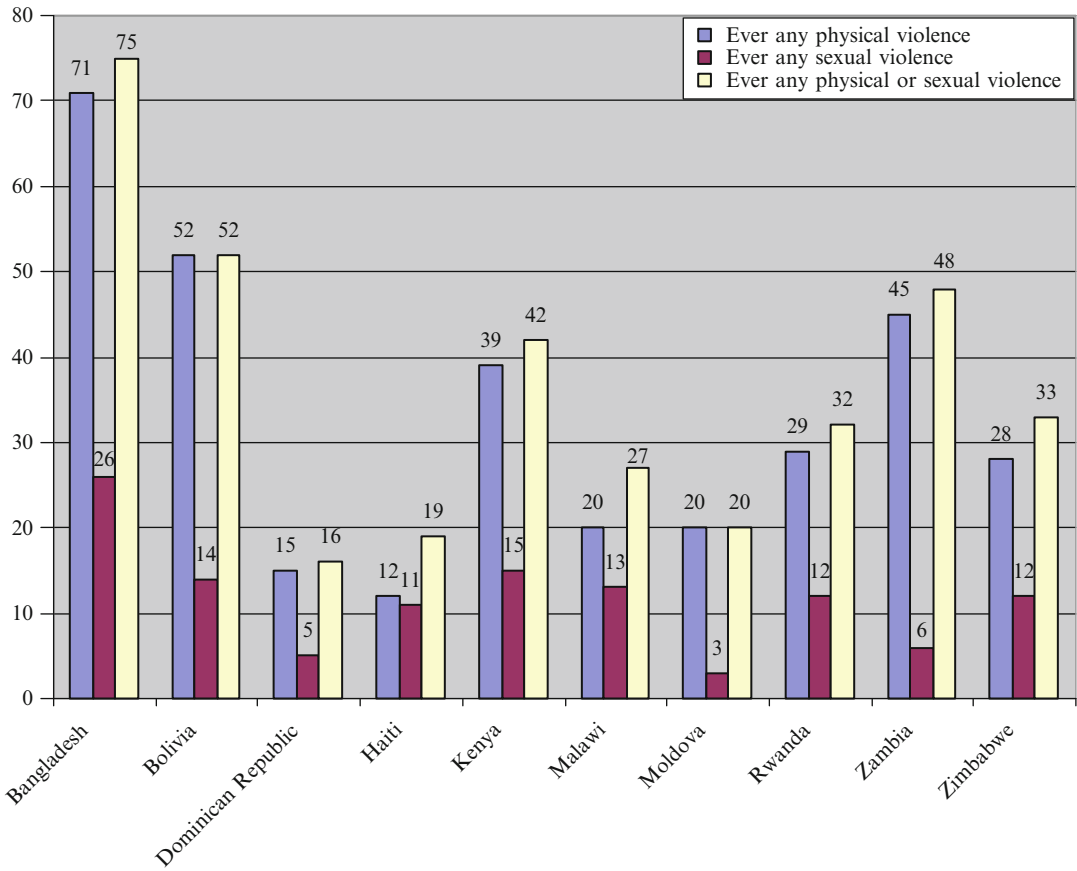
28 % in Tanzania, 30 % in rural Bangladesh, and 40 % in South Africa (WHO 2005) (Fig. 2).

Demographic and Health Surveys (DHS) are nationally representative household surveys that provide data for a wide range of topics in the areas of population, health, and nutrition. They also include a module on domestic violence assessing the prevalence of intimate partner violence (IPV) among married or cohabiting women. A comparative analysis was carried out based on data collected between 2001 and 2006 in ten developing countries including Bangladesh, Bolivia, Cambodia, Colombia, the Dominican Republic, Egypt, Haiti, India, Kenya, Malawi, Moldova, Nicaragua, Peru, Rwanda, Zambia, and Zimbabwe. The survey analysis showed that there was wide variation across countries in the prevalence of physical or sexual violence experienced by women and perpetrated by their partners: from 75 % in Bangladesh to 16 % in the Dominican Republic. The highest reported rates of physical violence were in Bangladesh (71 %), Bolivia (52 %), and Zambia

(45 %), and the lowest reported rates were in Haiti (12 %) and the Dominican Republic (15 %). The highest rates of sexual violence were reported in Bangladesh (26 %), Kenya (15 %), and Bolivia (14 %), whereas the lowest rates were reported in Moldova (3 %), the Dominican Republic (5 %), and Zambia (6 %) (Hindin et al. 2008) (Fig. 3).

Ongoing Demographic and Health Surveys (DHS) can be found in many parts of the world; the latest ones were finalized in 2011 in Angola, Ethiopia, Liberia, Madagascar, Nepal, and Rwanda (Measure DHS 2012).

A standardized survey is also being carried out in all 27 European Union countries including the candidate country Croatia on women's well-being and safety in Europe. The survey will involve standardized face-to-face interviews with about 1,500 women in each country. The interviews will cover women's experiences of violence including physical, sexual, and psychological violence, harassment, and stalking by current and former partners and non-partners. The survey will also look at violence experiences in childhood in order to create a comprehensive



Surveys on Violence Against Women, Fig. 3 Percentage of lifetime physical and sexual violence by intimate partners among currently partnered women age 20–44 (Source: Hindin et al. 2008, 18)

picture of women’s experiences of violence during their lifetime (FRA 2012).

Trend Analysis

When trends in violence against women are measured, the comparison is not done between countries but within countries based on information from different time periods. For many reasons, the level of violence against women differs significantly between countries, but that is not the focus of a trend analysis. Rather, the interest is in seeing if the data suggest that the level of violence is increasing, decreasing, or remaining about the same in countries across the globe. An important methodological advantage of comparing trends across countries (rather than levels at one point in time) is that this tends to minimize

the comparability measurement problems which hamper international comparisons of levels of violence (Marshall and Summers 2012).

A proper trend analysis is possible only if there is sufficient survey information on violence against women covering a prolonged time period. In a recent article Kangaspunta and Marshall (2012) found only seven countries with (repeated) victim surveys related to violence against women. They were Australia, Canada, Denmark, Finland, France, UK and the USA. There are also other countries which have conducted repeated violence against women surveys; however, they do not necessarily allow for a trend analysis. The case of Belgium presented below clearly shows some of the challenges of trend analysis.

Surveys on Violence Against Women, Table 1 Trend of different forms violence against women in Australia, Canada, Denmark, Finland, France, UK, and USA

	Australia	Canada	Denmark	Finland	France	UK	USA
Survey data source, violence against women	Targeted surveys 1996 and 2005	Combination of targeted survey, 1993 and general social surveys including a module on spousal violence 1999, 2004, and 2009	National health interview survey including a self-administered questionnaire on violence against women 2000 and 2005	Targeted surveys 1998 and 2005	Local victimization survey every 2 years 2001–2009 including a self-completion module on personal violence since 2007	General victimization survey including a self-completion module violence against women yearly since 2004/2005	Targeted survey 1996 and general victimization surveys 1993, 1996, 1999, 2002, 2005, and 2008
Survey-based trends in VAW							
General violence against women	–	NA	0	+	0 or +	NA	–
Partner violence	–	–	–	–	0	–	–
Sexual violence	–	–0	++	+	+	0	–
Trends in sexual assault and domestic violence – based on agency records							
Sexual assault	++	–	–	++	+	–	–
Domestic violence/IPV			+	+	+	+	
Trends in homicide – based on agency records							
Female victim homicide	–	–	–	0	–	–	–
Female partner homicide	+	–	NA	+	–	+	–

Source: Kangaspunta and Marshall 2012

In Belgium, three large-scale studies were conducted on the prevalence of gender-based violence in 1988, 1998, and 2009. The first study in 1988 analyzed incidences of violence against women; the second study in 1998 was extended to men. The 2009 survey studied the occurrence, forms, and severity of the physical, sexual, and emotional abuse to which women and men are exposed. These studies illustrate some of the difficulties performing trend analyses. While the 1998 study was exclusively concerned with violent experiences over the course of the respondents' whole life, the 2009 survey included also the experiences during the past 12 months. In the 1998 survey, a violent act was defined in a detailed manner listing 17 acts of physical and 24 acts of sexual violence; in the 2009 study violence was measured in a more synthetic manner, by grouping physical acts into three questions and one general question on sexual abuse.

In the Belgian surveys of 1988 and 1998, the age group included in the sample ranged from 20 to 50 years; the 2008 survey studied individuals between the ages of 18 and 75. In the case of the 2009 study, only abuse experienced in adult life (after the age of 18) was taken into account, whereas the figures for 1998 concern abuse experienced over the respondent's lifetime. Also, the 1988 survey was conducted using face-to-face interviews, while the more recent surveys were carried out either over the phone or online. In the report presenting the findings of the 2009 results, the authors conclude that – given the differences in methodologies – the surveys do not adequately support trend analysis (Pieters et al. 2010).

In the USA, the National Violence Against Women (NVAW) Survey was a onetime effort which has not been compared with any information collected after the survey. The US violence against women trend analysis has been conducted by using data from the Bureau of Justice Statistics' National Crime Victimization Survey (NCVS), 1993–2008 covering nonfatal and fatal violent crimes committed against females including nonfatal intimate partner violence (IPV), fatal IPV, rape and sexual assault, and stalking. The victimization survey included women and girls

age 12 or older (Kangaspunta and Marshall 2012).

The results of the trend analysis of violence against women in the most developed western countries suggest a drop in partner violence against women as measured by the standard surveys (Table 1). On the other hand, (non-partner) sexual violence appears to be increasing. This drop in partner violence is accompanied by an apparent increase in the level of female victimization reported to the police and other agencies. As Table 1 shows, homicides with a female victim were shown to be decreasing (possibly reflecting a general decline in homicides in general) (Kangaspunta and Marshall 2012).

Conclusions

The need to have sound and up-to-date information on violence against women is well recognized both at the national and international level. The United Nations as well as many other international and regional organizations has stressed the need to collect reliable and comparable data on gender-based violence. Many countries have responded to this call since 1993 when the first highly publicized violence against women survey was carried out in Canada. The first surveys were carried out in highly developed western countries only; however, soon after also countries in other regions followed with their own surveys. The latest efforts to conduct such surveys are covering more and more countries in all geographical areas. Also several new initiatives to collect comparative data have recently been launched.

Often, violence against women surveys have focused on intimate partner violence which raises particular concerns in many countries. Also other forms of violence have been studied including sexual violence and violence in public places. In general, all violence against women surveys have greatly increased our knowledge on the victimization of women, showing that high proportions of women in all parts of the world suffer from violence and abuse.

While some women face the dangers at home, most women need to be aware of the risks outside the private space as well. Women's everyday life contains numerous situations where harassment, abuse, and violence could take place such as during doctors' appointments, at the workplace, at school and other educational institutions, during leisure time – for example when participating in sports or just when having a walk in a park. Some of the new studies, particularly the survey on women's well-being and safety in Europe, aim at highlighting these latter forms of violence. The final objective of violence against women surveys should be to generate positive changes so that women all over the world could enjoy everyday life without fear of violence of any sort.

Related Entries

- ▶ [British Crime Survey](#)
- ▶ [Domestic Violence](#)
- ▶ [Feminist Theory in the Context of Sexual Violence](#)
- ▶ [Gendered Theory and Gendered Practice](#)
- ▶ [International Crime Victimization Survey](#)
- ▶ [National Victimization Surveys](#)
- ▶ [Police Family Violence Services](#)
- ▶ [Responses to Domestic Violence in India](#)
- ▶ [Understanding Cross-National Variation](#)
- ▶ [Understanding Victimization Frequency](#)
- ▶ [US National Crime Victimization Survey](#)
- ▶ [Victimization, Gender, and the Criminal Justice System](#)

Recommended Reading and References

ABS (2005) Personal safety survey Australia. Australian Bureau of Statistics, Canberra

Ardayfio-Schandorf, Elizabeth (2005) Violence against women: the Ghanaian case. Expert presentation at the expert group meeting organized by: UN division for the advancement of women in collaboration with: Economic Commission for Europe (ECE) and World Health Organization (WHO). "Violence against women: a statistical overview, challenges and gaps in data collection and methodology and approaches for overcoming them", Geneva, Switzerland, 11–14 April, 2005. <http://sgdatabse.unwomen.org/uploads/>

- [Violence against women in Ghana – paper by Elizabeth Ardayfio-Schandorf.pdf](#)
- Catalano S, Smith E, Snyder H, Rand M (2006) Intimate partner violence in the United States. Bureau of Justice Statistics
- Flatley J, Kershaw C, Smith K, Chaplin R, Moon D (eds) (2010) Crime in England and Wales 2009/10. Home Office Statistical Bulletin 12/10. London, Home Office. <http://www.homeoffice.gov.uk/rds/pdfs/10/hosb1210.pdf>
- FRA (2012) FRA survey on women's well-being and safety in Europe. European Union Agency for Fundamental Rights. http://fra.europa.eu/fraWebsite/research/projects/proj_eu_survey_vaw_en.htm
- Heiskanen M, Piispa M (1998) Usko, toivo, hakkaus. Kyselytutkimus miesten naisille tekemästä väkivallasta. Tilastokeskus. Tasa-arvoasiain neuvottelukunta. Oikeus 1998:12. Sukupuolten tasa-arvo. SVT. Helsinki
- Heiskanen M (2006) Miesten Naisille Tekemän Väkivallan Kehitys ja Kokonaiskuva. In Piispa, M., Heiskanen, M., Kääriäinen, J. and Sirén, R. (2006) Naisiin Kohdistunut Väkivalta 2005. Oikeuspoliittisen Tutkimuslaitoksen Julkaisuja 225; Yhdistyneiden Kansakuntien yhteydessä toimiva Euroopan Kriminaalipolitiikan Instituutti (HEUNI) Publication Series No. 51: Helsinki
- Helweg-Larsen K, Frederiksen ML (2008) Men's violence against women. Extent, characteristics and the measures against violence – 2007. English summary. Minister for Gender Equality, National Institute of Public Health
- Hindin MJ, Sunita K, Ansara DL (2008) Intimate partner violence among couples in 10 DHS countries: predictors and health outcomes. Macro International, Calverton, DHS Analytical Studies No. 18
- INHESJ (2010) La criminalité en France. Institut National des Hautes Etudes sur la Sécurité et de la Justice and Observatoire National de la Délinquance et des Réponses Pénales (ONDRP)
- Johnson H (2006) Measuring violence against women. Statistical trends 2006. Statistics Canada 2006. Catalogue no. 85–570-XIE
- Johnson H, Ollus N, Nevala S (2008) Violence against women: an international perspective. Springer, HEUNI, New York
- Kangaspunta K, Marshall HI (2012) Trends in violence against women: some good news and some bad news. In: van Dijk J, Tseloni A, Farrell G (eds) The international crime drop: new directions in research. Palgrave, Basingstoke
- Marshall HI, Summers D (2012) Contemporary differences in rates and trends of homicide among European nations. In: Liem MCA, Pridemore WA (eds) Handbook of European homicide research: patterns, explanations and country studies 2012, pp 39–69
- Measure DHS (2012) Demographic and health surveys. <http://www.measuredhs.com/What-We-Do/survey-search.cfm?pgtype=main&SrvyTp=year>

- Mc Lennan W (1996) Women's safety Australia. Australian Bureau of Statistics, Canberra
- Mucchielli L (2010) Are we living in a more violent society? A socio-historical analysis of interpersonal violence in France, 1970s-present. *Br J Criminol* 50:808–829
- Piispa M, Heiskanen M, Kääriäinen J, Sirén R (2006) Naisiin kohdistunut väkivalta 2005. Oikeuspoliittisen Tutkimuslaitoksen Julkaisuja 225; Yhdistyneiden Kansakuntien yhteydessä toimiva Euroopan Kriminaalipoliittikan Instituutti (HEUNI), Helsinki, Publication Series No. 51
- Phillips J, Park M (2006) Measuring domestic violence and sexual assault against women: a review of the literature and statistics. E-Brief: <http://www.aph.gov.au/library/intguide/sp/ViolenceAgainstWomen.htm>
- Pieters J, Italiano P, Offermans A-M, Hellemans S (2010) Emotional, physical and sexual abuse – the experiences of women and men. Institute for the equality of women and men, Brussels
- Simister J, Mehta PS (2010) Gender-based violence in India: long-term trends. *J Interpers Violence* 25(9):1594–1611, <http://jiv.sagepub.com/content/25/9/1594>
- Statistics Canada (1993) Violence against women – survey highlights and questionnaire package. Statistics Canada, Canada
- Tjaden P, Thoennes N (2000) Full report of the prevalence, incidence, and consequences of violence against women, findings from the national violence against women survey, national institute of justice, office of justice programs, U.S. Department of Justice, and the Centers for Disease Control and Prevention, Washington, DC
- UNFPA (2008) Combating gender based violence in the South Caucasus. <http://www.genderbasedviolence.net/content/show/4/about-us.html>
- United Nations (1993) Declaration on the elimination of violence against women. U.N. Document A/48/49, New York
- Van der Veen HCJ, Boogaerts S (2010) Huiselijk Geweld. WODC, Den Haag, Nederland
- Van Dijk JJM, van Kesteren JN, Smit P (2007) Criminal victimization in international perspective, key findings from the 2004–2005 ICVS and EU ICS. Boom Legal Publishers, The Hague
- Van der Veen HCJ, Bogaerts S (2010) Huiselijk geweld in Nederland: overkoepelend synthese-rapport van het vangst-hervangst-, slachtoffer- en daderonderzoek 2007–2010. Den Haag, WODC
- Walby S, Myhill A (2001) New survey methodologies in researching violence against women. *Br J Criminol* 41(3):502–522
- Wasileski G, Miller S (2010) Review essay: the elephants in the room: ethnicity and violence against women in post-communist Slovakia. *Violence Against Women* 16:99–125, <http://vaw.sagepub.com/content/16/1/99>
- WHO (2005) Multi-country study on women's health and domestic violence against women: summary report of initial results on prevalence, health outcomes and women's responses. World Health Organization, Geneva

Suspended Sentence

► Probation and Community Sanctions

Synthesizing Biological and Social Theorizing

Jon Heidt¹ and Gail S. Anderson²

¹School of Criminology and Criminal Justice, University of the Fraser Valley, Abbotsford, BC, Canada

²School of Criminology, Simon Fraser University, Burnaby, BC, Canada

Overview

The synthesis of biological and social environmental theories is a relatively recent phenomenon in criminology. While the early criminologists focused heavily on biological factors related to criminality, criminology eventually came to be dominated by sociological explanations of criminal behavior in the 1930s, and this continued for several decades. However, starting in the 1970s, criminologists began to reexamine biological findings relevant to understanding criminal behavior.

In modern criminology, there are several prominent attempts to synthesize biological and social environmental theorizing. In particular, these integrated theories in the biosocial and developmental research programs use a combination of biological, psychological, social, and environmental factors to explain criminality. These new theories have been successful fusing existing explanations of criminal behavior and explaining interactions between key variables from biology, psychology, and sociology.

The Reemergence of Biology in Criminology

For many years, biological explanations of criminal behavior were excluded from mainstream

criminology. This was largely a product of the domination of sociological and environmental explanations in the study of crime from the 1930s to the 1970s (Ellis and Walsh 2003). However, during the late 1970s, some criminologists began to acknowledge the importance of biology and individual differences in the formation of criminality. The emergence of the biosocial perspective offered a way to synthesize existing social environmental explanations with important biological findings relevant to understanding criminal behavior (Ellis 1990).

Recent attempts to fuse biological and social environmental theories can be found in developmental-life course theories of criminal behavior (e.g., see Farrington 1992; Moffitt 1993) and biosocial criminology (e.g., see Ellis 2005; Robinson and Beaver 2009). The notion of gene-environment interaction is important to these theories. Genes determine the different traits that are passed from parents to offspring. In some cases, these traits lead to the formation of an impulsive or aggressive personality creating a predisposition towards antisocial and/or criminal behavior (Walsh 2009). If the environmental conditions are right, an individual with an impulsive or aggressive personality may become involved in crime.

Many of these recent contributions also rely upon a risk factor approach to theorizing (Bernard and Snipes 1996). The goal of the risk factor approach is to identify important variables from different levels of explanation and explain how they interact with one another (Robinson and Beaver 2009). Before reviewing the key contemporary biosocial theories of criminal behavior, it is important to be familiar with the origin of the biosocial perspective in criminology.

Origin of the Biosocial Perspective in Criminology

In the late nineteenth century, biological and multiple factor explanations of criminality were quite popular and could be found in the work of many well-known positivist criminologists including Lombroso, Ferri, Garofalo, and

Hooton. However, starting in the early twentieth century, these explanations slowly started to lose favor and began to be displaced by theories that emphasized sociological and environmental explanations. Most prominent were theories offered by the Chicago School (e.g., social disorganization and differential association theory) and Merton's strain theory.

Modern fusions of biological and social environmental explanations can be traced back to Eysenck's (1964) theory of the criminal personality. He argued that criminality resulted from the interplay between neurophysiological and environmental conditions and that behavior was learned through classical (or Pavlovian) conditioning. More specifically, individual differences in neurochemistry may cause some people to condition less effectively which leads to the formation of various personality types some of which predispose people to criminality. He identified two personality traits, extraversion and neuroticism, as particularly important to proper conditioning. To summarize, Eysenck (1964) suggested that people with high levels of extraversion and neuroticism would have difficulties learning from punishment and other aversive stimuli, making them the most likely to engage in crime.

Eysenck (1964) claimed that extraverted individuals are especially prone to criminality because their behavior consists of attempts to stimulate their underactive arousal systems. The inactivity can be traced to a small bundle of neurons in the central part of the brain known as the ascending reticular activating system (Eysenck 1964). This system is responsible for keeping the cerebral cortex stimulated and alert. Extraverts are thought to have a pronounced insensitivity in this area of the brain. Consequently, they tend to seek out stimulation in their environment by engaging in risky, thrilling, and impulsive behaviors which often equate to various forms of criminal and/or antisocial behavior.

Neuroticism refers to how one reacts to stress. People with high levels of neuroticism are very sensitive to stress and have difficulties recovering from stressful situations. People with low levels

of neuroticism tend to be calm and emotionally stable even when stressors are present and recover quickly after experiencing stressful events. According to Eysenck (1964), high levels of neuroticism are connected to high baseline levels of activity in the autonomic nervous and limbic systems in the brain. These systems are responsible for triggering one's fight or flight response in emergencies and other dangerous situations. Neurotics are thought to be more prone to criminality because the emotional instability created by the high activity in these areas can cause overreactions or inappropriate reactions to stressful situations which may result in criminal and/or antisocial behavior.

C. R. Jeffery (1978) was another early proponent of the biosocial perspective in criminology. He noted that for nearly a half century, criminologists had ignored the findings of the biological sciences that were an important part of a full explanation of criminal behavior. While he was clearly a pioneer in the field of biosocial criminology, Jeffery failed to produce a specific theory of criminality that synthesized biological and social environmental factors. However, he did suggest the application of systems theory to the field of criminology (Jeffery 1990). The systems approach would become a key component of Robinson and Beaver's (2009) integrated systems theory of antisocial behavior which will be discussed later in this entry.

Another early example of an attempt to synthesize biological and social environmental explanations can be found in Wilson and Herrnstein's (1985) *Crime and Human Nature*. Wilson and Herrnstein argued that certain biological predispositions (e.g., aggression, impulsivity, and low IQ) affect how people respond to punishment and, consequently, their propensity to commit crime. In this theory, biological predispositions are thought to affect the operant (instrumental) learning process; this is in contrast to Eysenck's (1964) theory which stressed classical (or Pavlovian) conditioning. In other words, one's susceptibility to deterrence is affected by individual differences in biology. According to Wilson and Herrnstein (1985), behavioral problems begin with parents who fail

to discipline their children properly and continue with the criminal justice system which, in their view, had become overly lenient on criminals during the 1960s and 1970s.

It is clear that the work of both Eysenck (1964) and Wilson and Herrnstein (1985) are attempts to combine biological factors (i.e., predispositions) and social environmental factors (i.e., learning and conditioning) to explain criminality. One major shortcoming of these theories was a failure to seriously consider social factors involved in criminality outside of the family and psychological learning processes. In addition, these theories focused exclusively on explaining more serious crimes involving violence like murder and robbery while ignoring corporate and white-collar crime and other crimes common in the middle and upper classes (e.g., driving under the influence, tax evasion, and minor theft). It is interesting to note that early twin studies conducted by Mednick and his colleagues (1977) found stronger connections between heredity and petty crimes than between heredity and violent crimes. Finally, both theories seem to be geared towards explaining chronic and persistent forms of offending rather than sporadic or infrequent criminal behavior.

Contemporary theorists have attempted to address these issues by broadening the scope of their theories and by specifying interactions between key variables in their theories. These more recent theories will be reviewed in the following section.

Contemporary Syntheses

Farrington's (1992) integrated cognitive antisocial potential theory is one example of a recent attempt to fuse biological and social environmental explanations of criminal behavior. His theory also incorporates a developmental-life course approach meaning that it attempts to account for within-individual change as well as between-individual differences in offending. The key construct in this theory is the notion of antisocial potential (AP) which assesses one's potential to commit antisocial acts. In addition, AP serves as

a vehicle for integrating existing biological and sociological explanations of criminality.

Two types of AP are identified: long term and short term. Long-term AP refers to between-individual differences in individuals that are relatively stable over time. Factors that contribute to long-term AP include low cortical arousal, impulsivity, high levels of strain, weak attachments, and exposure to antisocial models (Farrington 1992). This is a clear attempt to combine variables derived from key biological findings with prominent sociological theories of criminality (e.g., strain, social control, and social learning). Short-term AP focuses upon within-individual variation in criminality. Factors for this dimension of AP are situational (e.g., anger, boredom, and intoxication) and are drawn out of the rational choice, routine activities, and opportunity theories (Farrington 1992).

Like Wilson and Herrnstein (1985), Farrington's (1992) theory addresses street crime committed by lower-class males, so the scope of the theory is quite limited. In other words, like its predecessors, this theory fails to account for middle-class and elite forms of crime and "normative" crime committed by adolescents. The main advancement here is the use of a developmental-life course approach and the focus on within-individual change in criminality. This emphasis allows the theory to account for changes in offending patterns over time.

Another theory that incorporates a developmental-life course approach is Moffitt's (1993) developmental taxonomy. She posits the existence of two types of offenders: life-course-persistent (LCP) and adolescence-limited (AL) offenders. According to Moffitt (1993), LCPs offend frequently throughout their life and account for approximately 5–10 % of all offenders. LCPs are thought to suffer from neuropsychological deficits that affect verbal (i.e., speech, reading, and writing) and executive (i.e., impulse control and attention span) functioning. These deficits also cause problematic social interactions, especially with parents, school officials, and other authority figures. Family interactions are particularly problematic since, in many cases, parents tend to have similar

neuropsychological deficits making interaction all the more difficult. In addition, these deficits predispose the individual to engage in risk-taking and impulsive behavior which often results in crime and other forms of antisocial behavior. Finally, these neuropsychological problems impede the learning of prosocial alternatives to crime, and eventually LCPs become ensnared by the consequences of crime (e.g., negative labeling and a reduction in legitimate opportunities that often lead to desistance from criminal activity). This portion of the theory is a clear attempt to explain how individual differences in biology (i.e., neuropsychological deficits) interact with various environmental factors (i.e., family and school interactions and negative labeling) to produce chronic offending patterns (Moffitt 1993).

Individual differences are thought to be much less important in explaining AL offending patterns; the key factors in AL offending are primarily social and environmental. AL offenders compose the vast majority of offenders (about 90–95 %) and commit crime during adolescent years but eventually desist by their early 20s. Moffitt (1993) argues that this type of crime is normative and is not the result of special traits or biological factors. Instead, AL offending occurs through a process of social mimicry of LCP behavior.

According to Moffitt (1993), modern society has created a "maturity gap" for teenagers in which they lack true roles. This also means that they are prohibited from engaging in adult activities like working, driving, voting, and living on their own until their late teens and early 20s. This role vacuum causes teens to search for other ways to feel mature and independent. For some of these adolescents, maturity and independence are defined as engaging in adult behaviors including drinking alcohol and having sex. Eventually, some of these individuals observe the LCPs getting rewards for their delinquent behavior (e.g., premarital sex, underage drinking, drug use, and possessions obtained through theft). For the ALs, these rewards symbolize freedom and autonomy. In order to obtain feelings of independence, the ALs begin to mimic the delinquent behavior of the LCPs.

AL offending has very little to do with individual differences per se; however, biological factors are still taken into account. Moffitt (1993) claims that societal changes have effectively lengthened adolescence by limiting or delaying roles that require responsibility and maturity. The factor of maturity and its role in development is the biological component in this part of the theory, while the notion of social mimicry and societal changes affecting developmental patterns of youth represents the social environmental component.

Ellis's (2005) evolutionary neuroandrogenic theory attempts to explain why sex is the strongest correlate of crime. To do this, Ellis (2005) suggests that high levels of testosterone give rise to competitive/victimizing behavior. Competitive/victimizing behavior is conceptualized along a continuum. On one end of the continuum are "crude" or criminal forms of behavior, which refer to attempts to injure people or deprive them of their property; in most cases, these acts have been criminalized by modern governments. Notably, white-collar or elite forms of crime are also addressed here as well. On the other end of the continuum are "sophisticated" or commercial forms of behavior, which include acts that "make no profits on the sale of goods or services, although those who administer and maintain the organization under which they operate usually receive much higher wages than do those who provide most of the day-to-day labour" (Ellis 2005: 288). According to this theory, the most serious criminality will be concentrated among adolescent and young adult males of low social status. Victimless forms of crime are left unaddressed by this theory.

This theory consists of two interrelated propositions. The evolutionary proposition suggests that aggressive and acquisitive criminal behavior evolved as an aspect of human reproduction. In short, females are thought to have mating preferences for higher status males that are capable of providing resources. These preferences have been observed in other mammalian species and are thought to hold true in humans as well. Consequently, the theory predicts that rape will always be more common among males who are unable to be stable providers.

The neuroandrogenic proposition identifies three important aspects of brain functioning that play a role in criminal behavior. First, high perinatal testosterone levels can alter brain functioning in ways that promote competitive/victimizing behavior. This is because higher levels of testosterone can affect brain development and may lead to suboptimal arousal levels, may make one prone to seizures, and may cause a rightward shift in neocortical functioning. These are all correlates that appear in other research on serious criminality (see, e.g., Eysenck 1964). Second, one's ability to learn or IQ can serve to inhibit criminal behavior. Third, executive functioning or planning ability is thought to inhibit "crude" or criminal forms of behavior. After adolescence, those with higher levels of learning and planning ability will move quickly from "crude" to "sophisticated" forms of behavior (Ellis 2005).

Because on it focuses upon testosterone and brain development, Ellis's (2005) evolutionary neuroandrogenic theory may appear to be a purely biological theory of criminality. However, the evolutionary proposition considers the role of social environmental factors relating to sex and social status. In contrast to some of the other formulations presented here, mainstream theories from the control, strain, and differential association/cultural deviance trajectories in criminology do not play a role in Ellis's (2005) evolutionary neuroandrogenic theory.

Robinson and Beaver's (2009) integrated systems theory is a wide-ranging and holistic attempt to combine biological and social environmental explanations into a unified theory. The approach used is composed of two key influences. First, the theory itself is derived from the evolutionary-ecological paradigm proposed by Vila (1994). Vila claimed that ecological, micro-, and macrolevel factors form the basis for the formation of personality (referred to as "strategic style") which determines the likelihood of criminal behavior.

Second, Robinson and Beaver's theory (2009) incorporates a systems theory perspective into the evolutionary-ecological paradigm. This is an interdisciplinary biosocial approach, meaning that it seeks to synthesize findings from

a variety of disciplines. Systems theory posits six levels of analysis: cell, organ, organism, group, organization, and community/society. Each of these levels represents an individual system, and all of these systems interact with other systems above and below them. This is essentially a risk factor approach similar to the one offered by Bernard and Snipes (1996) and used by the developmental-life course theorists (Farrington 1992; Moffitt 1993).

The biosocial basis of Robinson and Beaver's (2009) formulation is made clear in the first proposition of the theory:

All behaviours are the result of gene-environment interaction. Genes do not cause behavior; they predispose individuals to react to environmental stimuli in certain ways, meaning some will be more likely to behave in an antisocial manner. Genes are linked to numerous factors relevant for antisocial behavior, including but not limited to personality, drug use and abuse, IQ, violence, and mental illness. (367-368)

The rest of the propositions explain how various criminogenic factors contribute to antisocial and criminal behavior. A variety of factors and explanations from different disciplines are discussed and grouped into the various levels of analysis proposed by systems theory. Biological and psychological factors are considered to be part of the cell, organ, and organism levels. The cellular level of analysis refers to how genes may impact behavior and provide certain predispositions for antisocial behaviors. For example, genes determine IQ, neurotransmitter, and hormone levels all of which are connected to the brain. The main emphasis of the organ level is the brain since it plays a prominent role in behavior. The organism level is focused on the formation of personality which is considered to be a result of gene-environment interactions. This level also accounts for factors relating to diet and nutrition, drug consumption, and mental illness (Robinson and Beaver 2009).

The social and environmental factors are placed into the group, organization/community, and societal levels. The group level deals with small group interactions and incorporates explanations from sociological and social-

psychological criminology like the social learning and social control theories. Community variables are derived from sociological theories like social disorganization, routine activities, and lifestyle theory which focus on how certain areas and neighborhoods produce high levels of crime. This level of explanation also considers the impact of organizational level variables proposed by deterrence and labeling theories (i.e., related to the criminal justice system and societal reaction) might have on criminal activity. Finally, societal level explanations are drawn from macrosociological explanations of crime and incorporate variables from strain, anomie, sub-cultural, and culture conflict theories.

Robinson and Beaver (2009) also attempt to specify how the levels and variables within them interact with each other. Other theorists (Eysenck 1964; Farrington 1992; Moffitt 1993) have identified some of these interactions, especially those at the biological and psychological levels, but none have done so with the level of rigor applied here. For example, Robinson and Beaver (2009) explain that abnormal neurotransmitter, enzyme, and hormone levels impact the formation of the brain and personality, which can later lead to problems in social interaction with parents, friends, peers, and teachers. However, they identify other variables like destructive labeling, economic stress, diet, and pollution that can also affect brain development and may influence subsequent behavior. In addition, larger socioeconomic factors (e.g., tax and corporate policies) may contribute to social disorganization in certain communities. Levels of social disorganization can influence incidences of family conflict, which can also affect biosocial development. Further, social disorganization often gives rise to unsupervised peer groups which play role in the formation of criminal and antisocial behavior (Robinson and Beaver 2009).

Robinson and Beaver's (2009) theory does an effective job of organizing the important findings from a number of disciplines. The attempts made to specify the relationships between the different variables are also useful; however, further research will be needed to truly understand, identify, and refine all of these relationships.

Problems in previous theories (Farrington 1992; Moffitt 1993) are also avoided because the scope has been broadened, allowing this theory to more seriously consider how white-collar and corporate crime may be explained.

Conclusion

The biosocial and developmental approaches have proven to be extremely useful as a starting point for synthesizing biological and social environmental explanations of criminal behavior. However, there are still a number of lingering controversies and contentious debates that should be discussed. First, early biosocial theories (Eysenck 1964; Wilson and Herrnstein 1985) generally downplayed the role of social and environmental factors in criminality. Much of this was related to an implicit emphasis on identifying serious and violent chronic offenders rather than explaining more common types of crime committed by youth or less serious offenders. Again, it is important to bear in mind that early studies of twins found biological connections to petty, and not violent, crimes (Mednick and Christianson 1977). This disjunction between theory and research indicates that there is still an interesting puzzle for criminologists to solve.

Later theories (Farrington 1992; Moffitt 1993; Ellis 2005; Robinson and Beaver 2009) attempted to address this problem by more directly incorporating social factors into the mix of variables. However, biological factors and individual differences still play a central role in these formulations. Further, many of these theories fail to account for white-collar and corporate forms of crime, so their scopes are somewhat narrow. Notable exceptions here include Ellis's (2005) evolutionary neuroandrogenic theory and Robinson and Beaver's (2009) integrated systems theory.

Second, some commentators have suggested that there is currently an overreliance on the risk factor approach in criminology (Moffitt and Caspi 2006; Wikstrom 2008). Specifically, these critics charge that simply assembling lists of key risk factors does not advance the state of

criminological theory. The underlying issue here is confusion between correlation and causation; in other words, many risk factors are based on correlates of crime rather than causal factors. In order to break free of this problem, theorists have attempted to specify interactions between different variables and more specific causal processes that exist within theories (Robinson and Beaver 2009). Unfortunately, many of these relationships and causal mechanisms remain somewhat unclear and unspecified.

Despite these problems, biosocial and developmental-life course theories have done an admirable job of advancing criminological theories past the point of purely sociological explanations of criminal behavior. Recent attempts to integrate biological and social environmental explanations have clarified the relationships that variables have with one another and have identified some interactions that may exist between these variables. However, more unified theories will require more research that clearly specifies all of the potential relationships the various biological and social environmental factors may have with each other.

Related Entries

- ▶ [Career Criminals and Criminological Theory](#)
- ▶ [Cognitive/Information Processing Theories of Aggression and Crime](#)
- ▶ [Evolutionary Theories of Criminal Behaviors](#)
- ▶ [Eysenck's Personality Theory](#)
- ▶ [History of Criminological Theories: Causes of Crime](#)
- ▶ [Hormones, the Brain, and Criminality](#)
- ▶ [ICAP Theory](#)
- ▶ [Neurology and Neurochemistry of Crime](#)
- ▶ [Risk factors for Adolescent Sexual Offending](#)

Recommended Reading and References

- Beaver KM, Walsh A (2010) *Biosocial theories of crime*. Ashgate Publishing, Burlington
- Bernard TJ, Snipes JB (1996) Theoretical integration in criminology. *Crime Justice* 20:301–348

- Ellis L (1990) Introduction: the nature of the biosocial perspective. In: Ellis L, Hoffman H (eds) *Crime in biological, social and moral contexts*. Praeger, New York
- Ellis L (2005) A theory explaining the biological correlates of criminality. *Eur J Criminol* 2(3):287–315
- Ellis L, Hoffmann H (1990) *Crime in biological, social and moral contexts*. Praeger, New York
- Eysenck H (1964) *Crime and personality*. Methuen, London
- Eysenck H, Gudjonsson GH (1989) *The causes and cures of criminality*. Plenum, New York
- Farrington D (1992) Explaining the beginning, progress, and ending of antisocial behavior from birth to adulthood. In: McCord J (ed) *Facts, frameworks, and forecasts: advances in criminology theory*, vol 3. Transaction, New Brunswick
- Fishbein D (2001) *Biobehavioral perspectives in criminology*. Wadsworth Publishing, Belmont
- Jeffery CR (1978) *Biology and crime*. Sage, Beverly Hills
- Jeffery CR (1990) *Criminology: an interdisciplinary approach*. Prentice Hall, Englewood Cliffs
- Mednick SA, Christiansen, KO (1977) *Biosocial bases of criminal behaviour*. New York: Gardiner
- Moffitt T (1993) Adolescence-limited and life-course-persistent antisocial behaviour: a developmental taxonomy. *Psychol Rev* 100:674–701
- Moffitt T, Caspi A (2006) Evidence from behavioral genetics for environmental contributions to antisocial conduct. In: Wikstrom P-O, Sampson RJ (eds) *The explanation of crime: contexts, mechanisms, and development*. Cambridge University Press, Cambridge
- Rafter N (2008) *The criminal brain: understanding biological theories of crime*. New York University Press, New York
- Robinson M, Beaver K (2009) *Why crime? An interdisciplinary approach to explaining criminal behavior*. Pearson/Prentice Hall, Upper Saddle River
- Vila BJ (1994) A general paradigm for understanding criminal behaviour: extending evolutionary ecological theory. *Criminology* 32:501–549
- Walsh A (2009) *Biology and criminology: the biosocial synthesis*. Routledge, New York
- Walsh A (2011) *Social class and crime: a biosocial approach*. Routledge, New York
- Walsh A, Beaver K (2009) *Biosocial criminology: new directions in theory and research*. Routledge, New York
- Walsh A, Ellis L (2003) *Biosocial criminology: challenging environmentalism's supremacy*. Nova Science, New York
- Wikstrom P-O (2008) In search of causes and explanations of crime. In: King RD, Wincup E (eds) *Doing research on crime and justice*. Oxford Press, Oxford
- Wilson JQ, Herrnstein R (1985) *Crime and human nature*. Simon and Schuster, New York

Systematic Social Observation

John D. McCluskey¹, Roger B. Parks² and Stephen D. Mastrofski³

¹Department of Criminal Justice, Rochester Institute of Technology, Rochester, NY, USA

²Indiana University Bloomington, Bloomington, IN, USA

³George Mason University, Manassas, VA, USA

Overview

Systematic social observation (SSO) is the direct observation of social phenomena in their natural settings. It is often a group enterprise with many researchers using a systematized protocol to gather quantified data. Its application in criminology has been sparse compared to official data or survey methods, but it offers unique measurement capacity that can prove valuable for many issues in criminology.

Below, SSO is first defined, and then the method is described along with issues of validity and reliability. Next, a review of the application of SSO to criminological questions offers some insight into the variety of settings in which the approach has been used, and finally this entry concludes with speculation regarding the future configuration of SSO and its role in contributing to an understanding of important questions in criminology.

Definition of Systematic Social Observation

Many individuals engage in social observation, for example, standing on a street corner and people watching or observing audience members at a public event. Viewing persons in natural social settings (or even the setting itself, e.g., assessing the quality of neighborhood housing in a location one wishes to move) is an act of observation. What distinguishes SSO is the *systematic* application of rules and protocols that structure the act

of observation. Those explicit rules determine what constitutes relevant information for the observer and how that information is to be structured (coded) for subsequent analysis. Explicit rules for observation and coding make meaningful comparisons of events observed by multiple observers and at different times and places (Reiss 1971). One might say that SSO is structured watching with a purpose. SSO is a powerful tool for the study of human behaviors. Its power derives from the focus and structure it brings to observation of human behavior in its natural settings.

History

In the late 1920s and early 1930s, scholars studying early childhood social development practiced systematic social observation, using methods “. . . designed . . . to ensure consistent recordings of the same events by different observers . . .” (Arrington 1943, p. 83). Systematic social observation came to criminology at the hand of Albert J. Reiss, Jr., who encouraged social scientists to shed some “nonsensical” views about the limits and benefits of different forms of observing social phenomena (Reiss 1968, 1971). Reiss objected to the notion that direct observation of social phenomena in their natural setting was work for solo researchers using qualitative methods, while survey research was suitable as a group enterprise with many researchers using a systematized protocol to gather quantified data. Reiss argued that both direct social observation and survey research were in fact forms of observation that must confront the same set of challenges to producing interpretable information, that both were amenable to either solo or group practice, and that both could be used effectively for discovery or validation of propositions about social phenomena. Beyond these insights, Reiss’s important contribution to criminology in this area was the development and practice of the techniques of SSO. Reiss demonstrated how SSO could be used to answer important questions about what influences police-citizen interactions, with implications for theories about police-citizen

relationships and for public policies concerning justice, race relations, and crime control.

Suitability of SSO

What makes SSO especially valuable to researchers gathering data directly in the natural setting are precision of the observations and the independence of the observer from that being observed (Reiss 1971, p. 4). For example, some classic qualitative field research pioneered researcher access to the police occupation, but the necessarily selective samples of these solo researchers appear to have overstated the uniformity of police practice (Skogan and Frydl 2004, p. 27). SSO researchers have observed considerable variation in the way police use their authority, and some have shown the high degree of variability that may be found with the same officer over time. Precision is also accomplished through the sequencing of events and the detailing of context, matters that may not be well documented by official records or accurately recalled by participants when interviewed – for example, how police encounters with the public escalate into rebellion or violence (Sykes and Brent 1983). Sometimes SSO precision derives from the application of complex standards or expectations to the practices of persons with obligations to perform in particular ways. For example, the extent to which legal actors conform to constitutional standards or a professional standard can be assessed. Further, SSO can be used to determine the extent to which justice officials comply with the preferences of citizens they encounter or whether citizens comply with the preferences of justice officials in everyday situations (e.g., Mastrofski et al. 1996).

SSO is especially desirable when the question demands detailed knowledge of situations, conditions, or processes that are not otherwise well illuminated or where there is reason to question the validity of knowledge based on other forms of data collection. SSO may also be useful in studying people who might find it difficult to provide an objective or accurate account of what the researcher wishes to know (such as their behavior

and the context of that behavior in highly emotional situations). Where there are strong temptations to omit, distort, or fabricate certain socially undesirable features, such as illegal, deviant, or otherwise embarrassing situations, SSO offers an independent account. This is, for example, a limitation of survey-based citizen self-reports of encounters with police to deal with a problem caused by the survey respondent, and especially problematic if there is systematic variation in the degree of error across important subgroups within the sample, for example, according to race.

While much of the SSO research has focused at the level of individual persons as decision-makers, the 1980s saw the beginning of studies that use an ecological unit, such as the neighborhood block face, as the unit of SSO analysis. Noting that neighborhood residents find it difficult to offer accurate descriptions of their neighborhood's social and physical environment, Raudenbush and Sampson (1999) highlighted the value of an "ecometric" approach that uses SSO in conjunction with neighborhood survey research to more fruitfully characterize the state of neighborhood physical and social structure.

SSO may be especially well suited to situations and events where all of the relevant actors and events pertinent to the phenomenon of interest can be observed from start to finish in a limited, well-defined time period. For example, the police decision on how to deal with a traffic violator is clearly bounded in time and place. To the extent that the decision is heavily influenced by the context of the immediate situation (e.g., the offense, the evidence, the driver's demeanor), the decision on how to treat the traffic offender can be captured by SSO and has been especially useful in understanding informal sanctions, which are not completely captured in official records. In general, SSO lends itself to observing phenomena that occur either with high frequency, such as drivers' noncompliance with speed limits on public highways, or at predictable times and places, such as criminal trials, scheduled meetings between probation officers and offenders, or even field tests of prison security systems. Events that occur less frequently, such as acts of social disorder in public places, may require

considerably more observation time to obtain reliable estimates (Raudenbush and Sampson 1999), or they may be so infrequent and unpredictable as to make SSO simply impractical, such as the police use of lethal force or the life course of criminality in a sample of individuals. One of the most frequent uses of SSO has been investigating how criminal justice workers operate in the context of role expectations generated by their organization or profession. Primarily focused on police, SSO research in criminology has been very concerned with how officers negotiate the tension between the formal (legal, bureaucratic, and professional) standards set for them and those that issue from the occupational culture. SSO could also be applied to role conformance in the context of informal or illegitimate organizations, such as gangs.

SSO is often used in conjunction with other forms of observation. Some studies have used SSO to measure the extent to which treatment conditions in randomized trials have been maintained (Sherman and Weisburd 1995, p. 685). SSO data have been linked to other forms of data collection on research subjects, such as census data (on neighborhoods), survey interviews of police officers, and follow-up interviews with citizens who were observed in encounters with police. And sometimes SSO is used to supply data not otherwise available, such as objective measures of the physical and social disorder in urban neighborhoods (Raudenbush and Sampson 1999).

Perhaps the most frequent reason that criminologists have turned to SSO is their dissatisfaction with the data they could obtain by other means, such as official records and survey research (Buckle and Farrington 1984, p. 63). Self-report and victim surveys have a number of biases and limitations, but since these are different from those inherent in SSO, it can provide an alternative perspective on specific phenomenon relative to those approaches (Parks 1984).

Unit of Analysis and Sampling

Planning the selection of what is to be observed is an essential element for SSO. Like survey

interviewing, SSO requires a careful focusing of what is to be observed and makes it possible to estimate parameters and evaluate error. The first step requires establishing the unit of analysis. Given that much SSO focuses on social interactions, there are three distinct approaches (McCall 1984, pp. 268–269). One uses a time period as the unit of analysis, observing what happens within each discrete time segment, such as what behaviors police officers display in a 15 minute segment of time or the level of social disorder on a street segment during a one-hour period. Another uses a behavior or act as the unit, tracking the sequencing of different behaviors over time, such as the behavioral transactions between officers and citizens who engage each other. The third approach is to socially construct an event as a unit of observation, such as a face-to-face encounter between a police officer and citizen (Reiss 1971) or a public meeting between police and members of a neighborhood organization (Skogan 2006).

Once the unit of analysis is decided, the researcher must consider the sampling frame. The same principles of sampling apply to SSO as any other data collection method, such as survey research. The researcher must consider where and when the units of interest may be found and determine an efficient method of capturing a representative sample. An example of a straightforward sampling strategy is an SSO study of shoplifting that randomly selected shoppers entering a store, systematically varying the location of observers among entrances to the store (Buckle and Farrington 1984).

SSO researchers often use more complex sampling strategies focusing on geographic space. They have sampled police beats and specific days and times within them, oversampling places and times where higher levels of police-citizen encounters were expected. Some researchers rely upon the observed subjects making their own choices as to where observers conduct their observations. This makes sense when the object of study is a specific research subject, but when the object of study is the geographic entity itself, an independent sampling plan is required. For example, a study of public order in a park

required researchers to conduct hourly park patrols to observe and record activities of persons by location within the park (Knutsson 1997). Some researchers have used a smaller geographic unit than a police beat or park. Several studies use the face block to apply SSO to the measurement of disorder on public streets, defined in terms of traces of physical and social disorder (trash, graffiti, loitering, public intoxication) (Sampson and Raudenbush 1999). In one study, observers drove through street segments videotaping what was viewable from the vehicle (*ibid*). Others have performed live observation at the “epicenter” of each block face (best location to observe the most activity), randomly selecting short periods of time for observation from that location and recording them on check sheets (Weisburd et al. 2006). But observers could focus on single addresses, as might be done if one were interested in observing the extent of different kinds of desired and undesired social activity at crime hot spots. Even smaller spatial units have served as the sampling frame. A study of the relationship between crowding and aggression in nightclubs selected high traffic areas within the establishment (10 m²) to observe levels of patron aggression for 30-min time periods (Macintyre and Homel 1997).

While much of the extant SSO research must develop time- or area-based sampling frames that capture unpredictable or unscheduled events, some SSO studies have focused on scheduled events, such as the delivery of therapeutic community programs in correction institutions or the previously mentioned police-community neighborhood meetings. Sampling of regularly scheduled events is common in research on educational practices and physician behavior, a practice easily replicated for certain aspects of the legal process of interest to criminologists.

Sometimes the practicalities of conducting successful field observation make the research vulnerable to sample biases. In cases where consent of those to be observed must be secured, a clear bias is introduced when those who refuse to be observed differ in their behaviors from those who are willing to be observed. The practical requirements of observation can introduce bias

as well. For example, the observation of disorder on Chicago block faces required light that was sufficient for observation only between 7 am and 7 pm (Sampson and Raudenbush 1999), meaning that researchers were unable to measure many forms of disorder that mostly occur in the darkness. It would also be challenging to observe many aspects of law enforcement inquiry and exchanges in the investigative and prosecutorial processes, because much of the effort is not limited to face-to-face encounters, but rather occurs through telephone and computer, modes of communication that may necessitate very different sampling frames and observational methods. Particularly challenging are studies that require a sampling of *cases* rather than individual decision-makers, inasmuch as it is difficult to track and observe the behavior of many different persons who may be involved in making decisions about a case.

Instrumentation

Principles that apply to other forms of research also apply to the creation of instruments for structuring and recording SSO. Sometimes the instrument takes the form of a tally sheet or log for recording the frequency at which phenomena were observed, such as counting disorderly elements at block faces (Sampson and Raudenbush 1999) – or the timing and duration of events, such as police presence in a hot spot (Sherman and Weisburd 1995). Often the instrument takes the form of a questionnaire that is directed to the observer. A study of police use of force, for example, might ask observers to code a series of close-ended questions about the citizens involved in an incident (their personal characteristics, their appearance, their behavior), the behavior of police (how much and types of force used), and other features of the situation (location of the event, its visibility, the presence of bystanders).

SSO instruments have the desired effect of focusing observers' attention on items selected for observation. Field researchers have demonstrated a substantial capacity to recall the relevant features of long sequences of these events, given

the repetitive use of the protocols. Nonetheless, greater complexity in the coding system heightens the risk of error. The accuracy of such recall is undoubtedly variable, but researchers have not assessed most of the correlates of recall accuracy (e.g., observer characteristics, instrument characteristics, and the observational setting).

Recording Observations and Use of Technology

Two issues arise in recording of phenomena observed through SSO: (a) whether it is contemporaneous with the observation or later and (b) whether technological recording devices are employed. Resolving these issues requires choosing the highest priority and what must suffer as a consequence. The more contemporaneous the recording of an observation, the less the vulnerability to recall error and various forms of bias, but in many cases, the act of recording may increase the reactivity of the observed parties to the process of being observed, as, for example, when observers posing as shoppers follow actual shoppers to observe whether they are shoplifting (Buckle and Farrington 1984, 1994).

Employing technological aids is usually intended to increase the accuracy or detail of observation from that which would be otherwise available. Handheld electronic recording devices have been used in observing police-public interactions and in observing the social and physical environment of neighborhoods. Audiotaping of calls for service to police telephone operators has been used to gather data on police workload. Videotaping neighborhood block faces from a slow-moving motor vehicle has been used to observe neighborhood disorder. Use of handheld personal digital devices allows contemporaneous observation and recording of brief, frequent events and is most practical when the number of aspects to be observed per event is small in number, which minimizes the interference of recording events occurring in close succession with observing them (McCall 1984, p. 272).

The major advantage of initially recording events electronically and then encoding those records later for analysis is not only the elimination of recall problems but also that more detailed and accurate observations may be made, and the testing of interobserver reliability is facilitated. Further, field researchers may in some instances feel safer when they make the initial recording of their observations from the security of a moving vehicle or when events are recorded with a remote and unobtrusive device (time-lapse photography of a street corner). Nonetheless, there are a number of drawbacks that have concerned other researchers. Video recordings cannot exactly replicate what all of the available senses would communicate to an observer who was there “in the moment.” The use of drive-by photography can be expensive; it raises ethical concerns because denizens of the neighborhood find them intrusive and perhaps anxiety producing; it may raise legal and human subject protection issues if alleged criminal acts are recorded; and they may cause significant reactivity among those being observed. In other instances, however, the pervasiveness of already-present surveillance technology that records observations in readily shared (digital) formats (closed circuit television in public and mass-private settings or in-car police video cameras to record traffic stops) may afford researchers a relatively unobtrusive source of data that does not encounter these problems. Dabney, Hollinger, and Dugan (2004), for example, used augmented video surveillance to study shoplifters. The researchers followed a sample of drugstore customers at a single location equipped with high-resolution video surveillance cameras. They were able to determine which customers engaged in shoplifting and coded data on customers’ personal characteristics, as well as behavior. Increasingly, much technology-based surveillance derives its unobtrusiveness, not from its being unknown to subjects, but that it is taken for granted (Shrum et al. 2005, p. 11). And with the advent of nonlinear editing packages for digital video, the data itself (traditionally analyzed in quantitative or text format) can be readily manipulated and analyzed as images (Shrum et al. 2005, p. 5).

Error, Reliability, and Validity

SSO data are subject to the same range of threats that befall other methods. Error can be introduced by the observer, and issues of reliability and validity of the method must be addressed. Observers can introduce error intentionally (cheating) or unintentionally (bias or reactivity). Cheating is rarely reported in SSO, although its frequency is unknown. It seems likely that most instances of SSO cheating go undetected. Shirking, a more subtle form of cheating, may occur if observers attempt to reduce their workload by failing to record events that would require writing extensive narratives and structured coding. There has been no direct systematic assessment of the extent and impact of this form of shirking in SSO, but one researcher did examine the effects of time on the job (a proxy for burnout) on researcher productivity but found that productivity was not significantly related to time on the job lower (Spano 2005, pp. 606–608).

Sources of unintended biases in SSO are the mindset and prejudices that observers bring to the field or develop on the job. These may affect what they observe and how they interpret it. The research exploring these issues for SSO does not offer clear and consistent findings. Reiss (1968, 1971b, pp. 17–18) found that an observer’s professional background (law student, sociology student, or police officer) did have consequences for some types of information, but not others. A later study attempted to determine whether a statistical relationship between observed police orientation to community policing and officer success in securing citizen compliance could be attributed to observers’ own views on community policing (Mastrofski et al. 1996, p. 295). A clear association was not found between the observers’ attitudes and the effects that their analysis produced.

Some types of observation judgment are undoubtedly more vulnerable to personal bias than others. Some research, for example, required field observers to judge whether police officers applied excessive force against citizens. But a more effective strategy may be to bifurcate the process into (a) recording narrative accounts of what happened (without asking the field observer

to make a judgment about excessive force) and (b) having separate, specially trained experts review these accounts and make an independent judgment about whether they constitute a violation of some standard, legal, or otherwise.

No responsible researcher can casually dismiss the risks of reactivity in SSO, but a number suggest that it is dependent on the context of the observational setting and the nature of the relationship between observer and observed. In SSO of police it has been argued that reactivity to the observer can be reduced by the observer downplaying any evaluative role and emphasizing one's naivety as a "learner" (Mastrofski et al. 1998; Reiss 1968). Yet even this approach may generate more "teaching" or show-off activity in police subjects. Observed officers who engage in such teaching have been reported making contact with citizens to illustrate elements of police work to the observer (Mastrofski and Parks 1990, p. 487). And some types of observers (e.g., females in the presence of male police officers) may produce more of this effect than others (Spano 2007, p. 461).

One of the distinct advantages of SSO over solo field research is that it facilitates testing and improvement of the *reliability* of observations. Early on, much attention was given to the use of multiple observers and estimating their inter-rater reliability. Where many researchers can independently observe the same phenomenon by having multiple observers on scene or by using video recordings, the testing of inter-rater reliability is accomplished by measuring the extent of agreement among the pool of observers for the same set of events. Sometimes disparate independent observations of the same event are resolved by a process of discussion and negotiation. Where multiple independent observations of the same event are not possible, and that is often the case in situations where having more than one observer would be too disruptive, observers might be tested by using their detailed narrative descriptions to determine (a) if they are properly classifying phenomena according to the protocol and (b) the extent of agreement among persons who use those narratives to make classifications. For example, this has been done for

characterizing a wide range of police and citizen behaviors in predicting citizen compliance with police requests (McCluskey 2003, pp. 60–74).

Recently, SSO researchers have broadened their reliability concerns to incorporate measurement accuracy and stability. Raudenbush and Sampson (1999) apply psychometrics to the development of "ecometrics" to better understand the error properties of SSO data gathered from observing physical and social disorder in urban neighborhoods. They adapt three psychometric analytic strategies: item response modeling, generalizability theory, and factor analysis to illuminate the error structure of their observational data and to make judgments about the best ways to limit different sources of error in future observational studies. For example, they find that physical disorder can be more reliably measured at lower levels of aggregation than social disorder, due to the much lower frequency of the latter in their observations (Raudenbush and Sampson 1999, p. 30).

While textbooks often note that field studies are less vulnerable to validity problems than surveys because the method places the observer "there" while events are unfolding, they are of course only valid insofar as they produce data that measure what the researcher intends for them to measure. This suggests that issues of validity are related to the protocols, rules, and frameworks that guide the observation. For example, the validity of observations is weakened when observers make inferences about motives or psychological conditions, which cannot be directly observed.

SSO Contributions to Criminology and Future Opportunities

SSO data have made major contributions in two areas of criminological research: the behavior of rank-and-file police officers and the measurement of disorder in urban neighborhoods. SSO has dominated the empirical research on the discretionary choices of police officers: making stops and arrests, issuing citations, using force, assisting citizens, and displaying procedural

justice (Skogan and Frydl 2004:ch. 4). One of SSO's special contributions has been the scope of explanatory elements made available for the researchers' models. These include many details of not only the officer's behavior but also the context in which it occurs (nature of the participants, their behavior, the location, and the neighborhood). SSO has also been instrumental in detailing the nature of the *process* of police-citizen interaction, illuminating the interactive quality of temporally ordered micro-transactions or stages that may occur in even a relatively short police-citizen face-to-face encounter (Sykes and Brent 1983). And SSO has also allowed researchers to observe elements of organizational control and community influence on the work of police officers. For example, researchers can learn more about the influence of police supervisors on subordinates' practices (Engel 2000), the dynamics of police-community interaction and their consequences when police and neighborhood residents deal with each other at community problem-solving meetings (Skogan 2006).

A second area where SSO research has concentrated is the examination of neighborhood physical and social disorder. It has been used to test the impact of police interventions in hot spots, showing that police interventions in these "micro-places" not only reduce crime and disorder, they also diffuse those benefits to nearby areas (Sherman and Weisburd 1995; Weisburd et al. 2006). The largest project in this area has focused on Chicago neighborhoods and has produced a number of insights relevant to the testing and development of theories of the role of neighborhood disorder in causing crime in urban neighborhoods (Raudenbush and Sampson 1999). Using SSO-based measures of "objective" disorder described earlier in this entry, researchers have examined the sources and consequences of public disorder. The research has demonstrated the importance of "collective efficacy" in predicting lower crime rates and observed disorder, controlling for structural characteristics of the neighborhood (Sampson and Raudenbush 1999). Collective efficacy also predicted lower levels of crime, controlling for observed disorder and the reciprocal effects of

violence. The researchers found that the relationship between public disorder and crime is spurious, with the exception of robbery, which is contrary to the expectations of the well-known "broken windows" theory of neighborhood decline.

In general, SSO has afforded precision that has in many cases shown the phenomena of interest to be more complex than other forms of data collection had indicated. For example, SSO researchers have found rich variation among police officers in their patterns of discretionary choice and even noted the instability of those patterns for individual officers over time. And the independence of SSO observers from the phenomenon of interest has provided a means to understand the contributing factors to the social construction of phenomena, such as the contributions of a neighborhood's racial profile in assessing its level of disorder (Sampson and Raudenbush 2004).

There are many opportunities to expand the use of SSO to increase knowledge and understanding. Largely untapped is the observation of crime and disorder, especially at the microlevel, where observers have the opportunity to make detailed observations of offenders in the act. SSO studies of shoplifting and aggressive or disorderly behavior in bars and clubs show that this is feasible where observers can easily blend into the environment. Where that is not possible, access to unobtrusive surveillance technologies appears to offer opportunities for detailed observation that reduce reactivity concerns. It is highly likely that criminologists will take advantage of the ubiquity of electronic surveillance to capture events that would otherwise be costly to observe. For example, the growing sophistication of surveillance and identification technology may make it possible to use facial identification software to gather data for a network analysis of persons who frequent hot spots. This includes not only the growing use of video recording devices by government and private sector organizations but the now ready availability of miniaturized recording devices to the general public (through cell phone recording devices).

In searching for efficient ways to use SSO, criminologists will likely capitalize on the growing body of evidence about the predictability of crime and disorder occurring in small geographic spaces. Because much “street” crime is so highly concentrated in a relatively small portion of addresses or face blocks, the location of observers or observational devices can very efficiently generate lots of information on what occurs, especially in public areas. In addition, given heightened levels of obtrusive surveillance in public places, SSO should prove an excellent way to understand how security and surveillance operate, why certain methods are effective, and the collateral impacts of various methods of monitoring and control designed to increase public safety.

Another venue for SSO to be used fruitfully is in experimental studies. SSO can be used to measure key aspects of the process that presumably operate to link treatments to outcomes. For example, if the physical redesign of bars and serving practices of bartenders are intended to reduce violence in those establishments, do patrons in fact alter their patterns of behavior in the ways that are expected to produce less violence (Graham et al. 2004)?

Conclusion

Four decades ago Albert Reiss showed criminologists the utility of systematic social observation, but it remains a method used infrequently. This is due in no small part to two things. First, criminologists are rarely exposed to training and opportunities to do SSO during their course of study. Second, those who know a little about it may often expect that it requires more time and resources than they have available. This may indeed be the case. Many projects could be taken on a smaller scale with a narrower scope of questions than the better-known, large SSO projects, especially if technological advances are leveraged for both recording and coding purposes. Some researchers may decline to use SSO because of reactivity concerns, but the available evidence suggests that these problems are often

manageable and may be no more severe in any event than found with other data gathering methods.

Increased use of SSO will undoubtedly attract and stimulate greater scrutiny of its limitations, as well as its advantages. The error properties of most SSO data sets have been underexplored, and more attention here is needed. Expanding the use of SSO and more comprehensively assessing its strengths and limits could be fruitfully combined into a more comprehensive assessment of other methods of gathering data on crime and justice phenomena.

SSO deserves the consideration of researchers because of its many advantages. It offers enhanced prospects of validity for the study of crime and justice phenomena, and it increases confidence in reliability because of the researcher’s direct access to the phenomenon of interest and greater control and transparency of data encoding. It affords greater precision in capturing details of the phenomenon and its context, such as the sequencing of what happens before, during, and after those events. In many cases it may be the least problematic method for acquiring information. Criminology, which has strong roots in the traditions and methodologies of sociological research, remains heavily reliant on the use of sample surveys and official records (McCall 1984, p. 277; Reiss 1971). But as the field matures and diversifies intellectually, more of its researchers may with justification be inclined to make systematic social observation the method of first, not last, resort.

Related Entries

- ▶ [Defining Disorder](#)
- ▶ [Hot Spots and Place-Based Policing](#)
- ▶ [Measuring Police Performance](#)
- ▶ [Methodological Issues in Evaluating Police Performance](#)
- ▶ [Police Discretion in Law Enforcement](#)
- ▶ [Police Discretion in Providing Services and Assistance](#)
- ▶ [Police Legitimacy and Police Encounters](#)

Recommended Reading and References

- Arrington RE (1943) Time sampling in studies of social behavior: a critical review of techniques and results with research suggestions. *Psychol Bull* 40(2):81–124
- Buckle A, Farrington DP (1984) An observational study of shoplifting. *Br J Criminol* 24(1):63–73
- Buckle A, Farrington DP (1994) Measuring shoplifting by systematic observation: a replication study. *Psychol crime law* 1(2):133–141
- Dabney DA, Hollinger RC, Dugan L (2004) Who actually steals? A study of covertly observed shoplifters. *Justice Q* 21(4):693–728
- Engel RS (2000) The effects of supervisory styles on patrol officer behavior. *Police Q* 3(3):262–293
- Graham K, Osgood D, Zibrowski E, Purcell J, Gliksmann L, Leonard K, Parnanen K, Saltz R, Toomey T (2004) The effect of the safer bars programme on physical aggression in bars: results of a randomized controlled trial. *Drug Alcohol Rev* 23:31–41
- Knutsson J (1997) Restoring public order in a city park. In: Homel R (ed) *Policing for prevention: reducing crime, public intoxication and injury*. *Crim Justice*, Monsey, pp 133–151
- Macintyre S, Homel R (1997) Danger on the dance floor: a study of interior design, crowding and aggression in nightclubs. In: Homel R (ed) *Policing for prevention: reducing crime, public intoxication and injury*. *Crim Justice*, Monsey, pp 91–113
- Mastrofski S, Snipes JB, Supina A (1996) Compliance on demand: the public's response to specific police requests. *J Res Crime Delinq* 33:269–305
- Mastrofski SD, Parks RB (1990) Improving observational studies of police. *Criminology* 28(3):475–496
- Mastrofski SD, Parks RB, McCluskey JD (2010) Systematic social observation in criminology. In: Piquero A, Weisburd D (eds) *Handbook of quantitative criminology*. Springer, New York, pp 225–247
- Mastrofski SD, Parks RB, Reiss AJ Jr, Worden RE, DeJong C, Snipes JB, Terrill W (1998) Systematic observation of public police: applying field research methods to policy issues. National Institute of Justice, Washington, DC, NCJ 172859
- McCall GJ (1984) Systematic field observation. *Annu Rev Sociol* 10:263–282
- McCluskey JD (2003) Police requests for compliance: coercive and procedurally just tactics. LFB Scholarly Publishing LLC, New York
- Parks RB (1984) Comparing citizen and observer perceptions of police-citizen encounters. In: Whitaker GP (ed) *Understanding police agency performance*. U.S. Government Printing office, Washington, DC, pp 121–135
- Raudenbush SW, Sampson RJ (1999) Toward a science of assessing ecological settings, with applications to the systematic social observation of neighborhoods. *Sociol Methodol* 29:1–41
- Reiss AJ Jr (1968) Stuff and nonsense about social surveys and observation. In: Becker HS, Geer B, Riesman D, Weiss RS (eds) *Institutions and the person: papers presented to C. Everett Hughes*. Aldine Publishing, Chicago, pp 351–367
- Reiss AJ Jr (1971) Systematic observation of natural social phenomena. In: Costner HL (ed) *Sociological methodology*. Jossey-Bass, San Francisco, pp 3–33
- Sampson RJ, Raudenbush SW (1999) Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *Am J Sociol* 3:603–651, November
- Sampson RJ, Raudenbush SW (2004) Seeing disorder: neighborhood stigma and the social construction of “broken windows. *Soc Psychol Q* 67(4):319–342
- Sherman LW, Weisburd D (1995) General deterrence effects of police patrol in crime “hot spots”: a randomized, controlled evaluation. *Justice Q* 12(4):625–648
- Shrum W, Duque R, Brown T (2005) Digital video as research practice: methodology for the millennium. *J Res Pract* 1(1):1–19, <http://jrp.icaap.org/index.php/jrp/article/view/6/11>
- Skogan WG (2006) *Police and community in Chicago: a tale of three cities*. New York, Oxford
- Skogan W, Frydl K (eds) (2004) *Fairness and effectiveness in policing: the evidence*. National Academies, Washington, DC
- Spano R (2005) Potential sources of observer bias in police observational data. *Soc Sci Res* 34:591–617
- Spano R (2007) How does reactivity affect police behavior? Describing and quantifying the impact of reactivity as behavioral change in a large-scale observational study of police. *J Crim Justice* 35:453–465
- Sykes RE, Brent EE (1983) *Policing: a social behaviorist perspective*. Rutgers University Press, New Brunswick
- Weisburd D, Wyckoff LA, Ready J, Eck JE, Hinkle JC, Gajewski F (2006) Does crime just move around the corner? A controlled study of spatial displacement and diffusion of crime control benefits. *Criminology* 44(3):549–592
- Worden RE, McLean SJ (Forthcoming) Systematic social observation of the police. In: Reisig MD, Kane RJ (eds) *Oxford handbook on police and policing*. Oxford University Press, Oxford