Chapter 6
Polygraph Testing of Sex Offenders

Don Grubin

Introduction

From tentative beginnings in the 1990s, postconviction sex offender testing (PCSOT) has become increasingly incorporated into sex offender treatment and supervision in both the United States and United Kingdom. McGrath, Cumming, Burchard, Zeoli, and Ellerby (2010), for example, reported that nearly 80 % of community adult sex offender programs in the United States and over half of residential ones make use of polygraph testing to inform treatment or supervision, while in the United Kingdom, mandatory testing of high-risk sex offenders on parole was introduced in 2014 after a number of trials. Its spread to other countries is likely, with a number of jurisdictions actively considering its use.

The growing influence of PCSOT, however, is not without controversy. The speed with which it has been embraced by programs has tended to outpace evidence, with much of its impetus coming from clinical experience supported by a research base of limited robustness. Only recently have more well-designed studies been carried out. Although this is not unusual when new procedures are introduced, PCSOT carries with it significant baggage associated with polygraph testing more generally. Thus, while proponents claim that PCSOT makes important contributions to sex offender treatment and management by bringing to attention changes in risk, facilitating disclosures, and perhaps encouraging offenders to modify their behavior (Grubin, 2008; Levenson, 2009), others are more skeptical. Commentators, for example, have argued that the type of polygraph test used in PCSOT lacks

D. Grubin, M.D., F.R.C.Psych. (✉)
Institute of Neuroscience, Newcastle University, Gosforth,
Newcastle upon Tyne NE3 3XT, UK

Northumberland, Tyne and Wear NHS Trust, St Nicholas Hospital,
Gosforth, Newcastle upon Tyne NE3 3XT, UK
e-mail: don.grubin@ncl.ac.uk

D.R. Laws, W. O'Donohue (eds.), *Treatment of Sex Offenders*,
DOI 10.1007/978-3-319-25868-3_6

validation, is unscientific, and is potentially dangerous (Ben-Shakhar, 2008; Iacono, 2008); polygraph testing may adversely affect the therapeutic alliance between the offender and therapist or supervisor (McGrath et al., 2010; Vess, 2011); the entire process is based on manipulation or intimidation and potentially breaches a number of basic ethical principles relating to autonomy and non-malfeasance (Chaffin, 2011; Cross & Saxe, 2001; Meijer, Verschuere, Merckelbach, & Crombez, 2008); and, common to all critical commentaries, there is an absence of research to show that it is effective (Rosky, 2013).

To what extent, then, does PCSOT make a positive contribution to sex offender treatment and management, a question sometimes simplified to, "does it work?" As a first consideration, it must be able to differentiate truth telling from deception reliably, and it should facilitate the disclosure of clinically relevant information. If it meets these requirements, it then needs to be demonstrated that in doing so it has a beneficial impact on treatment and/or management. But even if PCSOT does "work" in this way, if in the process it crosses ethical or legal red lines, then it would be hard to justify continued reliance on it.

Polygraph Testing

As indicated above, there are two primary outputs from a polygraph test, each of which complements the other.

The first, and what people usually associate with polygraph testing, is test outcome, that is, whether an examinee "passes" or "fails" the test. Although the focus is typically on "lie detection," determination of truthfulness is equally important. In order to shift attention away from the polygraph as a "lie detector," therefore, many practitioners now refer to it as a means of "credibility assessment" (Raskin, Honts, & Kircher, 2014). The fundamental questions here, of course, are how accurate polygraphy is in detecting deception and confirming honesty and whether that level of accuracy is sufficient for the setting in which it is being used. Unfortunately, this second question is often overlooked, an important oversight when translating research findings regarding accuracy into practice—what may not be accurate enough in a national security context or in a court of law may be sufficient for investigating crime or when used postconviction.

The second output of a polygraph test is disclosure. Numerous studies have reported that individuals report information during a polygraph examination they would otherwise have kept to themselves. Critics sometimes dismiss this effect as being a "bogus pipeline to the truth" as they say it depends on an examinee believing that the polygraph "works" and that disclosures would not occur if examinees did not hold this belief. This assertion, however, begs two questions: the extent to which disclosures are in fact dependent on a belief in the accuracy of the polygraph test and, if they are, the level of accuracy required to trigger this effect. As will be discussed later in this chapter, while many social psychology studies have demonstrated that disclosures do increase when subjects believe they are attached to a "lie

detector," the strength of this effect is unclear. A third more philosophical consideration also arises in respect of this issue—if disclosures are a function of a belief in polygraph accuracy, but polygraphy is shown to meet the level of accuracy required to trigger this belief, is it still correct to refer to the phenomenon as a "bogus" pipeline?

Thus, although discussions about PCSOT often get bogged down in arguments about accuracy levels and the basis of disclosures, both issues are more complex than they appear at face value.

What the Polygraph Records

That there is an association between deception and physiological activity has been known for centuries. One of the earliest and clearest expressions of this was by Daniel Defoe, who when writing about the prevention of street robberies in the eighteenth century observed that:

> Guilt carries fear always about with it; there is a tremor in the blood of the thief that, if attended to, would effectually discover him … take hold of his wrist and feel his pulse, there you shall find his guilt; a fluttering heart, an unequal pulse, a sudden palpitation shall evidently confess he is the man, in spite of a bold countenance or a false tongue. (Defoe, 1730/ quoted in Matte, 1996)

Fairly, though, Defoe also noted, "The experiment perhaps has not been try'd."

While the phrase "a tremor in the blood" is so often quoted by those who write about the history of the polygraph that it is in danger of becoming a cliché, it nonetheless lays the groundwork for both the basis of polygraph testing and some of the misconceptions associated with it.

The involuntary physiological responses associated with guilt and deception recognized by Defoe are now known to be caused by activity in the autonomic nervous system. These responses, however, are not unique to deception—lots of things can make the blood tremor besides guilt and lying, and no physiological variable has yet been discovered that is specific to deception. Because of this, it is sometimes concluded that polygraphy, or any other techniques that rely on recording and interpreting physiological activity, cannot possibly work. But there need not be a unique physiological lie response for polygraph testing to be effective; instead, what matters is whether physiological reactivity recorded *in the context of a polygraph examination* discriminates truth telling from deception at levels sufficiently above the chance to make the technique meaningful and worthwhile. False-positive and false-negative findings occur with every test and investigation; more relevant is being able to quantify their frequency and ensure that whatever actions follow a test result take this error rate into account.

A second misconception that can be seen in Defoe's observations is that physiological responses associated with deception are the result of emotion, especially the emotions of fear and anxiety. This mistake leads some to argue that anxious individuals, either inherently or because they are made anxious by the circumstances

of the test, are likely to wrongly "fail" for this reason. Other critics are concerned that in order for the test to work, polygraph examiners must induce anxiety or fear in examinees, which is ethically dubious (BPS, 1986; Vess, 2011). There is also a belief that psychopathic individuals, because of their low levels of anxiety and emotional responsiveness, are more likely to "beat" the test. But though there is uncertainty regarding the mode of action of polygraphy and the neuropsychological basis of the physiological reactions it records, it is clear that emotional reactivity is only part of the story and that a number of cognitive processes associated with deception contribute to what the polygraph observes. Anxiety and fear, except insofar as they indicate that the examinee takes the examination seriously, are likely to be minor components at best. More will be said about this later.

Cardiovascular, respiratory, and electrodermal activities measured by recording devices as opposed to being observed indirectly began to be used as a means of detecting deception in the late nineteenth and first part of the twentieth centuries, mainly on their own but in some cases together, both in Europe and the United States (Alder, 2007; Krapohl & Shaw, 2015). Criminologists, psychologists, and physicians such as Cesare Lombroso, Hugo Munsterberg, Georg Sticker, Vittorio Benussi, Walter Summers, William Marston, John Larson, and Leonarde Keeler researched and applied their various techniques, sometimes with phenomenal claims of success. In the 1930s this work coalesced into instruments that could simultaneously record data from the three physiological systems, giving rise to what became known as the polygraph. Although the hardware has improved since then, and the process has become digitalized so that ink pens writing on moving paper are no longer required, little has changed in terms of the basic physiology that is recorded.

In what way is activity in these physiological systems associated with deception? Traditionally polygraph examiners are taught that what they are observing is a "fight, flight, or freeze" response caused by the fear of being caught out in a lie and the consequences that follow, implicitly accepting an emotional basis to the test's mode of action. There are a number of major problems with this explanation; however, response characteristics that are associated with deception on the polygraph test are not identical to what is seen in a "fight, flight, or freeze" scenario, deceptive responses are recorded even where there is little anxiety and no consequence attached to being caught out (e.g., in tests where examinees are simply told to pick a number and then to lie when asked if they have done so), and not all polygraph formats require lying at all but instead relate to the "recognition" of relevant items.

The reality is that we are well short of understanding the mode of action of the polygraph, indicated by the number of theories proposed to explain it (National Research Council, 2003; Nelson, 2015), although it is now accepted that a range of mental processes are involved in addition to emotion. Important are concepts and factors such as the "differential salience" (i.e., differing degrees of importance or threat represented by the questions asked on the test), cognitive work involved in lying and in inhibiting truth telling (truth telling being the default position), autobiographical memory, orienting to "threat," and attention (Nelson, 2015; Senter, Weatherman, Krapohl, & Horvath, 2010), which interact to produce arousal in the autonomic nervous system that can be seen in a number of peripheral physiological processes.

While a lengthy discussion regarding the physiological and psychological mechanisms underlying polygraphy cannot be pursued here, the fundamental point is that conducting a successful polygraph test is about more than simply attaching the recording hardware and then asking questions. Instead, the examiner must work at ensuring that whatever reactions are recorded are produced because the examinee is deceptive to the questions being asked, rather than by other possible causes of autonomic arousal. This is achieved in a lengthy pretest interview and requires examiner training and skill, in other words, a competent examiner. Given that the process is so heavily dependent on the examiner's capabilities, it has been argued that polygraphy should not be seen as a "scientific test" (BPS, 2004), but this is perhaps more of a semantic than a practical issue—operator skill is important in all forms of scientific testing. But whether "scientific" or not, what matters is whether, in the hands of a competent examiner, polygraph testing can be shown to be a reliable means of distinguishing truth telling from deception.

In terms of PCSOT, there is no need to induce anxiety in examinees, anxious individuals are no more or less likely to "fail" the test, and, because the generation of fear or anxiety is irrelevant, psychopaths are no more or less likely to wrongly "pass" the test (Patrick & Iacono, 1989; Raskin & Hare, 1978). Furthermore, as will be discussed later, the examinee does not need to be deceived about the accuracy of polygraphy nor manipulated in other ways for the test to be successful.

Polygraph Accuracy

While the physiological targets of polygraph testing have not changed much since the 1930s, numerous testing techniques, question formats, scoring systems, and specialized applications have emerged since then, often introduced with little empirical support. The plethora of approaches and the associated lack of standardization have made it difficult to provide clear estimates of polygraph accuracy.

A number of initiatives have meant that the situation has improved (Krapohl & Shaw, 2015). Chart scoring, as opposed to decisions based on a global overview of the polygraph chart, was introduced in the 1960s, a hardening of testing protocols took place between the 1960s and 1990s, increased acceptance of blind scoring of charts as a means of quality control to overcome the risk of examiner bias became more commonplace in the 1990s, research in the early 2000s better clarified response patterns that are indicative of deception (and just as importantly, response patterns that aren't) and the amount of variance explained by the different physiological channels, and in the late 2000s, the American Polygraph Association undertook an exercise to validate testing techniques (American Polygraph Association, 2011). All of this has provided a better scientific basis on which to evaluate the efficacy of polygraph testing.

The most definitive review of polygraph accuracy to date has been carried out by the National Academies of Science in the United States. It concluded that "polygraph tests can discriminate lying from truth telling at rates well above chance,

though well below perfection" (National Research Council, 2003, p. 4). Accuracy for the most commonly used test format, the comparison question test (a version of which is employed in PCSOT), was estimated to be between 81 and 91 %, which is highly supportive of a meaningful association between what the polygraph records, truth telling, and deception.

The National Academies Review was carried out on behalf of the US Department of Energy, triggered by allegations of espionage at the Los Alamos nuclear weapons facility, and was designed to advise on the use of polygraph testing for personnel security vetting. Its overall conclusion was that an error rate of 10–20 % was too high for this type of application given the low levels of deception likely to be found in the population to be tested (one hopes that there are not many spies working in federal agencies) and the disproportionate number of false-positive findings such an error rate would imply. Although polygraph proponents disagree with this conclusion, arguing that it is based on a misconception of the way in which security vetting is undertaken because in this setting it acts as an initial screen rather than providing a definitive outcome, more important in terms of PCSOT is the review's observation that polygraphy becomes viable when the underlying rate of deception is over 10 %—a rate which most observers, even those critical of polygraphy, would accept is probably exceeded in sex offender populations.

For a number of reasons, however, the National Academies Review is not the end of the story, at least in terms of PCSOT. Its estimate of accuracy is based on single-issue, "diagnostic" tests, that is, tests in which a single known issue is being investigated, for example, whether an individual was involved in a bank robbery. Although this is sometimes the case in PCSOT, as when the focus is on specific behaviors reported to have occurred during an offense or where the matter of concern is whether the offender is responsible for a new crime, the majority of tests carried out in PCSOT are screening in nature. In screening tests, a number of behaviors are explored, but there is not a known event that underpins the thrust of the exam.

Screening tests are generally considered to be less accurate then single-issue tests, although there are insufficient trials from which to determine their precise level of accuracy. Screening tests however tend to have higher false-positive rates (tests which wrongly label an examinee as deceptive). Two studies used anonymous surveys with sex offenders in the United States to ask about their experiences of being wrongly accused of deception and also of instances where deception had been missed (Grubin & Madsen, 2006; Kokish, Levenson, & Blasingame, 2005). The findings were very similar, with responses from offenders in both studies suggesting an accuracy rate for PCSOT between 80 and 90 %, reassuringly similar to the National Academies estimate.

Because of its likely error rate, the utility of PCSOT tends to be emphasized rather than its accuracy, with disclosures seen as more important than test outcome. In addition, it is recommended that outcome in screening tests is reported as "significant response" or "no significant response" rather than "deception indicated" or "no deception indicated" as it is in single-issue tests. However, a more recent initiative has expressed polygraph test outcome as a probability statement with confidence intervals derived from data normed on large sets of confirmed tests.

Referred to as the "Empirical Scoring System" (ESS), this allows a better judgment to be made about the degree of confidence one can have in a given test result (Nelson et al., 2011). Although the database on which ESS is built could be larger, and while it still requires independent validation, this type of approach provides greater clarity on polygraph test accuracy in environments such as PCSOT.

The error rate associated with polygraphy, and its screening function in most PCSOT settings, means that it is probably a mistake to talk about an individual 'passing' or 'failing' the test. One doesn't pass or fail a screening exam of any sort. The aim of screening is to identify those who require further investigation. In the case of PCSOT, significant responses to some questions are observed, which might be thought of as 'screening positive', but this is different from failing a test. It is therefore probably more sensible to think in terms of positive and negative predictive values: the former referring to the likelihood of a true positive (that is, deception) when an individual shows a significant response, and the latter to the likelihood of truthfulness when no significant responses are recorded. It is usually the case that one is higher than the other, providing an indication of whether one should be more confident in deceptive or truthful calls (the first where it the positive predictive value is higher, the second when the negative predictive value is).

There remains the problem of examiner competence and its impact on test accuracy. However, if properly trained examiners use correct techniques that are administered properly, then their accuracy rate should be similar to that reported in the research literature. Ensuring that this is the case requires a well-constructed quality assurance and quality control program, which unfortunately many PCSOT programs lack. But this is a reason to improve programs rather than to dismiss polygraphy. Provided it is in place, the important question becomes not whether polygraph is "accurate" but whether accuracy in the range of 80–90 % is accurate enough.

The answer to this question will depend on how test outcome is used. An error rate of 10–20 % is clearly too high to warrant sending someone to prison or taking away their livelihood but not too high to inform decisions about treatment engagement, changes in monitoring conditions, or the need for further investigation into possible transgressions. This is particularly the case when one remembers that typically we make these types of decision based on our own determination of whether or not someone is deceptive, even though in experimental settings the ability of the average person to do so accurately is rarely above 60 % (Bond & DePaulo, 2006; Vrij, 2000).

Utility and Disclosure

Polygraphy is known to increase the likelihood that an examinee will disclose previously unknown information. There are many anecdotal accounts of this phenomenon in both investigative and screening settings, but the best evidence for this effect is found in sex offender testing where numerous studies describe significant increases in self-report of previous offense types and victims, deviant sexuality, and risky behaviors (e.g., Ahlmeyer, Heil, McKee, & English, 2000; Grubin, Madsen,

Parsons, Sosnowski, & Warberg, 2004; Heil, Ahlmeyer, & Simons, 2003; Hindman & Peters, 2001; Madsen, Parsons, & Grubin, 2004). This work, however, lacks robustness in that comparisons are usually made in terms of what was known about an offender before and after polygraph testing rather than with contemporaneous comparison groups in which polygraph testing is not used. As critics readily point out, this makes it difficult to disentangle the effects of polygraphy from other factors such as treatment impact or changes in supervision.

The lack of a comparison group with which to determine polygraph efficacy in facilitating disclosures has been addressed in two large UK studies, both of which confirmed the findings of earlier work that showed increases in disclosure when polygraphy is used. In one of these studies, polygraph testing was voluntary (Grubin, 2010), while in the other it was a mandatory condition of a parole license (Gannon et al., 2014; Gannon, Wood, Vasquez, & Fraser, 2012).

In the trial of voluntary testing (Grubin, 2010), the supervision of nearly 350 polygraphed offenders was compared with 180 sex offenders from probation areas where polygraphy was not used. Just over 40 % of eligible offenders agreed to be tested, of whom 47 % were tested more than once. The majority were taking part in treatment programs. Probation officers reported that new disclosures relevant to treatment or supervision were made in 70 % of the first tests, compared with 14 % of the non-polygraphed offenders making similar types of disclosure in the previous 6 months. A similar difference was found in respect of retests (only in this case the comparison was with 3 months before). The disclosures made by polygraphed offenders were rated as "medium" or "high" severity (the former relating to behaviors indicative of increased risk, the latter to actual breaches or offenses) in over 40 % of cases. The odds of a polygraphed offender making a disclosure relevant to his treatment or supervision were 14 times greater than they were for non-polygraphed offenders.

Although the test and comparison groups reported in Grubin (2010) did not differ on demographic or criminological variables, the fact that those tested were volunteers could have introduced bias. Because of this the mandatory trial described by Gannon et al. (Gannon et al., 2012, 2014) was considered necessary before a decision could be reached about implementing mandatory testing nationwide (it was a requirement set by the UK Parliament). Like the earlier study, a comparison group was used. Unlike it, the mandatory trial was limited to high-risk offenders (defined as those released on parole following a prison sentence of a year or more), and though many had undertaken sex offender treatment in prison, relatively few were involved in community treatment programs. The focus of the mandatory trial, therefore, was on the impact of polygraph testing on supervision only.

There were over 300 offenders in each group, which again did not differ on demographic variables. Although the mandatory trial involved an overall higher risk group and many fewer were in treatment than in the voluntary trial, its findings were similar. Significant increases were found in the number of individuals who made what were referred to as "clinically relevant disclosures" and in the number of disclosures these individuals made in the polygraph group. This was particularly noticeable in respect of sexual- and risk-related behaviors. However, the odds ratio of a disclosure being made was lower at 3.1.

In both studies significantly more actions were taken by probation officers who managed offenders subject to polygraphy than by probation officers supervising comparison offenders. One interesting finding reported in Gannon et al. (2012) was that while 73 % of interviewed probation officers believed the offenders they supervised were "open and honest" with them, this was the case for only 25 % of the probation officers who supervised polygraphed offenders. This is perhaps an explanation for the finding in Grubin (2010) that whereas probation officers of polygraphed offenders were more likely to increase risk ratings, risk ratings were more likely to be decreased in the comparison group.

Although the impact of polygraph testing on disclosures is clear, the question still remains whether it is simply a "bogus pipeline" effect. As described earlier, this refers to the increase in disclosures being the product of a belief that the polygraph "works," the implication being that disclosures would dry up in the absence of such a belief. As one critic commented in a newspaper article, it relies on offenders "not knowing how to use Google" (London Daily Telegraph, 2012).

A number of social psychology studies have demonstrated that subjects who believe they are attached to a "lie detector" appear to be more honest in their answers to questions regarding attitudes and behaviors, which has been interpreted as a reflection of social desirability or acquiescence biases (Jones & Sigall, 1971; Roese & Jamieson, 1993). But the effect is not in fact that great—a meta-analysis of 31 published reports found a mean effect size of $d=0.41$, which is in the small-to-moderate range (Roese & Jamieson, 1993).

Another factor to take into account when considering the "bogus pipeline" hypothesis is that all of the bogus pipeline studies are based on the use of a near 100 % lie detector. It is not clear from them what would happen if, rather than being sold as being 100 % accurate, the "lie detector" was instead said to have an accuracy rate "well above chance, though well below perfection" as described by the National Academies in respect of polygraph testing (National Research Council, 2003). In a yet unpublished research, our group found that a "lie detector" claiming to have a 75 % accuracy rate (i.e., a level below that attributed to polygraphy) appears to elicit disclosures with a frequency similar to that of a near 100 % accurate lie detector. This would seem to suggest that if part of the increase in disclosures brought about by polygraph testing is due to a belief in its lie detecting properties, then whatever else it may be the pipeline is not a bogus one.

Regardless of the merits and impact of the "bogus pipeline effect," the much more psychologically interesting question is what makes individuals disclose in this setting anyway, bogus pipeline or not. It may be that offenders disclose because they believe they will be, or have been, "caught out" by the polygraph, which would be consistent with research showing that one of the best predictors of whether a suspect will confess to a crime is the belief that there is good evidence against them (Gudjonsson, Sigurdsson, Bragason, Einarsson, & Valdimarsdottir, 2004). As indicated above, however, the "bogus pipeline effect" itself is unlikely to be the entire reason for increased disclosures, explaining only a small part of the variance. It could be that a polygraph test allows the offender an opportunity to change his account in a face-saving manner (after all, he was found out by a "lie

detector") or it may simply be that the dynamics of the interview itself are different from what takes place in normal supervision. Whatever the reason, the effect deserves increased research attention the words, and consideration given as to how to enhance it.

One further issue to address in respect of disclosures is whether the circumstances of a polygraph test result in offenders making false admissions in order to please polygraph examiners or to explain a "failed" test. Because many of the disclosures made in PCSOT are in any case difficult if not impossible to verify (e.g., how can one determine whether or not an offender has been masturbating to deviant fantasies?), it can be a challenge to confirm their veracity. What little research there is in relation to this suggests that false admissions occur but not often. Two studies have addressed this question using anonymous surveys with sex offenders in the United States who were asked whether they had ever made false disclosures in a polygraph test (Grubin & Madsen, 2006; Kokish et al., 2005). In both studies fewer than 10 % of offenders indicated that they had done so; in the Grubin and Madsen (2006) study, those who reported making false admissions had higher scores on the NEO neuroticism scale and lower scores on the conscientiousness scale, suggesting that those who make false admissions during a polygraph test may share characteristics with those who make false confessions in police interviews (Gudjonnson & Pearse, 2011; Gudjonsson et al., 2004). In any case, while the issue is not trivial, it does not seem to be a major problem.

Proponents of PCSOT argue that whatever the reason for increased disclosure by offenders who undergo polygraph tests, the effect is genuine and valuable. They ask whether critics are really suggesting that this information should not be sought or used because of concerns regarding the evidence base for the mechanisms that generate it. But resolution of this issue perhaps depends more on how PCSOT is implemented than on the academic arguments regarding polygraph itself.

The Implementation of PCSOT

The initial use of polygraph testing with sex offenders was as a clinical assessment to assist treatment providers in gaining fuller histories with which to inform treatment plans. The term "postconviction sex offender test" started to be used in the 1990s in reference to tests administered to individuals under court order, court supervision, or court-ordered treatment, with the intention of enhancing treatment or improving supervision (Holden, 2000). Its aim was to generate more complete information about an offender's history, sexual interests and functioning, and offending behavior based on disclosures and test outcome. This has been referred to as adding "incremental validity to treatment planning and risk management decisions" with which to improve decision-making (Colorado, 2011) and can perhaps be thought of more simply as "information gain."

In the late 1990s, the "Containment Model" was developed by practitioners in Colorado (English, 1998). It has since become the basis of many PCSOT programs

in the United States, although it has not taken root in the United Kingdom. The Containment Model refers to a triangle formed by a supervision officer, treatment provider, and polygraph examiner, although others may also be involved, in which the offender is "contained." It depends on good communication between agencies, with information obtained by one informing the actions of others.

While the Containment Model has clear attractions from a public protection perspective, it implies that all sex offenders require high levels of external control to keep them from reoffending. Compliance in the immediate term may be obtained, but whether it leads to longer-term change is uncertain. And though some offenders may require "containment," others genuinely seek to improve their internal controls and engage with treatment and supervision. In other words, there are some offenders who work with treatment providers and supervisors, and there are others who work against them. For the latter group, containment may be necessary, with the polygraph serving primarily as a lie detector to indicate when risk is increasing (related to this is a finding of Cook, Barkley, and Anderson (2014) that recidivism rates were higher in offenders who avoided or delayed their polygraph), but for the former group of offenders, polygraphy can act as a truth facilitator, encouraging them to discuss problematic thoughts and behaviors and providing reassurance that their risk is stable. It should be remembered that polygraphy not only detects lies, it also catches offenders telling the truth.

Whether or not following a strict containment approach, PCSOT has moved away from being an accessory of treatment to assume a more central role in offender supervision. It remains, however, the servant of those working directly with the offender, functioning to provide information about whatever is most relevant at the time. In this respect, different test types are relevant depending on the offender's circumstances.

Test Structure

Before describing the types of test used in PCSOT, the basic structure of a polygraph session needs to be described. The typical format employed in PCSOT is the "comparison question technique." It consists of three phases: a pretest interview, the examination itself, and a posttest interview:

The *pretest* is the longest part of the examination and can take from 1 to 2 h. Among other matters, information is collected about the examinee's background and current behavior, and the test questions are established and reviewed in detail. Many disclosures take place during this part of the process.

The *polygraph examination* consists of 10–12 questions, of which just 3 or 4 target the areas of concern and are referred to as the "relevant questions." Responses to the relevant questions are compared with so-called comparison questions to determine whether or not they are indicative of deception. More will be said about this shortly. Polygraph questions need to be simple, answerable with a yes or no, and relate to specific behavior rather than mental state, intention, or motivation.

In the *posttest* interview, the outcome of the exam is fed back, with the examinee given an opportunity to explain deceptive responses. In the UK study of voluntary testing, one third of disclosures were made during the posttest (Grubin, 2010).

As referred to above, in the comparison question technique, relevant questions are evaluated against comparison ones. If physiological responses to the former are greater than the latter, the examinee is judged to be deceptive; vice versa, the examinee is considered truthful. The comparison questions often take the form of a "*probable lie*," that is, questions that the examinee is unlikely to be able to answer truthfully. Examples of probable lies are "have you ever lied to a loved one?" and "have you ever stolen from a family member?" The theory is that truthful subjects will find these questions more concerning than the relevant ones because of their implications and thus show greater responses to them, while the deceptive examinee will be more responsive to the relevant questions because they represent more of a threat. The strength with which relevant questions exert a greater pull on the examinee than the comparison ones has been called "relevant issue gravity" (Ginton, 2009), which is a tidy way of packaging the various cognitive processes that determine autonomic arousal in response to polygraph questions.

The probable lie approach has been criticized on a number of grounds. First, the underlying theory that the differential response to the two question types relates to truthful individuals being more worried about what are in effect less serious comparison questions is frankly implausible (Ben-Shakhar, 2008; National Research Council, 2003). But given that the technique has been shown to be able to identify deception, this suggests that we need a new theory, not that the technique itself is faulty. Others are concerned that the probable lie approach means the test is based on deceiving the examinee and requires the examinee to be forced into a position of having to lie (Cross & Saxe, 2001; Meijer et al., 2008; Vess, 2011). This ethical objection, however, is based on a misconception—the cognitive work of the probable lie doesn't arise from the lie but from the uncertainty associated with the question. Indeed, comparison questions can take the form of a "*directed lie*" in which the examinee is instructed to lie to a question such as "have you ever made a mistake?," which involves neither manipulation nor dishonesty. More will be said about directed lies later in this chapter.

Test Types

There are four basic types of polygraph test used in PCSOT, some of which have variants to them (American Polygraph Association, 2009).

Sex History Exams

The purpose of this test is to obtain a fuller and more accurate account of an offender's sexual history, including the type and range of deviant behaviors in which he has engaged, the age at which they commenced, and his history of involvement in

unknown or unreported offenses. There are two forms of this exam, one that focuses on unreported victims of contact offenses and the other on sexually deviant behavior more generally and offenses that don't involve forces such as voyeurism or Internet-related offending. The rationale for the separation is that the more severe potential consequences associated with the former behaviors may contaminate responses to the latter. Prior to the polygraph exam, the offender completes a sex history questionnaire, usually as part of sex offender treatment. The questionnaire is the focus of the examination, but only selected questions are asked during the test itself.

The intention of the Sex History Exams is to develop a better understanding of risk and of treatment need. There can be a tendency, however, for examiners to dig for much more detail than is needed to achieve these aims, making the procedure an unrealistic exercise in recall for the offender as well as a potentially humiliating one; more information is not necessarily better information. In addition, because it is based on a lengthy questionnaire which covers behaviors that have taken place over many years, the risk of false-positive outcomes (i.e., wrongly "failing" the test) is increased. This is an important consideration given that about half of the American community and a third of residential sex offender treatment programs for adult males require the Sex History Exam to be passed in order for the treatment to be completed successfully.

A further problematic issue associated with Sex History Exams is what to do about self-incriminating disclosures. Programs typically try to get around this ensuring that only general information about past offenses is obtained, but in some states even this minimal level of disclosure needs to be passed to the authorities. In reality, however, this is not a difficulty unique to polygraph testing and applies to treatment programs generally. Whatever solution works for the program should be sufficient for PCSOT.

The following are two examples of how Sex History Exams can be helpful to treatment (these and subsequent examples are taken from the UK polygraph trials):

An offender on parole following a conviction for the indecent assault of his stepdaughter disclosed during a Sex History Examination a large amount of previously unknown pornography use and cross-dressing. Subsequent to the test, he began to discuss this and his sexual fantasies more generally in treatment for the first time.

An offender in his 50s with no sex-offending history was convicted of Internet-related offenses. In a Sex History Examination, he admitted to stealing underwear from his sister's house, to sexual fantasies regarding schoolgirls, and to sitting in cinema car parks to watch young girls. Based on this and other fantasy-related informations he disclosed during the test, new treatment targets regarding fantasy and fantasy modification were identified and delivered.

Critics argue that information from Sex History Exams tell us nothing new in that it would be a surprise if offenders hadn't engaged in deviant behaviors besides their offenses and that there is little evidence to show that the additional information adds meaningfully to risk assessment or treatment provision (Rosky, 2013). This criticism seems odd, however, given that sex history questions are asked routinely in sex offender assessment and are considered an important part of the evaluation;

the only difference being that there is more likelihood of getting an honest account during a polygraph examination.

Instant Offense Exam

This exam type explores behavior that took place during the instant offense where there is inconsistency between victim and offender accounts or where the offender denies important aspects of what took place. A variant of this test relates to prior allegations where there hasn't been a conviction. Like the Sex History Exam, this test is directly relevant to treatment. Also like the Sex History Exam, there is a risk that the examiner will go on a fishing exercise seeking detail that doesn't take treatment any further. Used properly, however, it can overcome denial that is blocking treatment progress.

Below is an example of how an Instant Offense Exam assisted treatment in a perhaps unexpected way:

An offender was on license having committed an indecent assault on a child in a supermarket when intoxicated. He admitted the offense but denied any memory of having pushed his groin into the girl's back as reported by her mother even though he accepted this could have happened. Much time was spent in the treatment group trying to overcome his "denial." On an Instant Offense Exam, he was questioned about his lack of recall, and he was found truthful. The consistency of his self-report taken together with the test result led to his account of partial amnesia being accepting, allowing treatment to move beyond this issue.

Some critics believe this sort of information would be obtained anyway in the course of treatment, but whether or not this is the case, supporters of PCSOT argue that the disclosures come much earlier when polygraphy is used. There is little evidence with which to determine either of these claims.

Offenders may see the Instant Offense Exam as an opportunity to prove their "innocence" in the face of a wrongful conviction. Although there may be a time and place for this issue to be explored, PCSOT is not it. The Instant Offense Exam, therefore, must be used with caution.

Maintenance Exam

The Maintenance Exam is the workhorse of PCSOT. It addresses an offender's compliance with the terms and conditions of probation, parole, or treatment. It is a screening test that typically covers a wide range of issues in the pretest, following which 3 or 4 specific questions are asked on the test itself. Maintenance Exams can address sexual thoughts and fantasies so long as they are linked to masturbatory behavior. The aims of the test are to identify behaviors indicative of increased risk so that interventions can take place, confirm when offenders are not engaging in problematic behavior, and deter offenders from engaging in risky behaviors in the

first place. Its primary purpose is to prevent reoffending rather than to detect reoffenses after they have occurred.

Two examples of Maintenance Exams illustrate their potential value:

An offender on parole license disclosed he had recently started a relationship with a young woman (one of his license conditions being that he informed his probation officer of any new relationships). Although that was the extent of his disclosure, his offender manager met with the new girlfriend and found not only that she was a single mother but also that the offender was grooming her child in a manner similar to his instant offense. He was recalled to prison.

An offender with a history of involvement with sex offender networks had a license condition not to associate with known sex offenders. Following a deceptive test, he admitted to breaching this condition. When his probation officer later explored this with him, he admitted to marked feelings of loneliness and isolation following a move from a probation hostel. Steps were taken to address his isolation, and on his next Maintenance Exam, he said he was no longer reliant on his former sex offender contacts and much more settled in himself; he showed no significant responses to questions relating to associating with other sex offenders.

In neither of these cases can it be demonstrated that offenses were prevented, but it would be hard to argue that the outcomes were not worthwhile.

A difficulty faced by Maintenance Exams is how to respond to a deceptive result in the absence of disclosures. Given the 10–20 % error rate of polygraph testing, it is hard to justify sanctions such as prison recall based on a failed test alone (although this does occur in some US states, it is prohibited in the United Kingdom), but a deceptive test does provide a warning sign that all may not be well. Depending on the risk represented by the offender, the response could range from the probation officer addressing the issue in supervision with him to not relaxing restrictions such as curfews or exclusion zones to, in especially high-risk cases, putting the offender under surveillance.

Maintenance Exams are carried out regularly, to set protocols—for example, in the United Kingdom, they take place at 6 monthly intervals, but sooner if the offender fails, tests or concerns emerge between exams. This gives rise to a risk of habituation or sensitization, resulting in fewer disclosures and false-negative test results (Branaman & Gallagher, 2005). To counter this PCSOT, policies usually recommend that a different examiner is introduced after a set number of tests have been undertaken. Again, however, research relating to this issue is sparse, and it is not clear the extent to which habituation occurs or whether the suggested remedy is effective.

Monitoring Exams

Monitoring Exams are specific issue tests that take place where there is concern that an offender may have committed a new offense or breached a license condition. As in Maintenance Exams, no sanction follows a failed test in the absence of disclosure, but a failed test may indicate the need for further investigation. On the other hand, a passed test can offer reassurance to supervisors.

The following is an example of how a Monitoring Exam can contribute to management:

A 24-year-old man was on parole having been convicted of unlawful sexual intercourse with a 14-year-old girl. His probation officer believed he was still in a sexual relationship with his victim, but this was persistently denied by the offender, who was compliant with a night-time curfew and a tag. He denied any wrongdoing during the pretest interview, but he was deceptive on the test. In the posttest interview, he admitted to regular contact with the girl as well as a low level of sexual activity with her. The probation officer passed this information to the police and the offender was arrested. When interviewed by the police, the girl reported regularly spending a night a week in the offender's home (a place his tag confirmed him to be), where in addition to the sexual activity he had described she said they also engaged in sexual intercourse.

Beating the Test

Somewhat incongruously, the same critics who argue that polygraphy does not reliably differentiate truth telling from deception nonetheless also invariably raise the issue of countermeasures, that is, physical or psychological techniques, used to manipulate responses on the test to enable examinees to appear truthful when they are being deceptive (Ben-Shakhar, 2008; London Daily Telegraph, 2012). They argue that false-negative findings, whether the result of error or countermeasures, mean that "dangerous" offenders can "beat" the test and remain free in the community.

It is almost certainly the case that some offenders "beat" the test, but the reality is that without polygraphy, many more "beat" their supervisors and treatment providers. For example, as referred to earlier, in the absence of polygraphy, probation officers are more likely to reduce their risk assessments than they are when polygraphy is used (Grubin, 2010). Decisions, however, should not be based on polygraphy alone—PCSOT is just one part of the information package.

It is also the case that countermeasure techniques exist and can be taught, and there are a number of websites that offer to do so. But in order to be successful, practice is required—theory is not sufficient—and the examinee needs feedback when attached to the polygraph (Honts, Hodes, & Raskin, 1985). Most sex offenders do not have access to this type of coaching, and without it their charts usually show tell-tale signs of their attempts to manipulate the test. It should also be remembered that polygraph examiners read the same websites as their examinees.

Treatment Benefit and Risk Reduction

Probation officers like PCSOT. In the English probation trials (Gannon et al., 2014; Grubin, 2010), over 90 % rated polygraphy as being "somewhat" or "very" helpful, with very few tests considered by officers to have had either no or a negative impact.

But while subjectively probation officers may believe polygraphy makes their jobs easier, this is not the same as being able to demonstrate objectively that PCSOT results in improved treatment outcome or a genuine reduction in risk (Rosky, 2013).

Evidence regarding reduction in recidivism is extremely thin, although the absence of evidence should not be confused with evidence of absence. It is difficult to carry out randomized control trials of PCSOT for a range of reasons, not the least of which is a reluctance by criminal justice agencies to "experiment" with dangerous sex offenders. Furthermore, the low levels of recidivism that make treatment programs difficult to evaluate create similar problems for PCSOT, although significant increases in prison recall for breaches have been demonstrated (Gannon et al., 2014; Grubin, 2010).

Two early studies, although not of PCSOT per se, point in the right direction. Abrams and Ogard (1986) compared recidivism rates of 35 probationers (few of whom were sex offenders) from two counties in Oregon required to take periodic polygraph tests, with 243 offenders from a county where supervision did not involve polygraphy. Over 2 years, 31 % of the polygraphed men committed an offense or infringement compared with 74 % of those who were not polygraphed. But the number of polygraphed offenders was small; the samples were not matched nor is it clear whether there was selection bias in choosing those who underwent polygraphy. Also in Oregon, Edson (1991) reported that 95 % of 173 sex offenders on parole or probation who were required to undertake periodic polygraph testing did not reoffend over 9 years, but there was no comparison group in this study at all.

McGrath, Cumming, Hoke, and Bonn-Miller (2007) carried out the one randomized trial of PCSOT in the literature, comparing 104 sex offenders in Vermont who received treatment in programs that included PCSOT with 104 matched offenders in programs where polygraphy was not used. At 5-year follow-up, they found no difference in sex offense recidivism rates, but they did find a significantly lower rate of reconviction for nonsexual violent offenses. But though the study was well designed, its results are difficult to interpret because while the research was sound, the way in which PCSOT was delivered was not. Offenders undertook polygraph examinations on average just once every 22 months, dissipating the likelihood that polygraphy would have much of an impact on behavior. Even so, the reduction in violent offending is notable.

In trying to determine the impact of PCSOT, there is another issue to consider. It is well established in relation to sex offender interventions generally that to be effective, they should adhere to the "risk-need-responsivity" principle—that is, they should target high-risk individuals, reflect treatment need, and be responsive to cognitive and cultural differences between offenders (Andrews, Bonta, & Wormith, 2011). PCSOT does not tend to be delivered in this way because it is an assessment procedure rather than an intervention as such. After all, a screening technique for a medical condition is not judged on the basis of whether it improves survival rates for that condition—that is, the role of what follows—but on its success in identifying at-risk individuals. Expecting PCSOT to reduce recidivism may be an unreal expectation.

So how then is PCSOT to be judged? Rather than focus on recidivism perhaps, attention should be focused instead on the value of the information gained as one

would in an evaluation of screening instruments generally. The frequency and content of disclosures, the impact of test outcome on decision-making, and actions taken after a polygraph test could all form part of a cost-value analysis to determine the value added by PCSOT compared with the cost of administering it. In other words, to what extent does PCSOT better enable probation officers to monitor risk and initiate timely interventions, and are treatment targets better identified, when polygraph is used? The question then becomes, "is PCSOT worth it?"

Internet Offenders

Men who download indecent images of children from the Internet present a particular challenge for those carrying out risk assessments. Typically, little is known about relevant risk factors and they often have no criminal history. It is estimated, however, that around 50 % of men convicted of Internet offenses have committed undetected sexual assaults on children, and the majority show pedophilic sexual arousal patterns (Seto, 2013). It has been suggested that applying PCSOT techniques in a preconviction setting to men arrested for downloading offenses could assist in differentiating low- from high-risk offenders (where risk relates to contact offending against children), enabling police resources to be better focused and criminal justice interventions to be more accurately targeted in terms of custody and treatment. That this can be done was demonstrated in a small study in which 31 apparently low-risk Internet offenders underwent sex history-type polygraph examinations preconviction, where it was found that only 8 (26 %) could be confirmed as genuinely low risk (Grubin, Joyce, & Holden, 2014). A number of police forces in England are now exploring this application of polygraphy further.

Legal Considerations

The legal situation in the United Kingdom is more straightforward than it is in the United States. In the United Kingdom, the Offender Management Act 2007 sets out the statutory position regarding the mandatory testing of sex offenders on parole (Offender Management Act, 2007). Offenders must have been sentenced to a year or more in prison in order to ensure that the polygraph condition is proportionate. The legislation prohibits the use of evidence from polygraph tests in criminal proceedings, although this information can form the basis of criminal investigation, and it can also be used in civil proceedings. The act is supported by a statutory instrument containing polygraph "rules" which govern the conduct of polygraph sessions and set out the requirements that must be met by examiners. The 2007 legislation mandated a time-limited period to allow mandatory polygraph testing to be evaluated on a pilot basis in a small number of probation regions, after which the Secretary of State for Justice was required to return to Parliament for approval to extend

mandatory testing nationwide. Following the successful evaluation of the pilot (Gannon et al., 2012), Parliamentary approval was granted in 2013, and mandatory testing throughout England and Wales became effective in January 2014.

Although the Offender Management Act 2007 prohibits the use of the results of mandatory testing in criminal proceedings, there is no legislation that prevents polygraph testing in general from being used as evidence in the British courts. It is sometimes claimed that the law prevents the use of polygraph evidence, but this is not true (Stockdale & Grubin, 2012). Whether polygraphy evidence should be allowed in criminal proceedings is a too complicated issue to be explored here, apart from observing that while polygraphy can be a valuable investigative tool, it is not clear that it can add much to the decision-making process in court.

The position regarding PCSOT in North America is more haphazard. The main issue for the courts has been whether PCSOT breaches the Fifth Amendment rights against self-incrimination. In considering this question, the Supreme Court ruled in McKune v. Lile that it does not, albeit in a tight 5-to-4 decision. It observed that the treatment program of which it was part served "a vital penological purpose." On the other hand, in the United States v. Antelope (2005), the Federal 9th Circuit Appeal Court ruled that a paroled offender could not be compelled to waive his Fifth Amendment rights and take a polygraph exam with the threat of prison recall if he did not. This has made it even more necessary for programs to ensure that they properly address the self-incrimination issue, both in terms of PCSOT and more generally.

PCSOT is hardly used in Canada (McGrath et al., 2010), and it therefore does not appear to have been an issue for the Canadian courts, apart from one case where a prisoner applied for judicial review of a Parole Board decision not to release him partly on the basis that the decision was made before he had undertaken a polygraph examination—in this case the Court decided that the polygraph test results would not have changed anything in the Parole Board's decision (Aney v. Canada, 2005). In general, however, the Canadian Courts allow polygraph disclosures to be used in criminal proceedings so long as the jury is not told that they came from a polygraph test.

Ethics

Commentators rightly distinguish between practice standards and ethical principles, observing that the two do not necessarily overlap (Chaffin, 2011). Even where the delivery of PCSOT is well managed and delivered, potential ethical objections don't go away. When discussing PCSOT, a number of ethical issues are frequently raised. These tend to relate to a lack of respect for autonomy, intrusiveness, and compulsion, as well as special considerations that arise when testing special groups such as adolescents, the intellectually disabled, and individuals with mental disorder.

Some of these objections relate to a misconception of what happens in PCSOT, others to its questionable implementation. For example, Cross and Saxe (2001) refer to PCSOT as "psychological manipulation" on the basis that examiners deceive

offenders by telling them that the polygraph is error-free. While this may occur, it is certainly not good practice nor is there any reason for examiners to make out that the test is any more accurate than it actually is. Indeed, the British Psychological Society (2004) observes that participants should be informed of known error rates, a sentiment with which it is hard to disagree. There is no reason to believe that PCSOT would cease to be effective in these circumstances.

Cross and Saxe (2001), Meijer et al. (2008), and Vess (2011) all argue that the test itself is based on deception when the probable lie technique is used given the hypocrisy involved in demanding the offender to be honest. Vess (2011) and McGrath et al. (2010) wonder in addition what damage this might do to the therapeutic relationship. But as indicated earlier, the probable lie technique is not in fact dependent on the examinee lying even though this is what tends to be taught (indeed, as referred to above, other critics refer to this theory being deficient), but on uncertainty. Regardless, the use of "directed lies" overcomes this objection and also avoids the risk of the examinee admitting to transgressions that have nothing to do with his sexual risk.

Chaffin (2011), although concerned mainly with the testing of adolescents, focuses on PCSOT "extracting confessions" from examinees, stating "The polygraph is fundamentally a coercive interrogation tool for extracting involuntary confessions" (p. 320). PCSOT, however, need not, and should not, involve interrogation. It is instead an interview process in which lying is explicitly discouraged. The questions asked during PCSOT are asked by assessors and treatment providers anyway—the fact that PCSOT encourages disclosure of information relevant to treatment and risk management is in itself not an ethical issue.

Mandatory PCSOT is of course coercive in that there are penalties for noncooperation. But PCSOT examinees are convicted offenders, who by virtue of their criminal convictions are required to accept a range of restrictive and coercive measures such as conditions on where they live, limitations on employment, curfews, and treatment requirements. Indeed, the European Court of Human Rights has ruled that penile plethysmography (a technique in which penile arousal in response to sexual stimuli is measured and recorded) can be made a compulsory part of sex offender treatment on the grounds of public safety (Gazan, 2002); one might think this is considerably more "invasive" than polygraphy. Provided that the questions asked during the polygraph test are directly relevant to treatment or supervision, the process does not seem any more coercive then these other measures or any more morally problematic.

Another objection to PCSOT is that it carries with it the implication that sex offenders are not to be trusted and that this itself damages the relationship between supervisors and offenders. There is no evidence, however, that this is the case, while what evidence there is suggests it does not (Grubin, 2010). Indeed, this implication is often implicit in any case. One should not underestimate the benefits of an offender being able to demonstrate that he is being truthful in his dealings with those supervising him and the positive impact this can have on the therapeutic relationship.

There remains the question, however, of special groups. About half of adolescent treatment programs in the United States, for example, incorporate PCSOT (McGrath et al., 2010), and the American Polygraph Association PCSOT model policy allows

for testing juveniles down to the age of 12. As Chaffin (2011) points out, given the increased vulnerability of juveniles and adolescents to coercion and suggestion, and differences in the way that risk, treatment, and rehabilitation are conceptualized in this group, one can't assume that PCSOT approaches are appropriate for them. He could have added that it is not even clear that polygraphy itself works in the same way as it does in adults given differences in brain maturity and psychological development and that the American Polygraph Association age threshold appears arbitrary. Because of these and similar issues, mandatory polygraph testing in the United Kingdom does not apply to offenders who are under the age of 18.

Does this mean that polygraph testing of those under 18 is unethical? Testing offenders younger than 18 has its advocates (Jensen, Shafer, Roby, & Roby, 2015). Even Chaffin (2011), who considers the ethical concerns to be "substantial," doesn't go that far, although his view is contingent on the ability of those supporting its use in this group to prove that it provides more benefit than harm. Unless and until this evidence is produced, however, it probably makes sense to use PCSOT with great caution with those under 18, with decisions made on a consideration of individual cases rather than based on a blanket policy of PCSOT for all.

In terms of other special groups, such as those with intellectual disability and mental disorder, the position is similar. PCSOT has the potential to be of benefit, but caution needs to be used, by examiners who are aware of the pitfalls.

Finally, one might ask whether it is unethical *not* to use PCSOT in the treatment and supervision of sex offenders. If the information obtained during polygraph examination adds significantly to what is otherwise known about treatment need and risk, is it right to deny the potential benefits of PCSOT to an offender? When asked, many offenders themselves reported that they find polygraph testing to be helpful (Grubin & Madsen, 2006; Kokish et al., 2005). If PCSOT does reduce risk, how can one explain to a future victim why it did not form part of the offender's treatment and supervision package?

Conclusion

Does PCSOT increase community safety? Does it enhance sex offender treatment? Although the evidence is supportive, the benefits of PCSOT have yet to be conclusively demonstrated. Objections made by many of its critics, however, are based on opinion rather than fact. But what would count as definitive evidence? For ideological reasons, some will never be convinced.

Given the complexity of sex offender management, simply collecting data on numbers of disclosures, reconvictions, and the like will tell us little more than we already know. More thought needs to be directed to which offenders are most likely to benefit, the needs that PCSOT should target in those offenders, and whether modifications are necessary depending on the characteristics of the individual taking part. In other words, consideration should be given to how the "risk-need-responsivity" principle can be made to apply to PCSOT.

In the meantime, those who deliver PCSOT need to ensure that examiners are properly trained and supervised, protocols for the process are sound, and good quality control procedures are in place. In turn, those who make use of it must know the right questions to ask of it, how much weight to give its results, and how to integrate it with everything else they do with an offender. It should not be forgotten, however, that PCSOT remains just one tool in the box, and like any tool if it is not used with care it can cause harm.

References

Abrams, S., & Ogard, E. (1986). Polygraph surveillance of probationers. *Polygraph, 15*, 174–182.

Ahlmeyer, S., Heil, P., McKee, B., & English, K. (2000). The impact of polygraphy on admissions of victims and offenses in adult sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 12*, 123–139.

Alder, K. (2007). *The lie detectors: The history of an American obsession*. New York, NY: Free Press.

American Polygraph Association. (2009). Model policy for post-conviction sex offender testing. Retrieved from http://www.polygraph.org/files/model_policy_for_post-conviction_sex_offender_testing_final.pdf.

American Polygraph Association. (2011). Committee report on validated techniques. *Polygraph, 40*, 194–305.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2011). The risk-need-responsivity (RNR) model: Does adding the good lives model contribute to effective crime prevention? *Criminal Justice and Behavior, 38*, 735–755.

Aney v. Canada. (2005). 2005 FC 182.

Ben-Shakhar, G. (2008). The case against the use of polygraph examinations to monitor post-conviction sex offenders. *Legal and Criminological Psychology, 13*, 191–207.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 213–234.

Branaman, T. F., & Gallagher, S. N. (2005). Polygraph testing in sex offender treatment: A review of limitations. *American Journal of Forensic Psychology, 23*, 45–64.

British Psychological Society. (1986). Report of the working group on the use of the polygraph in criminal investigation and personnel screening. *Bulletin of the British Psychological Society, 39*, 81–94.

British Psychological Society. (2004). A review of the current scientific status and fields of application of polygraphic deception detection. Report (26/05/04) from the BPS Working Party. Retrieved from http://www.bps.org.uk/sites/default/files/documents/polygraphic_deception_detection_-_a_review_of_the_current_scientific_status_and_fields_of_application.pdf.

Chaffin, M. (2011). The case of Juvenile polygraphy as a clinical ethics dilemma. *Sexual Abuse: A Journal of Research and Treatment, 23*, 314–328.

Colorado. (2011). *Colorado Sex Offender Management Board: Standards and guidelines for the assessment, evaluation, treatment and behavioral monitoring of adult sex offenders.* Retrieved from https://cdpsdocs.state.co.us/somb/ADULT/FINAL_2012_Adult_Standards_120712.pdf.

Cook, R., Barkley, W., & Anderson, P. B. (2014). The sexual history polygraph examination and its influences on recidivism. *Journal of Social Change, 6*, 1–10.

Cross, T., & Saxe, L. (2001). Polygraph testing and sexual abuse: The lure of the magic lasso. *Child Maltreatment, 6*, 195–206.

Edson, C. F. (1991). *Sex offender treatment*. Jackson County, OR: Department of Corrections.

English, K. (1998). The containment approach: An aggressive strategy for the community management of adult sex offenders. *Psychology, Public Policy, & Law. Special Issues: Sex Offenders: Scientific, Legal, and Public Policy Perspectives, 4*, 218–235.

Gannon, T. A., Wood, J. L., Pina, A., Tyler, N., Barnoux, M. F., & Vasquez, E. A. (2014). An evaluation of mandatory polygraph testing for sexual offenders in the United Kingdom. *Sexual Abuse: A Journal of Research and Treatment, 26*, 178–203.

Gannon, T. A., Wood, J. L., Vasquez, E. A., Fraser, I. (2012). *The evaluation of the mandatory polygraph pilot (Ministry of Justice Research Series 14/12).* Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/217436/evaluation-of-mandatory-polygraph-pilot.pdf.

Gazan, F. (2002). Penile plethysmography before the European Court of Human Rights. *Sexual Abuse: A Journal of Research and Treatment, 14*, 89–93.

Ginton, A. (2009). Relevant issue gravity (RIG) strength – A new concept in PDD that reframes the notion of psychological set and the role of attention in CQT polygraph examinations. *Polygraph, 38*, 204–217.

Grubin, D. (2008). The case for polygraph testing of sex offenders. *Legal and Criminological Psychology, 13*, 177–189.

Grubin, D. (2010). A trial of voluntary polygraphy testing in 10 English probation areas. *Sexual Abuse, 22*, 266–278.

Grubin, D., Joyce, A., & Holden, E. J. (2014). Polygraph testing of 'low risk' offenders arrested for downloading indecent images of children. *Sexual Offender Treatment, 9*, 1–10.

Grubin, D., & Madsen, L. (2006). The accuracy and utility of post conviction polygraph testing with sex offenders. *British Journal of Psychiatry, 188*, 479–483.

Grubin, D., Madsen, L., Parsons, S., Sosnowski, D., & Warberg, B. (2004). A prospective study of the impact of polygraphy on high-risk behaviors in adult sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 16*, 209–222.

Gudjonnson, G. H., & Pearse, J. (2011). Suspect interviews and false confessions. *Current Directions in Psychological Sciences, 20*, 33–37.

Gudjonsson, G. H., Sigurdsson, J. F., Bragason, O. O., Einarsson, E., & Valdimarsdottir, E. V. (2004). Confessions and denials and the relationship with personality. *Legal and Criminological Psychology, 9*, 121–133.

Heil, P., Ahlmeyer, S., & Simons, D. (2003). Crossover sexual offenses. *Sexual Abuse: A Journal of Research and Treatment, 15*, 221–236.

Hindman, J., & Peters, J. (2001). Polygraph testing leads to better understanding adult and juvenile sex offenders. *Federal Probation, 65*, 8–15.

Holden, E. J. (2000). Pre and post-conviction polygraph: Building blocks for the future – Procedures, principles, and practices. *Polygraph, 29*, 69–92.

Honts, C. R., Hodes, R. L., & Raskin, D. C. (1985). Effects of physical countermeasures on the physiological detection of deception. *Journal of Applied Psychology, 70*, 177–187.

Iacono, W. G. (2008). Effective policing understanding how polygraph tests work and are used. *Criminal Justice and Behavior, 35*, 1295–1308.

Jensen, T. M., Shafer, K., Roby, C., & Roby, J. L. (2015). Sexual history disclosure polygraph outcomes: Do juvenile and adult sex offenders differ? *Journal of Interpersonal Violence, 30*, 928–944.

Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin, 76*, 349–364.

Kokish, R., Levenson, J., & Blasingame, G. (2005). Post-conviction sex offender polygraph examination: Client-reported perceptions of utility and accuracy. *Sexual Abuse: A Journal of Research and Treatment, 17*, 211–221.

Krapohl, D. J., & Shaw, P. K. (2015). *Fundamentals of polygraph practice*. San Diego, CA: Elsevier.

Levenson, J. S. (2009). Sex offender polygraph examination: An evidence-based case management tool for social workers. *Journal of Evidence-Based Social Work, 6*, 361–375.

London Daily Telegraph. (2012). *The Awkward Truth About Lie Detectors.* Retrieved from http://blogs.telegraph.co.uk/news/tomchiversscience/100173508/the-awkward-truth-about-lie-detectors.

Madsen, L., Parsons, S., & Grubin, D. (2004). A preliminary study of the contribution of periodic polygraph testing to the treatment and supervision of sex offenders. *British Journal of Forensic Psychiatry and Psychology, 15*, 682–695.

Matte, J. A. (1996). *Forensic psychophysiology using the polygraph: Scientific truth verification, lie detection*. New York, NY: J.A.M. Publications.

McGrath, R. J., Cumming, G. F., Burchard, B. L., Zeoli, S., & Ellerby, L. (2010). *Current practices and emerging trends in sexual abuser management: The Safer Society 2009 North American Survey*. Brandon, VT: Safer Society Press.

McGrath, R. J., Cumming, G. F., Hoke, S. E., & Bonn-Miller, M. O. (2007). Outcomes in a community sex offender treatment program: A comparison between polygraphed and matched non-polygraphed offenders. *Sexual Abuse: A Journal of Research and Treatment, 19*, 381–393.

McKune v. Lile. (2002). 536 U.S. 24.

Meijer, E. H., Verschuere, B., Merckelbach, H., & Crombez, G. (2008). Sex offender management using the polygraph: A critical review. *International Journal of Law and Psychiatry, 31*, 423–429.

National Research Council. (2003). *The polygraph and lie detection. Committee to Review the Scientific Evidence on the Polygraph*. Washington, DC: The National Academic Press.

Nelson, R. (2015). Scientific basis of polygraph testing. *Polygraph, 44*, 28–61.

Nelson, R., Handler, M., Shaw, P., Gougler, M., Blalock, B., Russell, C., … Oelrich, M. (2011). Using the empirical scoring system. *Polygraph,* 40:67–78.

Offender Management Act. (2007). Retrieved from http://www.legislation.gov.uk/ukpga/2007/21/pdfs/ukpga_20070021_en.pdf.

Patrick, C. J., & Iacono, W. G. (1989). Psychopathy, threat and polygraph test accuracy. *Journal of Applied Psychology, 74*, 347–355.

Raskin, D. C., & Hare, R. D. (1978). Psychopathy and detection of deception in a prison population. *Psychophysiology, 15*, 126–136.

Raskin, D. C., Honts, C. R., & Kircher, J. C. (Eds.). (2014). *Credibility assessment: Scientific research and applications*. San Diego, CA: Elsevier.

Roese, N. J., & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. *Psychological Bulletin, 114*, 363–381.

Rosky, J. W. (2013). The (F)utility of post-conviction polygraph testing. *Sexual Abuse: A Journal of Research and Treatment, 25*, 259–281.

Senter, S., Weatherman, D., Krapohl, D., & Horvath, F. (2010). Psychological set and differential salience: A proposal for reconciling theory and terminology in polygraph testing. *Polygraph, 39*, 109–117.

Seto, M. (2013). *Internet sex offenders*. Washington, DC: American Psychological Association.

Stockdale, M., & Grubin, D. (2012). The admissibility of polygraph evidence in English criminal proceedings. *Journal of Criminal Law, 76*, 232–253.

United States v. Antelope (2005). 395 F.3d 1128.

Vess, J. (2011). Ethical practice in sex offender assessment: Consideration of actuarial and polygraph methods. *Sexual Abuse: A Journal of Research and Treatment, 23*, 381–396.

Vrij, A. (2000). *Detecting lies and deceit*. Chichester: Wiley.