

---

# Forensic Psychotherapy for Sexual Offenders: Has Its Effectiveness Yet Been Demonstrated?

Harry M. Hoberman

---

## Introduction

To what degree has psychotherapy been empirically demonstrated to result in the prevention of future acts of sexual offending? That is, what scientific evidence exists that demonstrates that psychosocial interventions with sexual offenders consistently and effectively lead to enduring reduced rates of future sexual offending? Psychotherapy is generally conceived of as a process through which clients attempt to “change” problematic or maladaptive aspects of themselves through interactions with clinicians—persons with particular qualifications (e.g., training and experience). Such psychosocial interventions broadly involve clinicians providing various means of providing support, understanding, and “influence” so that help-seeking persons achieve some desired outcome by promoting self-understanding and/or exposing them to experiences and methods for changing their behaviors, thoughts, attitudes, and emotions. Consequently, it seems conceivable that psychotherapy has the potential to play some role in the management of sexual offenders, at least theoretically, by facilitating personal change in factors presumably related to the initiation and/or maintenance of sexual offending. As with other persons with significant behavioral problems, it is plausible that mental health professionals might be effective to some demonstrated degree in providing varied but identified means for sexual offenders to “change” in ways that their propensity for sexual behavioral problems is eliminated or substantially reduced. From a criminological perspective, it has been argued that potential evidence indicating general criminal recidivism can be reduced short term by some psychotherapeutic interventions suggests that sexual offenders too might respond to such interventions. However, while both conceiv-

able and plausible, ultimately it is an *empirical* question as to whether psychosocial interventions have been or can be proven on the basis of scientific study to affect change in sexual offenders or decrease the risk of sexual offense recidivism. The results of that empirical question also have critical and important policy implications. It has significant implications if psychotherapy has not or cannot be demonstrated to be an effective mechanism of personal change for sexual offenders. First, both therapists and society may demonstrate a false sense of security that once a sexual offender has been involved in a treatment program that their risk of sexual reoffending has been reduced. As a recent treatment review stated:

As a matter of social justice for the offend, and to provide reassurance to the community, it is essential that the treatments provided work and therefore inspire confidence that offenders who have completed treatment programs really are at reduced risk of sexual reoffending. (Dennis et al., 2012, p. 7)

Without the type of “proof” that is expected for other psychosocial interventions, then mental health professionals cannot make claims that their efforts at such treatment of sexual offenders matter. Further, in the absence of scientific demonstrations of psychotherapy effectiveness with sexual offenders, it becomes reasonable and necessary for alternative management approaches to be employed with sexual offenders as a means of preventing or reducing the risk of future sexual offenses.

For mental health professionals (MHP), generally, there is an assumption or belief that psychotherapy should and does make a difference to their clients; there is a strong expectancy effect that psychosocial interventions should be effective. In particular, for MHP providing psychosocial interventions for sexual offenders, it appears that clinicians have very strong beliefs that sexual offender treatment can and does have a powerful and enduring effect on their “clients” (e.g., Fortney, Baker, & Levenson, 2009). However, such beliefs should obviously not be assumed or taken on faith. For example, as scientific evidence appeared, the sexual offender field shifted from a reliance on unstructured clinical judgment (clinical intuition) to a reliance on

---

H.M. Hoberman, Ph.D., L.P. (✉)  
Forensic Psychology Evaluations, Professional Consultations,  
641 E. Lake Street, Ste 222, Wayzata, MN 55391, USA  
e-mail: [psycholegalevals@aol.com](mailto:psycholegalevals@aol.com)

empirically validated risk assessment approaches. Similarly, as with any intervention for a serious behavioral problem, the effectiveness of psychotherapy for sexual offenders should be clearly and consistently demonstrated by scientific study. Olver, Stockdale, and Wormith (2011) emphasized that the central purpose of offender treatment programs is reducing the recurrence of future criminal acts; thus, the primary purpose of sexual offender treatment is to affect a reduction in sexual offense recidivism. As Hanson et al. (2002) wrote: "If treatment is to be widely used in the management of sex offenders, then it is important that it works." As the Association for the Treatment of Sexual Abusers (ATSA, 2008, p. 1) has noted, treatment for sexual offenders is significantly different than that for other clients because of "a focus on the harm causes to the victims, the protection of future victims and the prevention of re-victimization" (emphasis added). For many psychological conditions or disorders, the evidence for the *efficacy* of psychotherapy is much more limited than often assumed or believed. While psychosocial treatments have been demonstrated to affect positive outcomes for some "emotionally distressed" persons, there is far less evidence that they have substantive short- or long-term effects for those with multiple or more severe "behavioral problems" that have led to significant impairments or distress. There is also evidence that beyond being ineffective, some psychosocial interventions may actually create "harm" for clients and society (e.g., Lillienfeld, 2007; Arkowitz & Lilienfeld, 2006). Thus, for a number of important reasons, MHP who provide psychotherapy for sexual offenders should have an accurate view regarding the empirical evidence that exists regarding the degree of effectiveness of the services they provide to sexual offenders. Other stakeholders concerned about sexual offending should also possess accurate knowledge about whether scientific evidence exists that sexual offender treatment might be effective.

Some may assume that the sexual offender who ends up as a participant in sexual offender treatment is the only client in that process. However, unlike other presenting problems, for sexual offending, most commonly there are several other "clients" invested in the importance of effective management of that problem. In addition to the sexual offender, the community is also a primary client of sexual offender treatment because ineffective treatment can directly create significant risks and consequences for public safety; unsuccessful sexual offender treatment may lead to additional future sexual offense victims. As Berliner wrote in 2002: "The big difference for sex offender treatment is that the price of failure is the victimization of an innocent person rather than continued suffering by the client" (p. 196). Similarly, Hall (1995) wrote: "...the expectation of psychological treatments for sexual offenders is no recidivism because of the serious effect of even a single act of sexually aggressive behavior.

Every act of sexual aggression adversely affects a person other than the perpetrator..." (p. 802). Moreover, the degree and relative persistence of harm that results from various forms of sexual offending can be profound; as Marshall et al. (2003) stated, such effects can be "devastating." This is a relatively unique state of affairs relative to almost all other types of mental health presenting problems where the consequence of ineffective treatment is primarily borne by the individual with the problem, leading to personal distress or impairment. Forensic psychotherapy is typically defined as that which pertains to "justice-involved clients," where the clinical outcomes of interest are most often focused on rule-breaking conduct and the determination of treatment benefit is most focused on a relatively specific outcome, namely, future criminality or reoffending (Mitchell, Simourd, & Tafrate, 2014).

Thus, almost all psychotherapy with sexual offenders would necessarily be regarded as "forensic psychotherapy," commonly understood as the psychological treatment of persons who have committed violent or aggressive offenses against others or themselves, who are often ordered into a therapeutic setting by the legal system, and who have particular sets of psychological and/or psychiatric characteristics that, to a large degree, define their criminogenic and treatment needs.

Further, to the extent that psychotherapy for a sexual offender is either required or funded by someone other than the offender, those sources are also a client. Certainly, public (government) and private (e.g., health insurance) funding sources of mental health programs are increasingly focused on empirically demonstrated effective interventions or evidence-based practice as a bottom line criterion for committing resources to publicly or third-party-funded interventions. Particularly, in times of limited economic resources, it seems unlikely that funding will be provided for treatment programs unless there is clear evidence of substantive effectiveness. Similarly, in addition to the primary victims of subsequent acts of sexual offending, other persons ("secondary victims") and entities bear the personal and financial responsibility of providing support and short- and long-term care and services for the effects experienced by the primary victims. In addition, almost all sexual offender treatment in North America is either explicitly or implicitly mandated by the criminal justice system; thus, again, it would be most appropriate to refer to sexual offender treatment as a form of forensic psychotherapy. Clearly, it is essential that all stakeholders relative to the management of identified sexual offenders—those persons involved in managing sexual offenders (particularly those who provide psychotherapy), those who are existing or potential direct victims or affected parties of sexual offending, and those entities that provide the funding for sexual offender treatment—have an accurate understanding of the nature and effectiveness of the existing

scientific literature on psychosocial treatment of sexual offenders.

Given the broad set of stakeholders involved with and affected by the effective management of sexual offenders, the issue of the relative value of sexual offender treatment as a component of that management system is of great significance. Unfortunately, the utility of psychotherapy for sexual offenders may, at best, be viewed as at a crossroads. For more than 30 years, there has been genuine and well-founded controversy about whether such interventions produce substantive and lasting changes for sexual offenders related to the prevention of future acts of sexual offending or reduced sexual offender recidivism. Relative to the conventional standards utilized to gauge treatment outcome studies, the existing scientific evidence does not yet provide support for the proposition that psychotherapy is an effective primary agent for “treating” or “changing” sexual offending or to reduce their potential for sexual reoffending. A series of reviews, including meta-analyses, have suggested—at best—“limited” or “cautious” evidence for the effectiveness of available psychosocial programs of sexual offender treatment for the typical sexual offender, at least as measured by the reduction of future sexual offense recidivism. That is, as even proponents of the efficacy of sexual offender treatment admit, the general results of studies of varying rigor have demonstrated only “small,” qualified positive outcomes for such interventions for select sexual offenders. Further, such proponents acknowledge that the evidence for such “small,” “promising” effects relies exclusively on scientifically “weak” studies and that the more rigorous scientific studies of sexual offender treatment have failed to show an effect of intervention. Even the Association for the Treatment of Sexual Abusers (ATSA) concluded in 2010: “After 50 years, the field of sex offender treatment cannot, using generally accepted scientific standards, demonstrate conclusively that effective treatments are available for adult sex offenders” (p. 1).

It is also important to consider broader contexts relative to the failure to demonstrate effectiveness of psychotherapies for sexual offenders. As Lilienfeld (2011) has pointed out, “Data indicate that large percentages of the general public regard psychology’s scientific status with considerable skepticism...” (p. 1); he notes that the widespread and long-standing public skepticism of psychology reflects the mental health profession’s failure to police itself and its problematic public face reflects the failure of the professional mental health field “to get its own clinical house in order and winnowing out the elements of our profession that are scientifically dubious, some of which have tarnished our hard-fought credibility...” (p. 125). As a function of some of these issues, the field of psychotherapy is facing an increasingly uncertain future, specifically the diminishing perceived value and utilization of psychotherapy. At a time when the demand for

mental health care is actually growing (almost doubling in the past 20 years), substantially less of it is being provided by nonmedical providers such as psychologists, social workers, etc. Very recently, several articles have pointed out that the field of nonmedical mental health providers has not made a convincing case for the use of psychosocial interventions and, in fact, by largely disavowing the need for evidence-based (largely scientifically evaluated) psychotherapies and effectively abandoning the mental health field to pharmacotherapy. Gaudiano and Miller (2013) noted that psychotherapy use is on the *decline* despite overall *increased mental health utilization*. They noted that from 1998 to 2007, there was approximately a 5 % decline in the use of psychotherapy alone and 8 % decrease in the use of psychotherapy with adjunctive medication. Several years ago, Baker, McFall, and Shoham (2009) pointed out that the lack of adequate training in and acceptance of the science of psychotherapy was leading to a greatly diminished role for psychotherapy in the mental health treatment field. In particular, they point to psychologist’s preference for valuing personal experience over research evidence—a “prescientific” perspective—as flying in the face of the evolution in health-care decision-making which places a premium on converging evidence that “a treatment is efficacious, effective-disseminable, cost-effective, and scientifically plausible” (p. 67). Similarly, Gaudiano and Miller place the responsibility for this decline in the utilization of psychosocial interventions primarily on psychotherapists’ tendency to rely on “personal experience” and “intuition” in performing their clinical work. Gaudiano and Miller argue that psychologists and other psychotherapists’ rejection of the principles of evidence-based practice largely stand in contrast to psychiatry’s training and practice model with its presumptive reliance on evidence-based medication research, primarily controlled studies involving random assignment of clients and similar scientific practices. Moreover, they point out that “the train has already left the station,” stating:

...as psychologists hem and haw about potential constraints placed on psychological practice by increasing scientific standards, and thus resist the notion of more prescriptive treatment approaches, the health care system has already adopted such an approach, is implementing it, and is holding psychologists accountable to it through reimbursement restrictions. (p. 816)

Thus, in the private sector, personal experience and judgment about “what works” with clients is being accorded increasingly little role in the endorsement of interventions. Moreover, psychotherapists and psychotherapies for various types of clients are progressively and increasingly rapidly being disenfranchised and excluded from possible treatment possibilities as a result of clinicians’ rejection or ignorance of currently available accepted empirical evidence and other supportive information related to such evidence-supported therapeutic practices.

Unfortunately, as a result of this longstanding failure to demonstrate clear effectiveness for psychosocial treatment, both policy makers and the more general community either are or are likely to be appropriately skeptical about practitioners' claims for sexual offender treatment effectiveness. In turn, the lack of demonstrated efficacy for psychotherapy for sexual offenders may increase reluctance to allow select sexual offenders to avoid incarceration or be released earlier from incarceration simply because they "participated in" or "completed" treatments, without scientific evidence that has demonstrated that such treatment results in decreased risk for sexual offense recidivism or relevant offender change. Further, in the absence of scientifically demonstrated results, the public and government stakeholders have been and are increasingly disinclined to endorse funding for the research of and/or implementation of existing or more novel programs of sexual offender treatment. In short, the relative role of psychotherapy as a component of a broad management approach for general sexual offenders necessarily remains in question.

This chapter is intended to provide a relatively straightforward, reasoned, and accessible review of the existing findings and issues regarding psychotherapy for sexual offenders. First, a brief synopsis of the research literature regarding psychotherapy in general shall be presented. Both the accepted methodological practices utilized in studying the effectiveness of psychosocial treatment as well as the results of the extant psychotherapy outcome literature will be summarized. Such a review provides a context for viewing the parameters for the more specific research literature on sexual offender treatment outcome. Second, the primary systematic reviews and meta-analyses of sexual offender treatment will be examined. The consensus of these reviews would appear to best be summarized as suggesting that to date the general efficacy of sexual offender treatment has not been scientifically demonstrated; few or no claims can be made for the "success" of such interventions. Third, a critical analysis of the methodological issues and inadequacies related to results of existing treatment research will be presented that provides perspective on the failure to yet demonstrate the effectiveness of sexual offender treatment.

---

### **The Nature, Methods, and Findings in General Psychotherapy Outcome Research Methodological Principles in the Scientific Investigation of Possible Outcomes of Psychosocial Interventions**

The available research on psychotherapy has been periodically summarized in the five sequential editions of the *Handbook of Psychotherapy and Behavior Change*, originally edited by Bergin and Garfield (1971, 1978, 1986, 1994)

and more recently by Lambert (2004, 2012). Kazdin (1986, 1994); Kendall, Holmbeck, and Verduin (2004); and others have described the nature of how models or theories of psychosocial intervention should be examined via a program of research that could delineate if and how psychosocial treatments might be effective with particular types of patients. Historically, the essential question for the scientific study of the effectiveness or efficacy of psychotherapy involves a scientific or empirical investigation to determine for persons with a particular presenting problem who want treatment: (1) whether the "average" person who participated in a particular treatment program had a better outcome than the "average" person who did not participate in that treatment *and* (2) if a benefit is observed, is it due to the intervention itself (or other factors). The goal of psychotherapy outcome research is initially to determine whether evidence can be obtained or demonstrated that particular treatments have *specific* effects, that is, an effect above and beyond those placebo/expectancy effects of those of nonspecific or common factors. Chambless and Hollon (1998) characterized psychotherapies as being efficacious if they work better than no treatment and as being specific if they are demonstrated to work better than nonspecific controls or credible alternative interventions. Hollon and Beck (2013) suggested that the term superior be applied when a given treatment outperforms all other viable alternative interventions.

According to both Kazdin (1986, 1994) and Kendall et al. (2004), as for other treatment outcome research methodologists, the essential approach to study of if and how psychotherapy might be efficacious is through randomized controlled trials (RCTs) of an intervention hypothesized to benefit a clinical population. Following the methodology of basic experimental science, a proposed psychosocial treatment approach is first studied or tested under specified controlled conditions; this provides an opportunity to determine the *effectiveness* of a proposed treatment approach—does it work with a relatively homogeneous group of persons with a similar presenting problem who are actively seeking (voluntary) treatment to address the presenting problem? Initially, such comparisons are typically conducted with more homogeneous individuals with the targeted presenting problem and may be comparable to other persons with more complex presentations or circumstances. Thus, Kendall et al. identified that to be most useful, treatment outcome studies required a controlled comparison of a specified intervention technique or program, with randomly selected clients exposed to the experimental treatment and control group(s) composed of relatively identical persons not exposed to that particular intervention. The first objective of outcome research is to determine if any consistent change occurs for persons receiving a positive treatment outcome relative to those not receiving that treatment [and not some unintended negative consequences as can happen (e.g., Rice & Harris, 2003;

Seto et al., 2008)]. Thus, some interventions may not result in desired change for clients as manifested on the relevant outcome measures of the presenting problem.

If treatment appears to show a benefit for those who participated under controlled conditions, an additional key objective is to determine if that change is, in fact, related to the specific treatment itself as opposed to other factors (e.g., spontaneous remission or the passage of time, dissimulation by clients, receiving a therapist's attention, the experience of repeated assessments). Such possible extraneous factors need to be "controlled" in a research study in order for one to have confidence that the treatment itself was responsible for any observed change. A control group provides the key means of potentially controlling for some factors (like characteristics of the subjects) that might be related to the outcome regardless of the experience. Consequently, a key issue in general psychotherapy treatment outcome research (similar to medication treatment studies) is to utilize a *control condition or group(s)*, particularly one that accounts for obvious potentially confounding factors such as client expectations or nonspecific factors related to interacting with a therapeutic agent. A "no treatment" control condition still may not protect findings from potentially confounding factors of an active treatment such as the anticipation of treatment, expectancy for change, and/or the act of meeting with a therapist. Even the so-called attention-placebo or nonspecific treatments, which may provide a reasonable measure of positive expectancy, may not be comparable to a condition where therapists provide a specific intervention to which they may be committed to ("believe in") as an effective treatment. Typically, only an alternative treatment condition (e.g., treatment as usual, another specific model of treatment) allows for controlling for nonspecific effects of treatment such as the length of treatment or client and therapist expectancies. Ideally, a treatment outcome study would involve at least three groups: a group that receives the treatment believed to produce a desired outcome, a group that receives an alternative credible intervention, and a group that does not receive any substantive intervention.

Almost all empirical, controlled studies begin with clients who are relatively compliant and motivated individuals; in addition, the treatment recruitment and delivery process typically induces expectancy bias for participants as well as their therapists. Given the possibility that client characteristics may strongly influence the outcome of an intervention, beyond controlled comparisons, the *random assignment* of clients in controlled psychotherapy trials is viewed as the second critical factor in treatment outcome research to ensure initial comparability between treatment and control groups. Random assignment of persons who are interested and motivated to address a particular presenting problem should eliminate unwanted potential effects of extraneous factors [demographic variables (such as age, socioeconomic status,

intelligence, education, and so on) as well as more substantive factors (such as the nature and degree of likely risk factors)]. Obviously, it would be problematic to allow the subjects themselves to select whether they are exposed to a particular experience or not exposed to a particular experience; the motivation and/or expectancy to either be exposed or not be exposed to the particular experience (or related characteristics) might determine the treatment outcome of participating subjects. Consequently, as Kendall et al. (2004) wrote:

...comparisons of persons randomly assigned to different conditions are required to ensure control of the effects of factors other than the treatment. Comparable persons are randomly placed into either the control condition or the treatment condition, and by comparing the changes evidenced by the members of both conditions, the efficacy of therapy over and above the outcome produced by the extraneous factors can be determined. (p. 19)

Then, "When treated clients evidence significantly superior improvement over non-treated clients, the treatment is credited with producing changes. This control procedure has desirable features and eliminates several rival hypotheses..." (p. 19). At the same time, randomization does not guarantee comparability, and the actual comparability of the participants in the treatment and control conditions should be examined. However, while the random assignment via RCT does not absolutely assure absolute comparability of the control and therapy conditions on all measures, it does maximize the likelihood of comparability. That is, randomized and controlled trials offer the best research design strategy for distributing pretreatment differences randomly; effectively, only randomization can eliminate the subtle selection biases that affect even the best alternative study designs. Almost all independent medical research groups (such as the Cochrane Collaboration) as well as various policy-making entities, including the US Center for Disease Control and Prevention and the US Food and Drug Administration, define and determine effective interventions based on the results of RCT. Further, as Howard et al. (1996) wrote, given the high degree of experimental controls imposed by RCT design: "...it is quite rare that a randomized experiment fails to conclude that the experimental treatment works" (p. 1060).

In addition, another preferred method for treatment outcome research involves the "intent-to-treat" design. That is, in more contemporary psychotherapy treatment outcome studies, the experimental or treatment group consists of *all* persons originally assigned to that group, whether or not they complete the intervention (e.g., those who complete and those who drop out of an assigned treatment or control group). The degree to which persons are retained in and complete the assigned treatment is considered an important aspect of the outcome or results of the treatment comparison; a treatment that loses a significant number of participants

and succeeds only for some persons would not necessarily be considered a successful or effective intervention (although it might provide useful information about which persons are most and least responsive to a particular treatment). Consequently, the outcome for individuals who drop out of or are terminated from treatment is typically counted as part of the intervention group's results. Thus, methodologically superior treatment outcome studies utilize "intent-to-treat analyses where the treatment group consists of all persons who began the treatment, including those who technically completed the program as well as those who dropped out after being assigned to the treatment group" (e.g., Chambless & Holon, 1998).

Standards exist and have achieved wide acceptance concerning the determination and rating of the methodological quality of treatment outcome research. For example, Sherman et al. (1998) developed a scale of methodological rigor, known as the Maryland scale, to provide a clear perspective on the quality of the scientific quality of crime prevention programs. The scale provides an assessment of the quality of the research design and whether study results can reasonably be used to draw conclusions about the effectiveness of sexual offender treatment. Thus, per the Maryland scale, a score of "1" indicates that a correlation exists between a treatment program and an outcome measure, a score of "3" indicates that the study included an intervention group and a comparison group, and a top score of "5" indicates that the study used both random assignment and an analysis of comparable intervention and comparison groups.

Once the *effectiveness* of a particular psychotherapy approach is initially established under controlled conditions to limit potential methodological confounds, the outcome comparison is typically the subject of attempts to replicate or cross-validate the results, ideally by other investigators than those who initially developed and tested the approach. Assuming multiple successful replications from RCTs performed by scientists of varying allegiance to that approach, the experimental intervention is tested with RCTs in more naturalistic situations with clinically representative clients with the primary presenting problem (e.g., those outside of university research settings, typically with more complex presentations and/or severity). Thus, once robust evidence exists that a psychosocial intervention is effective in RCTs (under more controlled conditions), the intervention can be systematically tested with an expanded group of clients with the presenting problems (e.g., those with more severe or comorbid conditions). At that point in time, modifications of treatment procedures may also be tested under controlled conditions to optimize the potential outcome with more heterogeneous clients. If positive outcomes consistently result in more naturalistic or clinically representative settings, then the psychosocial intervention is said to have demonstrated *efficacy*.

In the 1990s, the general psychotherapy field moved to endorse a model of Empirically Supported Therapies (ESTs).

As Arkowitz and Lilienfeld (2006) noted, this move was fueled by several considerations. First, ESTs are argued to protect clients against "a seemingly endless parade of fad therapies of various stripes..." (p. 45), a number of which have been found to be ineffective or even harmful. Second, ESTs are viewed as performing a quality control function for health-care agencies and policy makers to make scientifically informed decisions about which treatments should be reimbursed; "By placing the burden of proof on a treatment's proponents to show that it is efficacious, the EST list helps to ensure that therapies promoted to the general public have met basic standards" (p. 45). In 2005, the American Psychological Association (APA) issued a policy statement regarding evidence-based practice (EBP) in psychology, stating: "Evidence-based practice in psychology (EBPP) is the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences" (p. 13). They noted that this was similar to the definition of evidence-based practice adopted by the Institute of Medicine in 2001, where evidence-based practice (EBP) was defined as the integration of best research evidence with clinical expertise and consideration of patient characteristics; it was recommended that therapists determine the applicability of available research conclusions to the needs of particular help-seeking clients; thus, treatment should involve the application of available research evidence with probabilistic inferences for help-seeking clients based on current scientific knowledge.

### **To What Degree Are Psychosocial Treatments Generally Effective?**

To provide a context for considering the effectiveness of sexual offender treatment, it obviously makes sense to consider to what degree and in what ways psychotherapies are effective in the broader field of mental health problems. There has been a long controversy as to whether psychotherapy as a type of interventions has been demonstrated to be effective. Thus, in 1952, Eysenck published a review of 24 studies and concluded there was no research evidence to support the effectiveness of persons participating in psychotherapy compared to groups not participating in psychotherapy. In contrast, since Eysenck's publication, most studies evaluating the outcome of psychotherapy have been more positive despite the increasing methodological rigor that characterized those studies.

Subsequent to Eysenck's review, meta-analyses (MA) or analyses not of subjects but of existing studies began to appear. In MA, statistical methods are used to obtain a quantitative estimate of the overall or cumulative effect of a set of existing interventions on an outcome. By combining results of multiple smaller studies (e.g., in terms of sample size) and weighting them by size, it is hypothesized that the combined

results (now based on a larger number of subjects) provide greater power and might allow for identifying “effects” across studies that might be missed in individual studies, particularly those with small numbers of subjects. In addition, a potential strength of meta-analysis comes from the use of a standardized unit to compare outcomes from studies that may use different measures and by averaging effect sizes across different studies and comparisons. This increases the effective sample size for investigation and potentially minimizes the influence of extraneous factors in individual studies. Such a practice allows for a more precise evaluation of the efficacy of treatment programs. However, as Lambert (2013) pointed out: “Meta-analysis is not a panacea and cannot be used to create worthwhile information if it’s based on poorly designed studies or is biased” (p. 206).

As meta-analytic statistical techniques emerged, reviews of the expanding literature on psychotherapy have been subjected to a more sophisticated quantitative analysis. However, clear evidence that psychotherapy was associated with positive outcomes for general mental health problems did not emerge for 30 years after the Eysenck study. Smith et al. (1977, 1980) conducted a particularly significant MA of the extant psychotherapy literature. They analyzed more than 475 studies and demonstrated that the effects of psychotherapy were superior to no treatment and to placebo control conditions, typically for clients with some form of emotional distress or “neurotic” condition. Lipsey and Wilson (1993) reviewed 302 meta-analyses of a range of psychological, educational, and behavioral treatment and found a strong positive effect. They utilized more stringent criteria to examine a limited sample of studies (156 meta-analyses) and found that the average treatment effect size was 0.47.<sup>1</sup> They concluded that the evidence from this MA indicated that psychosocial treatments “generally have positive effects” (p. 141) on those who participated in them relative to a control condition.

However, based on an earlier analysis of the literature, Shadish et al. (1997) suggested that previous meta-analyses had overestimated the effects of treatment because they calculated *unweighted* effect sizes, which gave more importance to studies with larger N’s. Thus, they recalculated the effect sizes for the Smith et al. (1980) data set and found an effect size of 0.60 (a medium effect) as opposed to the 0.85 effect size (a large effect) originally reported by Smith et al. (1980), and Wampold et al. (1997) also reanalyzed previous meta-analyses and noted that the effect size of psychotherapy compared to no treatment was 0.82 (considered a large effect); however, the effect size of psychotherapy compared to a placebo condition was 0.48, and the effect size of placebo vs. no treatment was 0.42. Thus, the relative effectiveness of psychotherapy was reduced to a medium effect when compared to placebo condition; further, a placebo con-

dition, in and of itself, produced a medium-sized effect. These results suggested that a significant mechanism for the positive effects of psychotherapy was nonspecific and that, for predominantly emotional problems (such as anxiety or depression), positive benefits of psychotherapy may be largely due to factors such as clinical attention and/or expectation of change. Westen et al. (2005) pointed out that when investigators have compared two bona fide intent-to-succeed treatments, the outcome effects are generally small (e.g., an average ES or d of 0.20). That is, when two meaningful interventions are compared to one another (as opposed to a no treatment condition), the effect size was substantially reduced for clients.

Lambert and Ogles (2004) concluded that studies that were representative of clinical settings and conditions (e.g., more varied clients with comorbid conditions) appeared to produce generally similar effects to those that were not representative of clinical conditions. However, they also noted that higher-quality RCTs of treatment for actual clinical conditions were generally lacking; most extant positive studies for psychosocial interventions were conducted in research settings with more pure and circumscribed client samples. Most recently, Lambert (2013) wrote: “From 40 to 60 % of clients show a substantial benefit in carefully controlled research protocols, *although far fewer attain this degree of benefit in routine practice*” (p. 204, emphasis added).

Another critical question regarding psychotherapy concerns whether clients maintain whatever measured “gains” or initial response that they are reported to have made in treatment. Nicholson and Berman (1983) conducted the earliest and most influential meta-analysis regarding follow-up outcome of persons treated with psychotherapy. In their study of 67 studies, while noting some divergence in the studies, they reported that treatment gains were maintained (largely for the treatment of problems of emotional distress). Later, however, Lambert and Ogles (2004) identified several methodological concerns that prevent reaching broad conclusions about the maintenance of treatment gains. First, they noted that client attrition from the end of treatment to follow-up data collection was a critical issue (as well as attrition during treatment itself); that is, a significant number of persons who participate in treatment studies either leave the study during or after the controlled intervention phase. Consequently, only smaller groups are available for study at points distal to the end of psychotherapy. Second, in the majority of cases, most studies do not continue to follow subjects in control groups after treatment ends making follow-ups “naturalistic” (and *not* controlled or “comparative”). Westen and Morrison (2001) found that only 36–38 % of persons treated for depressions remained improved at a 2-year follow-up and that there were low levels of “sustained efficacy”; if those individuals who began but did not complete psychotherapy were included for study, the improvement rate dropped to approximately 25 %. They noted that

<sup>1</sup>Cohen (1992) identified the benchmarks for effect size for comparing independent means as 0.20 (small), 0.50 (medium), and 0.80 (large).

the available follow-up results were worse for clients with anxiety disorders. Westen and Morrison (2001) argued that psychotherapy provided to relatively pure samples of depressed and anxious clients, with rigorous inclusion and exclusion criteria, results in improvement/initial response of pathological states (as distinguished from disorders) which was approximately 50 % for those persons who complete psychotherapy but that the majority of clients do not show sustained improvement over 1–2 years, particularly for “generalized affective states.” That is, the average client will maintain a mild but clinically significant level of symptoms after intervention, but “a substantial number of patients will continue to be highly symptomatic” (p. 885).

At best then, the available evidence suggests that various types of psychosocial interventions are somewhat effective in treating persons with relatively unidimensional presenting problems of emotional distress (primarily for anxiety and depressive conditions) most commonly found in typical clinical practice (i.e., for which people seek treatment). However, the evidence for the effectiveness of psychotherapy for more severe problems is much less clear. Compared to treatment for more circumscribed problems such as emotional distress, Lambert and Ogles (2004) showed that the average effect size for efficacy is much lower (e.g., approximately 50 % lower) for more severe problems such as schizophrenia, alcoholism, and delinquency and for persons characterized by “social detachment.” Lynch, Laws, and McKenna (2010) in a meta-analysis of well-controlled RCTs found that CBT was not effective in reducing symptoms or preventing relapse for schizophrenia or in reducing relapse in major depression or bipolar disorder. Even in the treatment of major depression, they found that the effect size for reducing symptoms was small. Hollon and Beck (2004) concluded: “It remains unclear just how effective CBT (including relapse prevention) is in the treatment to substance abuse. It typically outperforms minimal treatment control, but is has a more inconsistent record relative to attention placebos and rarely exceeds alternative interventions” (p. 474). As noted, Kopta et al. (1994) showed that patients with significant characterological issues (e.g., maladaptive personality traits or personality disorders) required much “stronger” doses of treatment over a longer period of time prior to showing symptomatic improvement (e.g., treatment sessions that were more frequent and of longer duration); similarly Tyrer and Johnson (1996) also showed that clients with comorbid personality disorders have the highest initial levels of symptoms and improved the least over follow-up. Clarkin and Levy (2003) reported that clients with a greater number of personality disorder traits also had difficulty staying in active treatment and would drop out at a higher rate. Multiple reviews of the treatment of personality disorders, particularly Borderline and Antisocial Personality Disorders have found no or little scientific evidence that such conditions can be treated efficaciously (e.g., Binks et al., 2006; NICE, 2009a, b; Duggan et al., 2007; Gibbon et al., 2010; Stoffers et al. 2012). Further, in general, available studies show

that individuals with maladaptive traits or personality disorders have much higher relapse rates when compared to patients with no such comorbid problems. In addition, methodological problems of lower power and attrition are more common among across studies of persons with more severe and/or chronic problems. While it can be said there is some evidence for psychotherapies involving persons with severe or chronic problems having a relatively positive effect on some elements of their problems and on “satisfaction” with the therapy experience, in a significant number of cases, “treated” clients continued to manifest ongoing symptoms of varying degrees of severity and/or to convert to other significant psychiatric conditions.

Currently, there is little evidence that any specific type of psychotherapy [e.g., cognitive-behavioral therapy (CBT)] is more effective than another therapy, particularly when the allegiance (expectancies of treatment success) of the investigator and study therapists is controlled. Wampold et al. (1997) and Wampold (2001) in reviewing their own MAs and those of others concluded that there was little evidence that specific ingredients are necessary to produce change as a result of exposure to psychotherapy. [Note that many of these studies involved the treatment of emotional distress, again typically anxiety or depression.] In a later MA, Wampold et al. (2002) also found that CBT was not more effective than other bona fide (credible) psychotherapies for unipolar depression and that all bona fide psychological treatments were equally effective in mood improvement. In their review in 2004, Lambert and Ogles concluded: “There is a strong trend toward no differences between techniques are modes in amount of change produced which is counterbalanced by indications that, under some circumstances, certain methods (generally cognitive behavioral) or modes (family therapy) are superior” (p. 164). They concluded that extant research “shows surprisingly small differences between the outcomes for patients who undergo a treatment that is fully intended to be therapeutic” (p. 164). With some exceptions, research generally supports that somewhat greater effectiveness of CBT over alternative psychotherapies has been demonstrated for clients with anxiety or depressive disorders and particularly for individual but not group psychotherapy. However, the mechanism of action for such outcomes is unclear regarding CBT. Per Longmore and Worrell (2007), review of CBT identified three empirical anomalies in the CBT empirical literature:

Firstly, treatment component analyzes have failed to show that cognitive interventions provide significant added value to the therapy. Secondly, CBT treatments have been associated with a rapid symptomatic improvement prior to the introduction of specific cognitive interventions. Thirdly, there is a paucity of data that changes in cognitive mediators instigate symptomatic change.... A comprehensive review of component studies finds little evidence that specific cognitive interventions significantly increase the effectiveness of the therapy.... Although evidence for the early rapid response phenomenon is lacking, there is little empirical support for the role of cognitive change as causal in the symptomatic improvements achieved in CBT. (p. 173)



More generally, collectively, research findings indicate that substantive behavioral change both precedes and lays the foundation for later cognitive change. Finally, it must be noted that very recent evidence suggests that modern CBT clinical trials appear to provide smaller decreases in depressive symptoms as compared with earlier research trials.

Measurement of clinical change has also been problematic in the general psychotherapy outcome literature. The percentage of persons considered “improved” has been shown to have more to do with particular rating scales and sources of information (e.g., global, self-report ratings) rather than actual behavioral change (e.g. Hill & Lambert, 2004). When more specific problems and behaviors were rated for change, there is less evidence of significant change or improvement. Weiss et al. (1996) reviewed 41 studies and found a basic lack of agreement regarding the nature of improvement; when agreement between client and therapist was found, it was not high. Both Pekarik and Guidry (1999) and Rosenblatt and Rosenblatt (2002) reported very similar results. In contrast, agreement was higher between clients and external raters or judges, clearly suggesting that clinicians were poor judges of treatment-related behavioral change (e.g. Johnson & Friberg, 2015). In other research, Gregerson et al. (2001) looked at ratings of treatment made pre- and posttreatment. They found that the difference in the size of treatments of pre- and posttreatment suggests that retrospective (post) evaluations of treatment change “overestimated treatment effects” by a *factor of two* compared to actual pre-/post-measurements. “Life records” and real outcome measures would be considered to be the least reactive of available assessment methods. Hill and Lambert (2004) noted that differences in outcome results have been found to be a function of a source (e.g., client, therapist, expert judges, and significant others) and not content (the actual functioning of a client). They concluded that therapist ratings of treatment outcome and global ratings of change are associated with an illusory “perception of greater effectiveness” of treatment compared to more specific and more distal measures. In their review, Hill and Lambert also pointed out that data from therapists who are aware of the treatment status of clients produce larger positive ratings than those from virtually all other sources. Similarly, they found that global ratings of change produce larger estimates of change than ratings on specific dimensions, symptoms, or problem areas and that proximal ratings lead to more positive ratings of change than distal ones. Physiological measures, in contrast to those by therapists or “unblinded” evaluators (those who know the treatment status of the client), typically show small effects of treatment, even when they are the targets of treatment. They noted in their review that global ratings of treatment goals are characterized by multiple methodological problems. Among them were high correlations among goal ratings (a “halo” effect); the use of relative perceived goal change as opposed to absolute, well-defined, standardized

criteria for change; and a confounding between therapist expectancy and their ratings. They recommended that to the extent, global ratings are utilized to measure outcome that follow-up evaluators be as independent as possible from therapists/goal setters so that there is maximal independence of and objectivity in ratings.

Other issues have been identified relative to determination of aspects of the effectiveness of psychotherapy. There has been a consistent finding in the general treatment outcome literature that the investment of a researcher/therapist to a particular model of intervention accounts for a significant amount of the measured outcome in treatment studies that find particular interventions effective. Recently, Munder et al. (2013) conducted a MA of 30 studies of Researcher Allegiance (RA). They found that the mean RA-outcome association was statistically significant ( $r=0.26$ ) corresponding to a moderate effect size and that this relationship was robust across several moderating variables including characteristics of treatment, population, and the type of RA assessment. Munder et al. concluded that the RA-outcome association is substantial and robust. In addition, Lambert and Ogles (2004) reviewed several large treatment outcome studies that attempted to “dismantle” or study components of interventions. Results indicated that treatment outcome was not related to which specific components clients received or the acquisition of skills (symptoms improved before skills training and potential behavioral change). More recently, a meta-analysis was conducted on both additive and dismantling studies, which examined their effect both at the end of formal treatment and at follow-up. Bell et al. (2013) found that for dismantling studies, there were no significant differences between the full treatments and the dismantled treatments. For additive studies, the treatment with the added component showed a small but significant effect at completion and a large effect at follow-up. However, this was only true for specific problems that were targeted for intervention. Thus, some specific intervention components, directly related to the primary treatment target, made only a modest contribution at outcome. In short, other than investigators “finding” what they expect or want to, there are significant questions about what elements of psychotherapy “matter” or “work” relative to “symptom relief” or “behavior change.”

In research settings, treatment dropout or attrition has averaged to 47 % and is even higher in actual clinical settings (e.g., Lambert & Ogles, 2004); per a meta-analysis, approximately 47 % of patients dropped out of psychotherapy (Wierzbicki & Pekarik, 1993). Clarkin and Levy (2003) identified that clients with maladaptive personality traits (such as those with personality disorders) were at high risk for premature dropout, with dropout rates varying from 40 to 67 %. Three client variables found to be particularly related to negative outcomes were overall problem severity at intake, interpersonal difficulties, and comorbid personality disorders

(e.g., Lambert & Ogles, 2004). Lillienfeld (2007) reported on a number of psychotherapies for specific problems that actually demonstrated harmful outcomes for clients. More generally, as many as 10 % of clients' problems worsen as a result of their participation in psychotherapy (e.g., Lambert & Ogles, 2004).

Research has also examined therapist and client variables related to treatment outcome. Lambert (1992) concluded in an earlier review that as much as 40 % of client improvement may be attributed to client variables and extra-therapeutic influences. Thus, number and severity of maladaptive personality traits and social detachment were also found to be associated with poor psychotherapy response (e.g., Clarkin and Levy, 2003). Some clinicians appear to be "outliers" in terms of their increased effectiveness as psychotherapists (Lambert & Ogles, 2004; Lambert, 2013); that is, in particular, it appears that specific therapists account for a disproportionate percent of "successful cases" in treatment outcome studies, leading to suggestions that there should be increased study of the "empirically validated therapist." Per Lambert and Ogles (2004), the importance of the so-called therapeutic alliance is a necessary but not sufficient condition for change in psychotherapy. They view the therapeutic alliance as a manifestation of the critical role of common factors in effective psychotherapy. However, they determined: "...we simply do not know enough yet about the therapist factor to specify when and how it makes a difference, nor when it matters more than technique" (p. 168). Similarly, Crits-Christoph, Johnson, Connolly Gibbons, and Gallop (2013) concluded:

Despite extant research, there are mixed reviews on the importance of the therapeutic alliance in treatment outcome;" they pointed to a recent MA that found a "small to moderate relationship between the [therapeutic] alliance and therapeutic outcome ( $r=0.27$ )." (p. 302)

Crits-Christoph et al. also pointed to research which suggests that early positive change in symptoms is the actual *cause* of a positive or improved "therapeutic alliance," as opposed to the opposite process.

### Summary of Psychotherapy Outcome Literature

Relative to control conditions, psychotherapy has been found somewhat to be moderately beneficial for persons motivated for change in their lives, particularly for persons who seek treatment primarily because they themselves are disturbed by moderate to high degrees of emotional distress (e.g., they "feel badly"). For many persons who seek psychotherapy for anxiety or depression, the effects of treatment appear to be somewhat enduring (albeit these have typically been relatively "pure" clients by virtue of exclusion criteria that screen

out significant—and typical—comorbidity, for example). In contrast, for persons seeking treatment to address more complex or severe behavioral problems, there is little to some evidence for the relative effectiveness of psychotherapy typically for particular features of those conditions (and not necessarily changes in key signs or symptoms). In general, greater problem severity and chronicity, comorbid psychiatric conditions (in particular, maladaptive personality traits), and functional impairment in everyday life were each associated with decreased response to psychosocial treatments. Researcher allegiance appears to account for a significant amount of variance in outcome; those invested in a particular intervention are more likely to find it effective. There is *decreasing* evidence that *specific types* of psychotherapy produce differential degrees of improvement, including even the treatment of emotional distress. Thus, most interventions are equally effective for persons with emotional distress. Little superiority of CBT has been demonstrated for more severe and/or behavioral conditions. To be effective, psychotherapy needs to be provided in a sufficient dose relative to the severity of the individual's presenting problem; a greater number of and/or more severe problems require more intense and/or higher doses of psychotherapy. Clearly, client characteristics have a particularly significant role in or influence on the outcome or benefit realized in psychotherapy. Therapist characteristics also impact the outcome of psychosocial treatment for common presenting problems; some individuals appear to be much more effective with clients than other clinicians. In 1986, Lambert concluded that common (therapeutic) factors accounted for 30 % of the therapeutic effect, technique 15 %, expectancy (placebo-effect) 15 %, and spontaneous remission 40 %. More recently, Lambert (2013) suggested that improvement from psychotherapy is a function of the following four factors to the indicated degree: client/life situation (40 %), common factors (30 %), client expectancy (15 %), and (specific) techniques (15 %).

---

## Key Reviews of Sexual Offender Treatment Outcome Reviews

### General Systematic Reviews of Sexual Offender Treatment

Systematic reviews (SR) of treatment research involve a particular approach to the examination of scientific literature, one that attempts to identify and appraise available studies regarding interventions for a particular problem or condition. SRs include a clearly formulated question; use systematic and explicit methods to identify, select, and critically appraise relevant research; and collect and analyze data from the studies that are included in the review. Specific statistical methods (e.g., such as meta-analysis) may or may not be used to

analyze and summarize the results of the included studies. In most fields of medicine or mental health more specifically, SRs are limited to a focus on high-quality studies such as RCTs. In certain cases, SRs involve simply a sequential discussion of selected studies with a critical discussion of the apparent results across studies. A second type of review is, in effect, a subsection of systematic reviews and often relies on meta-analysis (as noted previously, a particular statistical technique which appraises the combined results of varied studies utilizing common metrics). It is worth noting that a recent review of 300 studies by Moher et al. (2007) found that all systematic reviews were not equally reliable. Moher et al. concluded that the quality of reporting in such reviews was often inconsistent. For therapeutic reviews, the comparison of Cochrane<sup>2</sup> and non-Cochrane reviews provided discouraging results and suggested little improvement in the quality of reporting of non-Cochrane reviews over time. It was found that many non-Cochrane reviews did not report key aspects of systematic review methodology. Further, strong evidence of bias in outcome reporting was noted for non-Cochrane reviews.

In the first modern SR of sexual offender treatment, Furby et al. (1989) found few well-designed studies of sexual offense recidivism, including those where offenders received specialized sexual offender treatment or generalized treatment. In particular, they noted that the most common design for studies they reviewed were single-group, posttest-only designs; these were investigations where a group of sexual offenders were provided treatment and a recidivism rate was determined for that group. Thus, these studies did not include a “no intervention” control group to compare sex offense recidivism rates for comparable sexual offenders who did not receive treatment. Furby et al. concluded that there is “as yet no evidence that clinical treatment reduces rates of sex reoffenses” (p. 27).

White et al. (1998) developed the first Cochrane review (CR) of “Managements for people with disorders of sexual preference and for convicted sexual offenders.” White et al. attempted to identify all relevant randomized controlled trials and could identify only three methodologically sound studies of the type typically considered for medical efficacy treatments and only two of these were psychological interventions: Romero and Williams (1983) compared psychody-

amic group treatment to probation, while the Sex Offender Treatment Evaluation Project (SOTEP) Marques et al. (1994) was the preliminary report of what would eventually be the largest RCT of CBT-RP specific to sexual offender treatment. White et al. concluded:

It is disappointing to find that this area lacks a strong evidence base, particularly in light of the controversial nature of the treatment and the high levels of interest in the area...large, well-conducted randomized trials of long duration are essential if the effectiveness or otherwise of these treatments are to be established. (Abstract)

Alexander (1999) reviewed 79 studies of rates of sexual offense recidivism of sexual offenders ( $n=10,988$ ) as a means of opining whether sexual offender treatment might make a difference in such recidivism. She explicitly rejected applying a meta-analytic approach due to methodological issues regarding the lack of standardized research designs, making it problematic to determine whether observed differences were the result of exposure to treatment or to other study or group differences (e.g., follow-up periods, offender samples, recidivism criteria, or other design features). Further, as she noted: “The current subject pool does not include subjects who dropped out or were terminated during the course of the treatment. Dropouts/non-completers were excluded due to the lack of consistency with which data on these subjects were reported in their various studies” (p. 103). Alexander reported a very *slight* difference in sexual offense recidivism in favor of treatment ( $d=0.12$ ). Again, the majority of studies included no control group let alone subjects randomized to treatment; thus, the treated and untreated sexual offenders (the “quasi” control group), in most cases, were from different samples. As a result, it was unclear what kind of comparative conclusion could be reached. Hanson et al. (2002) indicated that a valid criticism of Alexander’s results was that there was too much method variance across studies to allow for clear conclusions.

In a SR, Gallagher et al. (1999) examined 25 published and unpublished studies on the effects of sexual offender treatment on sexual reoffending. Of these, 22 are related to adult sexual offenders. They found that 11 or 44 % of the studies included no comparison group and 9 or 36 % included “nonspecialized” treatment. Further, only 2 studies used random assignment (RCT), only 5 used subject level matching, and only 12 % included treatment dropouts. The authors conducted a meta-analysis but provided little detail of their particular methodology. Overall, they concluded that most treatment groups fared better than comparison groups relative to sexual offense recidivism. They found a “medium” effect size, but they also found that effect sizes varied greatly, “suggesting genuine differences in treatment effect estimates across studies” (p. 22). [Of note, they considered the earliest publication of the SOTEP study, which showed more promising results than the final version.] Gallagher et al.

<sup>2</sup>The Cochrane Collaboration consists of systematically reviewing randomized trials of the effects of varied health interventions. Cochrane reviews (CR) are viewed a scientific investigation in itself, with a pre-planned method section and an assembly of original studies (predominantly randomized controlled trials and clinical controlled trials). The results of these multiple primary investigations are synthesized by using strategies that limit bias and random error. The Cochrane Collaboration provides a handbook for systematic reviewers of interventions across various mental and physical health domains. Over 150,000 RCTs exist in the Cochrane library, and they are recognized as a key resource in evidence-based medicine.

showed that neither the strictly behavioral nor the augmented behavioral treatment produced significant reductions in recidivism. They reported that cognitive-behavioral treatment programs appeared to be effective in reducing sexual offense recidivism. Gallagher et al. found no difference between studies using cognitive-behavioral therapy alone or with relapse prevention methods. They concluded that despite heterogeneity of effects and various methodological issues, there was “sufficient evidence” to suggest the effectiveness of CBT for sexual offenders. However, as Hanson et al. (2002) indicated, the Gallagher et al. review included six studies in which biases in favor of a treatment effectiveness might be expected; in addition, the Gallagher review was also based on the preliminary results of studies in which the final or later results were more negative for the same studies (e.g., SOTEP).

Grossman et al. (1999) attempted to review what they regarded as available key papers presenting data on outcomes for sexual offenders in treatment programs. They noted that generally results suggested that biological and psychosocial interventions appeared to reduce sexual offense recidivism. However, they concluded: “Although some forms of treatment for sex offenders appear promising, little is known definitively about which treatments are most effective for which offenders, over what time span, or in what combinations” (p. 358). “In particular, they noted that available findings appeared to suggest that the more high risk a sexual offender was, the less confident we can be that treatment will have lasting benefits” (p. 359). Grossman et al. urged caution in “unfolding the implications of the positive treatment findings in the literature,” stating that while treatments exist and results indicate some potential, “They are, however, complex, difficult to interpret and cause for cautious optimism as best. If mental health professional and society at large are to accept the challenge of promoting treatment for sex offenders, vigorous ongoing research efforts are mandatory” (p. 359).

Also in another SR from 1999, Polizzi, MacKenzie, and Hickman (1999) observed that the “The recent reviews and meta-analyses concerning the efficacy of sex offender treatment provide conflicting viewpoints” (p. 370). They compared prison-based to community-based sexual offender treatment programs. A key feature of this review is that they utilized the so-called Maryland criteria to assess scientific rigor. Initially, they began with consideration of 21 studies. However, the investigators rejected 8 studies as “too low in scientific rigor,” leaving just 13 studies to examine. Polizzi et al. identified that approximately 50 % of the remaining studies showed statistically significant findings supportive of sexual offender treatment in reducing sexual recidivism. Most of these studies employed a CBT approach to treatment. They concluded that community-based programs were “effective.” However, they only identified two studies that they characterized as possessing “scientific merit” [one of

child molesters (from 1988) and exhibitionists (from 1991)]. More importantly, in the studies examined in their SR, Polizzi et al. did not control for the effect of dropouts/refusers on the recidivism rates of untreated comparisons. They concluded that “non-prison-based sex offender treatment programs using cognitive-behavioral treatment methods are effective in reducing the sexual offense recidivism of sex offenders.” Thus, they claimed community-based CBT for sexual offenders “works.” In contrast, they concluded that prison-based programs using CBT were “promising,” “but the evidence is not strong enough to support a conclusion that such programs are effective” (p. 20). Of note, the authors included the SOTEP study as a community-based program, whereas the participants were actually prison inmates whose treatment site was a state hospital. The authors noted that there were too few studies focusing on particular types of sexual offenders to draw conclusions about whether treatment was effective for rapists, child molesters, or “high-risk” sexual offenders. Polizzi et al. concluded: “Any conclusions drawn from this review must remain tentative. With a heterogeneous population, it is difficult to provide general conclusions about the effectiveness of sex offender treatment programs” (p. 372).

Bilby, Brooks-Gordon, and Wells (2006) conducted a SR of quasi-experimental and nonrandomized controlled trials with matched and non-matched controls, including 21 quasi-experimental studies from the UK, USA, Canada, and Europe. They noted that due to the wide variety of outcome measures, they felt that they could not conduct a meta-analysis. They pointed out that although the majority of these studies were matched studies: “The problem with type of study is that, to match successfully, investigators need to know about all the relevant factors which may influence outcome, and this is unlikely to be the case, leading to potential differences between experimental and control groups” (p. 470). They also noted that 13/21 studies did not specifically match participants and that control groups were drawn from very different samples. In a later article, Brooks-Gordon and Bilby (2006) wrote:

Most participants in matched trials where a significant treatment effect was found were allocated to treatment groups according to sentencing decision and post-sentencing risk assessment. Most of these studies were matched retrospective trials carried out on offenders in the criminal justice system; matching was done retrospectively. Matching offenders with a control group is problematic and can threaten the quality of the research. The results here were equivocal: more studies found no statistically significant treatment effect than found a significant effect. (p. 5)

Bilby et al. (2006) found that 7 studies showed a statistically significant treatment effect and 10 did not, while in 4 studies the data were not clear enough for analysis.

Brooks-Gordon et al. (2006) conducted a SR of RCTs regarding the effectiveness of psychological treatments for

sexual offenders. They found nine RCTs (all reported before 1998 and totaling 567 offenders), 231 of which had been followed up for 10 years. They concluded: "Analysis of the nine trials showed the cognitive behavioural therapy (CBT) in groups reduced re-offence at 1 year compared with standard care ( $n=1,555$ ) but increased re-arrest at 10 years" (p. 442). They noted that if the Romero and Williams (1983) study had had only a few more rearrests in the intervention group, it could be suggested that treatment was less effective than doing nothing. Brooks-Gordon et al. wrote that their findings were "likely to be controversial as there is a huge investment in sexual offender treatment programmes, and many policy-makers erroneously and unreservedly assert that sexual offender treatment therapy is effective—whereas our findings show that uncertainty about effectiveness of treatment remains" (p. 460). Further, they stated:

The ethics of providing this still-experimental [sexual offender] treatment to a vulnerable and potentially dangerous group of people outside of a well-designed evaluative study are debatable...Psychological interventions could help or they could harm sex offenders...In an environment of limited resources it would seem imprudent to allocate funds to unproven and potentially harmful interventions. (p. 461)

Kenworthy et al. (2003, 2004) initiated an updated CR of White et al.'s (1998) earlier study, noting that there was significant political and institutional pressure to prove that treatment works. However, they concluded: "To date, no positive treatment effects have been found in quasi-experimental institutional programmes" (abstract). They examined nine random assignment studies involving treatment of over 500 sexual offenders that were available as of 2002; thus, they evaluated the same studies as Brooks-Gordon et al. (2006). However, Kenworthy et al. found that a lack of relevant data made it impossible to draw conclusions for clinicians, concluding:

Limited data make recommendations difficult. One study suggests that a cognitive approach results in a decline in re-offending after one year. Another large study shows no benefit for group therapy and suggests the potential for harm at ten years. The ethics of providing this still-experimental treatment to a vulnerable and potentially dangerous group of people outside of a well-designed evaluative study are debatable. This review proves such studies are possible. (abstract)

The Institute for Health Economics (IHE) in Alberta, Canada, provides evidence in health technology assessment to assist in health policymaking and best medical practices. The IHE, like the Cochrane Collaboration, is an independent, not-for-profit organization that performs research in health economics and synthesizes evidence to assist health policymaking and best medical practices. The IHE published a Health Technology Assessment (HTA) Report entitled "Treatment for Convicted Adult Male Sex Offenders" (Corabian, Opsina, & Harstall, 2010a). [Subsequently,

Corabian et al. (2010b) provided an e-journal summary of the IHE study.] The IHE identified eight SRs conducted on the effectiveness of sexual offender treatment interventions that met their inclusion criteria; all eight focused on the use of psychotherapy and one also included studies of surgical castration and hormonal medication (e.g., Losel & Schmucker, 2005). These studies were selected as meeting the IHE inclusion criteria, which by virtue of design and quality of reporting were most likely to provide "high levels of evidence." They concluded that a subset of the studies showed "small but statistically significant reductions in sexual and general recidivism rates among convicted adult male sex offenders treated with various cognitive behavioural therapy (CBT) approaches..." (p. iv). Yet they noted when analyses were restricted to the few available RCTs, a mean effect was shown, but it was *not* statistically significant. Further, the IHE also stated:

Confidence in these findings, however, must be tempered as the available evidence is based mostly on poor quality primary research studies...Given the methodological problems of the available primary research it is difficult to draw strong conclusions about the effectiveness of sexual offender treatment programs using various CBT approaches for such a heterogeneous population. (p. iv)

In addition, the IHE stated: "SOT programs neither cure sexual offending nor guarantee a complete cessation of offending..." (p. 37). At best, they noted that such interventions represent but one element in a comprehensive risk management strategy for sexual offenders. The IHE further noted: "Overall, the results reported by the selected SREs provide little direction regarding how to improve current treatment practices...There are still uncertainties regarding the most useful elements and components of a SOT program for convicted adult male sex offenders." They concluded that the available research indicated "...more and better research was needed to clearly answer the set of remaining questions" (p. iv).

Later in the IHE report, they noted:

...since the evaluated programs were not sufficiently documented...it was not possible to identify if any characteristics or elements contributed more or less to the success or failure of a program and who of the involved offenders were most likely to benefit from or be harmed by treatment. SOT programs typically work within a broad CBT framework but may vary in terms of resources, philosophy of a program and its treatment objectives, timing, duration, format, intensity, and content of treatment, level of worker expertise and treatment fidelity/integrity as well as the referred sex offenders' characteristics and selection criteria for participation in the program (which can be based on various risk assessment modalities or no risk assessment at all). (p. 33)

Ultimately, the IHE concluded: "Any conclusions drawn from this overview of SRs remain tentative. Given the methodological problems of the available primary research, it is

difficult to draw strong conclusions about the effectiveness of SOT programs using various CBT approaches for such a heterogeneous population” (p. iv).

In 2011, the Swedish Council on Health Technology Assessment HTA (identified by its Swedish acronym of SBU) was assigned by the Swedish government to conduct a SR of “Medical and Psychological Methods for Preventing Sexual Offenses Against Children.” This review provided an extensive and detailed report of the existing SRs and meta-analyses of sexual offender treatment. The SBU found that in examining seven previous SRs:

...the debate in the scientific literature on what sexual offender treatment interventions works for adult male sexual offenders remains divided...Although some of the selected SRs suggest a positive effect for CBT on both sexual and general recidivism, methodological problems, inconsistency results, and a lack of high-quality primary research studies included in the SRs raise uncertainty about which of the available approaches work for adults male sex offenders. (p. 32)

The SBU SR stated that the available evidence provided evidence for some effectiveness of treatment in reducing sexual offense recidivism, noting that existing SRs showed small reductions in such recidivism for sexual offenders after undergoing CBT. However, the SBU concluded:

Major deficiencies were found in the evidence concerning effective medical and psychological interventions for individuals that have committed sexual offences against children. This is serious since the purpose of this treatment is to prevent new offences... For adults that have committed sexual offenses against children the scientific evidence is insufficient for determining which treatments that could reduce sexual reoffending. The lack of evidence concerns both benefit and risk for pharmacotherapy and psychological treatment programmes. (pp. 4–5)

In addition, the SBU concluded: “Sexual offender treatment programs neither cure sexual offending nor guarantee a complete cessation of offending, and they represent one element in a comprehensive risk management strategy designed for convicted adult male sex offenders...Not all sexual offender treatment interventions and programs are effective in reducing sexual/non-sexual recidivism in this population” (p. 37).

Most recently, Langstrom et al. (2013) conducted a systematic review of medical and psychological interventions of sexual offenders who committed sexual offenses against children. They reviewed 1,447 abstracts, retrieved 167 full text studies, and finally included eight (8) studies with low to moderate risk of bias. They concluded that there was “weak evidence for interventions aimed at reducing offending in identified sexual abusers of children...For adults, evidence from five trials was insufficient regarding both benefits and risk with psychological treatment and pharmacotherapy.” Langstrom et al. noted: “Despite severe consequences for victims and society, this systematic review identified

remarkably little research of acceptable quality on individual-level prevention of child sexual abuse” (p. 3). Of more recent studies, they identified only one RCT involving offenders who had sexually abused children. Overall, effectively, they concluded that no evidence exists of the effectiveness of cognitive-behavioral treatment or pharmacological interventions, noting “the remarkable lack of quality research studies in sexual abuser of children...” (p. 4). They expressed the hope that such treatments might be found to have some positive effects if and when large, methodologically rigorous, studies are implemented. However, they are also warned of the potential consequences of denying treatment to offenders for whom it might have benefit and, conversely, of providing unproven treatment that might increase the risk for future sexual offending.

### Meta-Analyses of Sexual Offender Treatment

Kendall et al. (2004) identified that meta-analytic statistical techniques could be useful because they synthesize results across multiple studies by converting the results of each investigation into a common metric (usually, the “effect size”). Such a method increases the potential power of experimental studies by combining the results of a number of investigations (typically with relatively small numbers of subjects) to increased statistical power to determine if there is a trend or clear effect over the aggregated studies. As noted previously, such an effect size (ES) provides a measure of the magnitude or “strength” of the experimental effect; in and of itself, the effect size is not an indication of causality.<sup>3</sup> The outcomes of different treatment comparisons can then be compared with respect to the magnitude of difference reflected in such statistics. As noted, a key issue that arises in meta-analytic studies has to do with whether studies of inferior methodological quality should be included or omitted. Kendall et al., among others, agree that it is important to eliminate those studies whose quality does not allow them to contribute meaningful findings as a result of basic inadequacies in research design. A recommendation that a particular approach is effective or more effective than an alternative approach cannot be determined if that recommendation is based on inadequate research:

If the research evidence is methodologically unsound, it is insufficient evidence for a recommendation; it remains inadequate as a basis for either supporting or refuting treatment recommendations, and therefore it should not be included in cumulative analyses...Caution is paramount in meta-analyses in which various studies are said to provide evidence that treatment is superior to

<sup>3</sup>The effect size (ES) is typically derived by computing the difference between the mean of the treatment group and the control group at post-treatment and then dividing the difference by the pooled standard deviation of the 2 group.

controls. The exact nature of the control condition in each specific study must be examined...Meta-analyzers cannot tabulate the number of studies in which treatment is found to be efficacious in relation to controls without examining the nature of the control condition. (Kendall et al., 2004, p. 37)

Thus, a critical issue for interpreting the results of any MA is reliant on the quality of the specific investigations that compose the MA; the inclusion of methodologically weak or inadequate studies limits any conclusions drawn from that MA.

Hall (1995) conducted the first MA of sexual offender treatment studies that appeared after the review by Furby et al. He studied only studies that include some comparison group and utilize recidivism rates between each treatment and comparison groups (alternative or no treatment) as the outcome measure. Of 92 studies available, 80 were eliminated from consideration because they had fewer than 10 subjects, lacked a comparison or control groups, or did not report sexual offense recidivism rates. For the twelve studies Hall deemed adequate for evaluation, his MA revealed a “small” but statistically significant overall treatment effect ( $r=0.12$ ); however, the treatment ES across studies were significantly heterogeneous. Effect sizes were significantly greater in studies of outpatients than for studies of institutionalized offenders, potentially an effect of the severity of participant psychopathology. Hall concluded that comprehensive cognitive-behavioral treatments (CBT) showed better outcomes than purely behavioral treatments. Hall noted conservatively that 36 % of those eligible for participation in sexual offender treatment were typically excluded from participating in treatment: “In general, the most pathological participants were excluded from samples (e.g. extensive offense history, psychotic, organic brain syndrome, denied offenses, management problem in prison, withdrew from treatment program)” (p. 803). Consequently, he wrote, “Thus, the currently reviewed treatments may be less effective with the most pathological sexual offenders” (p. 808). In addition, it was found that 1/3–2/3 of participants refused hormonal treatments, while refusal and dropout rates for CBT were found to be about 1/3 of eligible participants. Hall concluded that his meta-analysis results suggested “the effect of treatment with sexual offenders is robust, albeit small...” (p. 808).

In 2002, Hanson et al. published the first report from a Collaborative Outcome Data Project (CODP) established by the Association for the Treatment of Sexual Abusers (ATSA). They noted that a primary objective of the CODP was “to promote professional debate concerning the relative quality of treatment outcomes studies for sex offenders” (p. 173). Hanson et al. conducted a MA that combined data on 43 psychosocial treatment programs involving 9,454 sexual offenders who were either assigned to either sexual offender treatment, were untreated or were provided other

interventions. Of the treatments reviewed, 23 were offered in institutions, 17 in the community, and three in both settings; the major sponsor of the programs studied were departments of corrections ( $n=26$ ). The treatment studies considered were delivered between 1965 and 1999; only 23 studies had been published in either a book or a journal and approximately ½ were from the USA (Canadian samples made up another 16 studies). Approximately 80 % of the sexual offenders received “current” treatment (defined as CBT offered after 1980 or behavioral, other psychotherapeutic, and/or mixed treatments delivered between 1998 and 2000). The median length of the follow-up was 46 months for both treatment and comparison groups or just less than 4 years. Sexual offense recidivism was defined by reconviction in 8 studies, rearrest in 11 studies, while 20 studies used broad definitions (e.g., including parole violations, readmissions to institutions, unofficial community reports, or all of these). Thirteen programs reported outcome only on sexual recidivism, five reported only on general recidivism outcomes, and 25 reported on both.

Hanson et al. (2002) grouped the studies that they considered into several categories. The first category was based on the strongest method for comparing different comparison groups, random assignment; in these studies, persons were randomly divided into groups who received treatment and those who did not. The second category that Hansen et al. considered relative to treatment outcome was referred to as “incidental assignment” to treatment. In these studies, sexual offenders who were provided with sexual offender treatment were compared to varied comparison groups that were “created” from some pool of sexual offenders available to investigators. Per the 2002 review, such control groups were matched in various ways to those who received treatment. Thus, the control samples were selected according to varied criteria in specific studies, including offenders who (1) had been released before the implementation of the treatment program (5 studies); (2) had received no treatment or received treatment judged to be lower in quality, due to such administrative reasons such as too little time remaining on their sentences (5 studies); (3) matched from archives of criminal history records (3 studies); or (4) had received an earlier version of the treatment (2 studies). Hansen et al. labeled these 17 studies as involving “incidental assignment” because it was theorized or believed that there was no “obvious” or “a priori” expectation that the treated and untreated offenders should differ in risk and thus have no “obvious” bias in group composition. In addition, an additional category of subjects considered by Hanson et al. was those deemed “assignment based on need,” where treatment [was] given to those assessed as requiring treatment. Finally, they compared any treatment attendance (including dropouts), treatment completers to treatment dropouts, and treatment dropouts to treatment refusers.

Results of the MA of 5,078 treated and 4,376 untreated sexual offenders found that the *unweighted* averages across all studies indicated the sexual offense recidivism rates were lower for the treated groups (12 %) than for the comparison groups (17 %). The chief conclusion drawn by Hanson et al. (2002) from these results was that “there was a *small* advantage for the treated versus the untreated offenders,” and this finding was statistically significant (p. 181, emphasis added). However, this overall analysis included the results of sexual offender treatment for juvenile sexual offenders, which, were more likely to show a positive outcome for offenders (albeit largely for multiple trials of one particular method of treatment). Further, Hanson et al. noted considerably variability across studies, with treatment effects much larger in studies that had not been published. Of significance, when only the four methodologically superior RCT studies were examined, *no* treatment effect was found. In contrast, evidence for treatment effectiveness was found only in the results from the incidental assignment studies which, on average, showed statistically meaningful reductions in sexual offense recidivism, albeit with more variability than expected by chance. Perhaps oddly, Hanson et al. then combined the nonsignificant findings from the methodologically superior random assignment studies with the significant effects of the 17 methodologically inferior incidental assignment studies of “current” treatments and, on this basis, concluded that “current” treatments were associated with significant reductions in both sexual (from 17.3 to 9.9 %). Thus, the results that Hanson et al. found for treated sexual offenders over a mean 4-year follow-up (12 %) were comparable to the rates of sexual offender recidivism that had been found for largely untreated sexual offenders in the two MA of risk factors for sexual offender recidivism; respectively, Hanson and Bussiere (1998) and Hanson and Morton-Bourgon (2004, 2005) indicated that the mean 5-year rates of sexual reoffending for the two large samples of almost exclusively untreated sexual offenders were 13 and 14 %. Thus, the results for sexual reoffending that Hanson et al. (2002) reported for the treated sexual offenders in their treatment MA were equivalent to the rate of much larger samples of untreated sexual offenders in varied comparison groups.

According to Hanson et al. (2002), other findings from their meta-analysis included:

studies comparing treatment completers to dropouts consistently found higher sexual and general recidivism rates for the dropouts, regardless of the type of treatment provided. Even in studies where there was no difference between the treatment group and the untreated comparison groups, the treatment dropouts did worse. (p. 182)

Although it was determined that dropouts were approximately twice as likely to sexually reoffend, in their analysis of the untreated “comparison” groups, Hanson et al. did not account for or control for such dropout effects; there was no

analysis of intent to treat. Hanson et al. also reported “offenders who refused treatment were not at higher risk for sexual recidivism than offenders who started treatment” (p. 182), a finding that conflicts with that of the other reviews. Interventions that were viewed as “current” treatments were found to be associated with greater reductions in recidivism. In contrast to what Polizzi et al. (1999) reported, Hanson et al. (2002) found that both institutional and community treatments showed equal results regarding the degree of recidivism associated with the different types of programs.

However, a key finding by Hanson et al. was that “Offenders referred to treatment based on perceived need had significantly higher sexual recidivism rates than the offenders considered not to need treatment” (p. 182). The odds ratio was 3.4 (with an outlier study removed), and there was no significant variability, indicating that this was a robust phenomenon. Hanson et al. concluded: “Studies that compared sex offenders who ‘needed’ treatment to less needy offenders consistently found *worse outcomes for the treatment group*. It appears that evaluators are better able to identify high risk offenders than to change them” (p. 187, emphasis added).

Hanson et al. offered a considerably measured conclusion to their MA, writing, “We believe the balance of available evidence suggest that current treatments reduce recidivism, *but that firm conclusions awaits more and better research*” (p. 186, emphasis added). They indicated that when random assignment and incident assignment studies were combined, there was a *reduction* in sex offense recidivism and “These reductions were not large, but they were statistically reliable and large enough to be of practical significance” (p. 187, emphasis added). They concluded that the absolute reduction in recidivism rates was *modest* even among the better-designed studies of current treatments and that no treatment effect was found among the best-designed studies. The results reported by Hanson et al. suggested that treatments that appeared effective for adult sexual offenders were more “current” programs providing some form of CBT. They also reported no “setting” effect for sexual offender treatment; both institution-based and community-based programs for adults were found to be associated with reductions in sexual recidivism of adult sexual offenders. Hanson et al. did not identify specific interventions that provided guidance on the effectiveness of any sexual offender treatment interventions for different types of sexual offenders (e.g., “rapists” vs. child molesters or mixed offenders). Finally, the authors concluded that the results of their meta-analysis provided little direction in terms of how to improve current practice.

Several years later, Losel and Schmucker (2005, 2008) conducted another meta-analysis of both published and unpublished sexual offender “controlled” outcome studies available as of 2003, involving either psychosocial **or** biological



treatments. Losel and Schmucker (2005, 2008) reviewed 69 studies with more than 22,000 subjects; unpublished investigations comprised 36 % of the study pool. Of those studies, approximately 18 % were analyses of biological treatment (e.g., hormonal treatment and surgical castration). Of the remaining studies identified, 46 % were of CBT and 18 % were “classical” behavioral psychosocial treatments. Per their 2008 paper, however, “60 % of the identified studies used clearly non-equivalent control groups” (emphasis added, p. 16, emphasis added). About one-third of these studies had been reported since 2000, but the actual program implementation started earlier (e.g., in the 1990s). Approximately 70 % of these studies were conducted in North America. The definition of recidivism varied across studies: arrest (24 %), conviction (30 %), and charges (19 %). Recidivism was recorded after an average follow-up period of more than 5 years. Sexual recidivism outcomes were reported in 74 of the 80 comparisons. Although most treatments were specifically designed for sexual offenders, the authors found it difficult to rate whether treatment was implemented reliably, as three-quarters of the studies did not provide information on program integrity. Residential (institutional) treatment was somewhat more frequent than outpatient treatment; approximately one-half of the studies were implemented in an institutional setting. Although a group format was most frequently used, almost 50 % of the programs included at least some individualized treatment. Sexual offenders who received treatment participated voluntarily in most studies; however, 30 % of the comparisons referred to offenders who were at least partially obliged to attend treatment. In more than 50 % of the primary research studies, the authors were affiliated with the treatment program that was implemented (raising the question of allegiance issues).

Methodologically, approximately one-third of the comparisons contained fewer than 50 sexual offenders as subjects, while 46 % included 100 subjects or less. Only seven comparisons were based on random assignment and just six studies received a Level 5 designation of the Maryland scale. Conversely, 60 % of the treatment comparisons were at Maryland Scale Level 2 such that treatment and control group could not be considered equivalent; in an additional 24 % of studies, the equivalence of the two groups was simply assumed by the original investigators. In approximately 24 % of the comparisons, the control group consisted of treatment refusers.

Of note, when recidivism rates were calculated for treated and comparison subjects initially using unweighted averages, a treatment effect was found. However, when weighted averages were utilized (e.g., taking into account relative numbers of persons in treatment and comparison groups), “the difference in recidivism rates vanished completely (11 % each for treated and comparison participants),” (p. 127) although the authors dismissed this issue.

Subsequently, Losel and Schmucker (2005) utilized the mean effect size, which showed that the majority of effects were positive; they then calculated odds ratios.<sup>4</sup> Including treatments for **both** biological and psychological treatments, the mean odds ratio for sexual offense recidivism was 1.7, which was highly significant so that the absolute difference in sexual recidivism between the “any” treatment group (e.g., biological **and/or** psychosocial intervention) and the heterogeneous control groups was 6 %. The rate of sexual recidivism for the overall treated groups (e.g., psychosocial **and/or** biological treatments) was 11 % (the control groups showed an average sexual offense recidivism rate of 17.5 %). However, as with the Hanson et al. (2002) meta-analysis, there were considerable, statistically significant differences in effect sizes across the comparisons studied indicating considerable heterogeneity beyond what would be expected by chance. Large effects of treatment were found more frequently in studies with small sample sizes. Of note, medical treatments (e.g., hormonal treatments or surgical castration) were found to have considerably higher effect sizes than those for CBT (e.g., 2–10 times larger, respectively). Of psychosocial interventions, only cognitive-behavioral and “classic behavior therapy” generally showed a significant impact on sexual offense recidivism. An important issue relative to how the Losel and Schmucker (2005) results are typically discussed is that the widely reported reduction in sexual offense recidivism as a result of “treatment” includes the combined results of both biological and psychosocial interventions; the reported 6 % (or “37 %” relative reduction) for sexual offense recidivism resulted from a comparison of *both biological and psychosocial interventions* and would not apply to just psychosocial interventions. Losel and Schmucker (2005) reported that after removing those studies involving surgical castration (which had the highest effect size of all treatments), the effect size for treatment generally decreased and the relative sexual offense recidivism “drop” for nonsurgical treatments decreased by approximately one-third.

In contrast to the Hanson et al. (2002) findings, more modern sexual offender treatment programs were no more effective than earlier programs. Losel and Schmucker and Losel (2008) noted that “Some recent evaluations have revealed rather small or no positive effects...As follow-up one of the soundest evaluations has also found no positive effect...” (p. 136). That is, as with the Hanson et al. (2002) meta-analysis, Losel and Schmucker’s (2005) review did not include the final results of the Marques, Wiederanders, Day, Nelson, and Ommeren (2005) study that showed no effect of sexual

<sup>4</sup>Odds ratio is the *odds* of being in one group relative to the *odds* of being in a different group and is used with categorical variables (e.g., treatment, no treatment; reoffend, not reoffend). The odds ratio is not the chance or likelihood but the ratio of the odds, not the percentages. The odds of an event occurring is the probability of an event divided by the probability of an event not occurring.

offender treatment utilizing a RCT. No difference was found between group and individual treatment programs by Losel and Schmucker (2005). Per their 2008 report, “Only outpatient treatment showed a significant effect” (p. 13). The odds ratio for institutional sexual offender treatment was considerably lower than that for outpatient treatment and *not* significant. Thus, similar to Polizzi et al.’s findings, prison-based programs or hospital-based sexual offender treatment programs showed outcome results that indicated little difference between sexual offender treatment participants and nonparticipants. No significant difference was found for treatments for adult and those for adolescent sexual offenders. In their two reports, Losel and Schmucker (2005, 2008) found that only sexual offender treatment programs involving *voluntary participation showed a significant effect*; programs that involved “a more or less coerced treatment” *did not show a significant effect* (2008, p. 13). Further, “Whether treatment was terminated regularly or prematurely had an impact on sexual recidivism. Whereas “regular” completers showed better effects than the control groups, dropouts did significantly worse. Dropping out of treatment doubled the odds of relapse...” (2005, p. 132). However, effect sizes that referred to treatment completers revealed considerable heterogeneity. Various methodological differences related to sample size and design quality were identified in comparisons between those who completed and those who did not complete sexual offender treatment. However, those differences were neither uniform or provided clarity as to their implications. However, those differences were neither uniform or provided clarity as to their significance.

Losel and Schmucker (2005) noted that their analyses repeatedly indicated problems of confounded moderators. Consequently, they tested to see to what degree the treatment effects were confounded with methodological and other study characteristics. Methodological characteristics accounted for a considerable amount of variance in outcome for treatment (e.g., 45 %). Of these methodological characteristics, general characteristics of treatment were most important, including specificity of treatment for sexual offenders, involvement of authors in the program, and a group format contributed to a 9 % increase in explained effect size variance. Thus, for example, treatment studies in which the study author(s) was in some way involved in the program delivery more likely showed significant treatment effects, but programs that were evaluated by independent researchers did not; this strongly suggests the so-called allegiance effects. They concluded that “... methodological factors play an important role and seem to be confounded with treatment and offender characteristics. This problem of confounded moderators is rather general and difficult to solve...” (p. 138).

In their conclusions, Losel and Schmucker (2005) stated: “Bearing the methodological problems in mind, one should draw very cautious conclusions from out meta-analysis. The most important message is an overall positive and signifi-

cant effect of sex offender treatment” (p. 135); however, it appears that this conclusion was *inclusive* of surgical castration and hormonal treatments that were more effective than psychosocial treatments. As the authors pointed out in 2005, differences in treatment and comparison groups “most probably” related to their inclusion of both medical and psychological modes of treatment because, “The average effect of physical a treatment is larger than that of psychosocial programs” (p. 135). Further, the authors also cautioned, “One must bear in mind that outcomes of treatment often decline when model projects are transformed into routine practice” (p. 137). In 2008, Losel and Schmucker wrote:

The size of the [treatment] effect is small to moderate... However, the evidence is based on studies that mostly apply a weak methodological standard. *Restricting the analysis to a few randomized trials shows a comparable mean effect but it does not render it statistically significant...* Obviously we need more high quality evaluations on the whole range of sexual offender treatment to come to unequivocal conclusions. (p. 17, emphasis added)

In their 2008 paper, Schmucker and Losel noted methodological study characteristics explained the largest proportion of variability in effect size variance; they concluded that “Overall, findings are promising but more differentiated evaluations of high quality are needed” (p. 1).

Hanson et al. (2008, 2009) completed an updated but somewhat different MA of sexual offender treatment relative to the 2002 paper. At the outset, they stated:

All reviews have concluded that more and better studies are needed. Few studies have used strong research designs (i.e. random assignment), and there are even fewer studies with strong research design examining interventions consistent with contemporary standards. Consequently reviewers are forced to consider whether the less than ideal studies are “good enough.” (p. 866)

Hanson et al. (2009) considered sexual offender treatment in the specific context of treatment of general criminal offenders (e.g., a criminological perspective) and not that of psychotherapy outcome research per se. More particularly, they examined the utility of the risk, need, and responsivity (RNR) model (e.g., Andrews & Bonta, 2006) which states that “... treatments are most likely to be effective when they treat offenders who are likely to reoffend (moderate or higher risk), target characteristics that are related to reoffending (e.g. criminogenic needs), and match treatment to offenders’ learning styles and abilities (responsivity; cognitive-behavioral interventions work best)” (p. 866). Thus, Hanson et al. (2009) addressed the question of whether the principles of effective general criminological interventions also applied to the psychological treatment of sexual offenders. They also examined whether different results were found in better-quality studies than in studies that met only minimum standards of acceptability (e.g., weak designs) and relied on the Collaborative Outcome Data Committee (CODC, 2007a, b). Thus, they

noted that a “strong” study would be one that involved “a well-implemented random assignment study (e.g. uncorrupted random assignment, 5 or more years of follow-up, sample size >100, < 20 % attrition, no preexisting differences between the groups found post hoc)” (p. 869).

Hanson et al. (2009) found that no studies reported findings for different intensities of treatment services within the same setting. Studies were therefore coded as adhering to the risk principle if their treatment group was higher risk than average for sexual offenders. Programs were considered to meet the need principle if the majority (51 %) of the treatment targets were criminogenic needs. It was assumed that CBT programs adhered to the responsivity principle. Of the 23 studies accepted for analysis, 14 were published and 9 were not. Most of the studies were Canadian (12) or American (5). Only 19 studies focused on adult sexual offenders. Of the 23 programs, 10 were offered in institutions and 11 in the community; 16 programs were sponsored by corrections. In total, 22 studies examined 3,121 treated sexual offenders and 3,625 untreated sexual offenders.

Regarding results, the sexual offense recidivism rate for treatment groups had an unweighted mean of 11 % and for comparison groups 19 %. The odds ratio for sexual offense recidivism with a fixed weighted mean was 0.77, but there was more variability than would be expected by chance. For 22 studies examining the sexual recidivism rate, results from both fixed-effect and random-effect analyses indicated significantly lower sexual recidivism rates in the treatment groups than in the comparison groups. However, of note, the combined rate of sexual and other violent offenses was not significantly lower for treatment groups relative to comparison groups, while general recidivism rates were lower for the treatment groups. Again, Hanson et al. (2008) found no differences as to whether treatment was delivered in the community or in institutions; recent treatments were found to be more effective. The treatment effects on both sexual and violent recidivism were smaller in the good-quality studies than the weak studies.

Only two studies were each rated as following two of the three RNR principles. Analyses found that programs were more effective when they targeted criminogenic needs or delivered in a manner that was likely to engage sexual offenders (e.g., responsivity via CBT). Support was demonstrated for both the need and responsivity principles. However, Hanson et al. (2009) did not find that the risk principle was supported; that is, available program results were *not* significantly more likely to be effective when they treated offenders who were rated as at higher risk to reoffend. Overall, regarding sexual offense recidivism, results indicated that the relative effectiveness of sexual offender treatment increased according to the degree that treatment adhered to the RNR model, except for the risk principle. However, for the 10 studies that examined both sexual and violent recidivism as the outcome variable, there was no significant difference based on adherence to the RNR model.

Hanson et al. (2009) reported that the sexual and general recidivism rates for the treated sexual offender were lower than for comparison groups (based on unweighted averages). However, for a median follow-up of 4.7 years, the results for sexual offense recidivism that Hanson et al. (2008) found for treated sexual offenders (11 %) were again comparable to the rates of sexual offender recidivism that had been found for largely untreated sexual offenders in two MA of risk factors for sexual offender recidivism [respectively, Hanson and Bussiere (1998) and Hanson and Morton-Bourgon (2004, 2005) indicated that the mean 5-year rates of sexual reoffending for the two large samples of almost exclusively untreated sexual offenders were 13 % and 14 %]. Thus, the results that Hanson et al. (2009) found for the treated sexual offenders in their treatment MA were largely equivalent to the rate of much larger samples of untreated sexual offenders. Unfortunately, Hanson et al. (2009) noted that not one “strong” study of sexual offender treatment could be identified per CODC criteria. Hanson et al. concluded:

Confidence in the findings, however, must be tempered by the weak research designs. Even after excluding the worst 80 % due to inadequate study quality, still only 5 of the remaining 23 studies were rated as good according to the CODC guidelines (18 were weak). The effects tended to be stronger in the weak research designs compared to the good research designs. *Reviewers restricting themselves to the better-quality, published studies...could reasonably conclude that there is no evidence that treatment reduces sex offense recidivism.* (p. 881, emphasis added)

In 2012, Dennis et al. (2012) authored another CR of psychosocial interventions for adults who had been sexually offended. In keeping with similar reviews, their selection criteria involved randomized trials comparing psychological interventions with standard care or another psychological therapy provided to adults in either institutional or community settings for sexual behavior. The authors stated: “While this review adopts the Cochrane principles of examining only evidence from RCTs, we do so without any apology, in the belief that other types of trial evidence are likely to inflate the positive findings for the intervention” (p. 27). They found ten studies that met their criteria involving a total of 944 male adults, of which four compared CBT with no treatment or wait list control and one which compared CBT with standard care. Four other studies involved behavioral programs and one study compared psychodynamic intervention with probation. For CBT, Dennis et al. reported: “The result of comparing reconviction for sexual offences between conditions was not statistically significant” (p. 23). Similarly, for the psychodynamic intervention, there was no difference in rate of sexual rearrest at 10-year follow-up. Thus, the investigators reported: “The main finding of this systematic review is that there was no evidence from any of the trials in favour of the active intervention in a reduction of sexual recidivism—the primary outcome” (p. 25). For both CBT and psychodynamic interventions with meaningful follow-up

data, Dennis et al. noted that "...neither showed any benefit for the intervention. Thus, neither...appeared to reduce sexual recidivism" (p. 35). They stated:

The inescapable conclusion of this review is the need for further randomized controlled trials. While we recognize that randomisation is considered by some to be unethical or politically unacceptable (both of which are based on the faulty premise that the experimental treatment is superior to the control—this being the point of the trial to begin with), without such evidence, that area will fail to progress. Not only could this result in the continued use of ineffective (and potentially harmful) interventions, but it also means that society is lured into a false sense of security in the belief that once the individual has been treated their risk of reoffending is reduced. Current available evidence does not support this belief. Future trials should concentrate on minimizing risk of bias, maximizing quality of reporting and including follow-up for a minimum of five years 'at risk' in the community. (p. 3)

### Summary of Reviews of the Effectiveness of Sexual Offender Treatment

In his meta-analysis, Hall (1995) found a "small" statistically significant effect for sexual offender treatment for highly screened subjects and concluded that such treatments might be less effective with more "severe" sexual offenders. He found that outpatient sexual offender treatment appeared more effective than institutional treatment. Like Hall, Hanson et al. (2002) concluded from their meta-analysis that "there was a *small* advantage for the treated versus the untreated offenders" (emphasis added) and this finding was statistically significant. However, they found that if only those studies that utilized random assignment of sexual offenders to treatment were examined, no treatment effect was apparent. In addition, Hanson et al. identified a "robust" finding that sexual offenders referred to sexual offender treatment based on "perceived need" (e.g., likely higher-risk sexual offenders) had substantially higher sexual offense recidivism rates than those with less need and concluded such offenders were less responsive to sexual offender treatment. Hanson et al. concluded that "We believe the balance of available evidence suggest that current treatments reduce recidivism, but that firm conclusions await more and better research" (p. 186). Losel and Schmucker (2005, 2008) also concluded from their meta-analysis that a majority of treatment studies (a combined set of biological and psychosocial treatments) suggested a positive effect for sexual offender treatment. Losel and Schmucker found that "obligatory participation" in treatment resulted in no treatment effect. Thus, according to the conclusions of Hanson et al. and Losel and Schmucker, both higher levels of need and mandated participation were associated with no treatment effectiveness. In contrast to Hanson et al., Losel and Schmucker found no differences between "current" and older sexual offender treatment pro-

grams. Another difference between the Hanson et al. (2002) and Losel and Schmucker (2005, 2008) meta-analyses was that the latter identified a trend for lower effectiveness in institution-based programs, whereas the former did not. This was similar to what Polizzi et al. (1999) concluded, namely, that "...the evidence is not strong enough to support a conclusion that [prison-based programs] are effective." Relative to the effectiveness of sexual offender treatment, Losel and Schmucker (2005) concluded "one should draw very cautious conclusions from our meta-analysis" (p. 135), while Hanson et al. (2002) opined "firm conclusions await more and better research" (p. 186). Hanson et al. (2008) found that unweighted rates of sexual offense recidivism were lower for the treated sexual offender than for comparison groups (based on unweighted averages). They did not find support for the risk principle—treatment was not more effective with more high-risk sexual offenders—comparable to their earlier finding about perceived need. They found no studies of adult sexual offenders that targeted risk, needs, and responsivity. Hanson et al. (2009) concluded that much could be done to increase confidence in outcome studies on sexual offender treatment. The IHE report interpreted the available data to suggest that sexual offense treatment had been shown to provide "small" reductions in sexual offense recidivism. However, the SBU found that the scientific evidence was insufficient for determining whether that such treatment could reduce sexual offending. Finally, one of the most recent, most rigorous reviews of treatment, for both CBT and psychodynamic interventions with meaningful follow-up data, Dennis et al. (2012) noted that "...neither showed any benefit for the intervention. Thus, neither...appeared to reduce sexual recidivism" (p. 35). This finding was also confirmed by Langstrom et al. (2013).

A key issue identified by most SRs and MAs was the dearth of high-quality research methodology in the available studies. Losel and Schmucker (2005, 2008) noted that only 6 of 69 studies available were considered to meet the Maryland Level 5 standard; as did most recent and prior reviewers, they emphasized that most of the studies included in their meta-analysis were of poor methodological quality. In addition, as noted, Losel and Schmucker did not include the final report from the Marques et al. study, the only contemporary RCT for psychosocial treatment for sexual offenders. Further, Hanson et al. (2002, 2009) pointed out that the treatment effects on sexual recidivism were, in fact, smaller in the good-quality studies than in the weak studies, suggesting that it was low-quality studies that inflated the already small positive outcome. Similarly, Losel and Schmucker (2005, 2008) showed that larger effects of treatment were found more frequently in studies with small sample sizes. In addition, they reported that the largest treatment effect was found for Maryland Level 3 studies in which the equivalence of comparison groups was assumed; thus, their results were

similar to those of the Hanson et al. meta-analysis where only incidental assignment (“assumed equivalence”) showed an effect for sexual offender treatment. Hanson et al. (2009) also found that approximately 80 % of included studies were characterized by weak research designs and that more positive results were associated with more methodologically flawed studies. They concluded that if only higher-quality studies were considered, it would be reasonable to conclude that there was no evidence that psychosocial treatment decreased sexual offense recidivism. Langstrom et al. (2013) noted “the remarkable lack of quality research studies in sexual abuser of children...” (p. 4). Thus, all of the systematic reviews and meta-analyses, to date, have concluded that the field of sexual offender treatment was characterized almost exclusively by poor-quality methodology (primarily lack of RCTs and/or small numbers of subjects), that little information was provided about program integrity, and/or that no or minimal information was found as to elements of sexual offender treatment that were related to the outcome of the interventions. There was unanimity in the SRs and MAs that there was a strong need for more research of sexual offender treatment characterized by significantly higher scientific rigor.

Thus, at best, if one considers quasi-experimental research studies (consistently viewed across reviews as methodologically weak) only, a relatively small effect regarding decreased recidivism is sometimes demonstrated for treating low- to moderate-risk sexual offenders.<sup>5</sup> However, if only higher-quality, methodologically rigorous research studies (such as RCTs) are considered, from a scientific perspective, no definitive evidence has yet been presented by any researchers that psychotherapy is associated with any substantive reduction in sexual offense recidivism.

### **A Critical Perspective on the Results of Existing Systematic Reviews and Meta-Analyses of Sexual Offender Treatment**

From an empirical perspective, no clear evidence of a scientific nature has yet been found via rigorous scientific study that psychotherapy is associated with a consistent, meaningful effect in the reduction of sexual offense recidivism; no substantive or strong proof yet exists that psychosocial inter-

<sup>5</sup>While not intended as a study of sex offender treatment outcome, Helmus (2009) in her meta-analysis of factors in base rate variability in sexual offender samples found that the effects of sex offender treatment (among other variables) on sex offense recidivism. She compared offenders who completed treatment, dropped out of treatment and those who did not attend treatment. There was no difference in the rate of sexual reoffending relative to treatment participation; similarly, there was no difference between offenders who started treatment and those who completed treatment.

ventions “work” in reducing future sexual offending at this time. In their own MAs, Hall (1995), Hanson et al. (2002), and Losel and Schmucker (2005, 2008) acknowledged that the effect sizes obtained in their meta-analyses for psychosocial treatments of sexual offender were “small”—despite the inclusion of and primary reliance on studies with acknowledged poor quality (e.g., problematic control groups) as well as relying on offender samples that likely enhanced the probability of obtaining positive outcome found for treatment conditions (e.g., predominantly low-risk offenders with few additional psychiatric or psychosocial issues). Langstrom et al. (2013) came to the conclusion that no evidence exists of the effectiveness of cognitive-behavioral treatment relative to sexual offender treatment for child molesters. Dennis et al. (2012) concluded that neither CBT nor other psychosocial interventions with meaningful follow-up data showed any benefit for the intervention; they did not appear to reduce sexual recidivism. Thus, to date, no investigator or scientific authority has produced or found what he or she consider to be a rigorous scientific evidence for and/or concluded that sexual offender treatment has been demonstrated to be “very” or “greatly” effective for sexual offenders. Rather, at best, if quasi-experimental research studies (e.g., incidental assignment) are included, the current findings indicate that such treatment might be “somewhat” “slightly” effective with voluntary, low-need/low-risk sexual offenders who are volunteers and are not “mandated” for participation in sexual offender treatment. Each available SR or meta-analysis has commented on the poor quality of the existing treatment outcome literature, and each has strongly recommended the need for additional studies, better designed with strong methodological qualities, particularly random assignment of subjects to treatment and control groups. In 1997, Hanson wrote: “Meta-analyses rely on the quality of the original studies, and skeptics can claim that there is an insufficient number of well-controlled studies to justify meta-analytic review” (p. 139). Even at present, it appears that an insufficient number of such studies exist.

Further, as Berliner wrote of the Hanson et al. meta-analysis at the time of its publication in 2002:

The conclusions of this study, however, should not be exaggerated nor considered the final word on sex offender treatment. The studies measure reductions in recidivism and its elimination. *The effect sizes for recidivism reduction are not large, thus there will still be failures, the cost of which will be born by victims. It is not at all clear that these results can be generalized to the highest risk offenders. Even if they could be applied to these offenders, a moderate effect size reduction would still mean that high-risk offenders continue to be dangerous.* (p. 196, emphasis added)

Such comments are equally applicable to the subsequent meta-analysis by Losel and Schmucker (2005, 2008). [In addition, as will be reviewed in more detail, if one was to factor in (1) the sexual offenders who denied consideration

for treatment initially and/or (2) those who refused participation, and/or (3) those who dropped out of or were demitted from treatment, the minimal effects of sexual offender treatment outcome studies would almost certainly be even smaller or potentially nonexistent.] Beggs (2010) pointed out something obvious that is largely not addressed in the available sexual offender treatment literature—"the fact that residual post-treatment reoffending occurs at all indicates that not everyone who completes the same treatment will derive the same benefit" (p. 369). Of note, recent systematic reviews by the IHE, the SBU, Dennis et al. (2012), and Langstrom et al. (year) concluded that there was, at best, slight and, at worse, *no* scientific evidence of the effectiveness of sexual offender treatment. At this date, at best, no reviewer has concluded there is strong or even moderate empirical support for the effectiveness of psychosocial treatment for sexual offenders relative to reducing future sexual offense recidivism; at worse, the more consistent conclusion has been that there is no strong empirical evidence for the effectiveness of sexual offender treatment. This finding stands in marked contrast to the available scientific literature on psychotherapy more generally; as Westen et al. (2005) wrote: "The data are now clear that virtually anything researchers do for 10–20 sessions with patients that they firmly believe will be efficacious in fact leads to better outcomes than experimental conditions not intended to work..." (p. 428). Yet, despite what one presumes to be the best intentions of those providing sexual offender treatment, from the perspective of a "hard" outcome that matters most to stakeholders—reducing future sexual offending—there is little indication that psychosocial interventions are efficacious for participating sexual offenders.

In their article "Psychotherapy on Trial," Arkowitz and Lilienfeld (2006) enumerated a variety of reasons as to how clinicians can be misled into concluding that an ineffective psychotherapy is in fact efficacious. They identified that several phenomena that can make psychosocial interventions "appear" effective as justification as to why scientific psychotherapy outcome research is necessary: spontaneous remission, placebo effects, regression to the mean, treatment/programming interferences, selective attrition, effort justification, and demand characteristics. Each of these factors can be viewed as applicable to sexual offender treatment as well as general psychotherapy research. Since Furby et al.'s (1989) initial review of sexual offender treatment outcome, a number of general and specific methodological concerns have been raised regarding interpretations of the existing reviews of sexual offender treatment, including the select meta-analyses typically relied upon as the basis for the claims that sexual offender treatment "works." Beyond the failure of available sexual offender treatment outcome studies to empirically establish clear effectiveness of such interventions, a number of serious methodological issues undermine results obtained to date, which further qualify

claims made in support of the effectiveness of sexual offender treatment. Such methodological issues in sexual offender treatment literature include limitations of meta-analysis, inadequate length and methods of follow-up of subjects, failure to utilize survival analysis in outcome measurements, allegiance effects, the general failure to use RCT designs (e.g., to use random assignment of motivated or genuinely help-seeking subjects to treatment and control groups), and distinct problems in the existing choices of control groups. Each of these factors seriously qualifies the already uncertain findings of the available sexual offender treatment outcome literature, particularly the last two.

### Issues with Meta-Analyses

An initial issue for extant reviews that relied on meta-analysis is the well-known limitations of that method of evaluation. In the general psychotherapy outcome literature, a number of criticisms have been offered regarding meta-analytic studies of treatment outcome and the limitations of existing meta-analyses. Sharpe (1997) identified the primary criticisms of meta-analysis: (1) mixing dissimilar studies, (2) publication bias (including published studies which typically favor those with positive outcome), and (3) inclusion of poor-quality studies. Chambless and Hollon (1998) emphasized that in the absence of sufficient high-quality studies available for study, the results of meta-analyses were not dependable. Lambert and Ogles (2004) also opined that while there had been recent improvements in meta-analytic methodology, significant and problematic variability in meta-analytic methodology remains. The results of any meta-analysis of treatment outcome studies will be dependent upon the essential quality of available studies for analysis such that "summarizing" poor-quality or methodologically limited studies is not likely to be particularly informative. As Kendall et al. (2004) wrote, "Meta-analyzers cannot tabulate the number of studies in which treatment is found to be efficacious in relation to controls *without examining the nature of the control condition*" (p. 37, emphasis added). Specifically, virtually all of the SRs, including the meta-analyses, have noted the poor quality of existing sexual offender treatment outcome studies with virtually not of the available studies rating high on the Maryland scale or any other metric of study quality. As Eysenck (1994) stated, "...a good meta-analysis of bad studies will still result in bad data" (p. 789). He went on to state: "Meta-analyses are often used to recover something from poorly designed studies, studies of insufficient statistical power, studies that give erratic results, and those resulting in apparent contradictions...Effect sizes summed over such exceedingly heterogeneous data can hardly be accorded any validity, yet these data are often cited as proving the efficacy of psychotherapy" (pp. 791–792). As noted, Hanson (1997) offered a similar

opinion. Craig et al. (2003) noted the “considerable variability” in sample selection among studies of sexual offense recidivism and that sexual offenders are a particularly heterogeneous group of offenders. Many experts note the high degree of selectivity—investigator allegiance—that operates in the selection of which studies are included or disqualified and note that the problem of remaining “blind” in meta-analytic research has not been adequately addressed in such investigations (e.g., Eysenck, 1994; Westen & Morrison, 2001). Further, Matt (1989) demonstrated that judgmental factors are involved in selecting effect sizes from a meta-analysis. Average of varied effect sizes from the same studies showed variability; this can have a very significant influence on the reported results, to the point of reducing reported effect sizes by half. In addition, Hemphill (2003) noted: “It is important to recognize that different effect sizes do not produce results that are necessarily interchangeable. The magnitude of effect size cannot even be generalized across time within a single study because long follow-up periods increase observed base rates, which in turn influence magnitudes of effect sizes.” Given that, like other criminal and violent outcomes, the base rate of sexual offense recidivism increases with longer follow-up periods (e.g., tripling over a 15–20 years interval), this suggests that current effect sizes would not provide a meaningful measure of treatment effectiveness, even if the current effect sizes were empirically meaningful. At present, even with a greater number of studies available, virtually no modern methodologically adequate—e.g., meaningfully controlled—studies have been conducted. Consequently, the low methodological quality of existing meta-analysis constitutes a “rate limiting factor” study (particularly considered within the context of issues in meta-analysis generally) that will continue to qualify any interpretation of the results of meta-analyses of sexual offender treatment outcome studies.

### Issues with “Official” Recidivism as an Outcome Measure and Sample Censorship

Sexual offense recidivism is the primary outcome variable of interest relative to the efficacy of sexual offender treatment, that is, the primary concern as to whether psychotherapy “works” for sexual offenders as opposed to more common psychotherapy goals such as symptom relief or reduced personal distress. Such offense recidivism is the key metric for determining the efficacy of such interventions for several reasons. First, most generally, such recidivism best captures what Westen and Morrison (2001) referred to as sustained efficacy, the ability of treatment to produce lasting changes rather than an apparent positive initial response. Second, as forensic psychotherapy, the intention of such interventions, as well as the basis for providing public funding of sexual offender treat-

ment, is public safety, specifically the prevention of future harm to possible victims.

The conventional means of measuring sexual offense recidivism in existing studies of sexual offender treatment is typically one rearrest or reconviction as measured by existing official criminal records. Thus, there is no “count” as to whether those “treated” sexual offenders who “failed” by sexual reoffending after treatment had one or multiple victims, the number of times they victimized one or more victims or the degree of harm that resulted from the sexual offense for primary or secondary victims. Rather, it appears that almost all existing treatment studies have relied on available criminal justice outcome measures and typically for follow-up periods of no more than 5 years. However, as Douglas et al. (2006) pointed out, “Sole reliance on official records will invariably underestimate actual criminal behavior” (p. 545) and lower base rates of actual recidivism. Regarding violent behavior generally, Douglas and Ogloff (2003) found that when criminal records were supplemented by other archival sources, the base rate *quadrupled* from approximately 10 to 40 %. Similarly, Monahan et al. (2001) found that the inclusion of information from official records, other collateral sources, and self-report increased recidivism rates by a factor of six! For sexual offense recidivism specifically, the results of Prentky, Lee, Knight, and Cerce (1997) reported they found a marked underestimation of sexual offense recidivism specifically depending on whether the criterion was based on charges, conviction, or imprisonment. Further, some offenders commit multiple sex offenses or victimize the same individual repeatedly over a follow-up period. Further, it is near universally agreed that such official rates of sexual offense recidivism “miss” most sexual offenses, because such offenses are not reported by victims or not processed through the criminal justice system (e.g., Craig et al., 2003). As Hanson (1997) noted that while “detected” recidivism is a credible measure, it is “an insensitive measure,” pointing out that since most sexual assaults, particularly those against children are never reported to police; “It is impossible to study...that which remains hidden...Rarely will sexual offenders be falsely reconvicted, but many sexual offenses will go undetected” (p. 131). Craig et al. (2003) concluded: “...sexual recidivism could be underestimated by as much as 40 % in some studies” (p. 72). In addition, numerous studies via self-report in varied contexts have shown that under conditions created to maximize veracity, sexual offenders of all types reported substantially greater frequency (and diversity) of sexual offending (e.g., Abel, Blanchard, & Becker, 1978; English et al. 2000; Heil, Simons, & Ahlmeyer, 2003; Ahlmeyer et al., 2000; Hindman and Peters, 2001, 2010). Yet sexual offender treatment studies rely on the relative minority of actual sexual offenses that are detected, reported to, and processed by the criminal justice system. Consequently, using sexual offense recidivism

as measured by official records of arrests and/or convictions provides a grossly insensitive index of outcome, leading to a likely significant underestimate of both the frequency and severity of sexual reoffending.

In addition, it is the consensus that sex offense recidivism rates increase substantially with increased periods of follow-up (e.g., Rice & Harris, 2003); per Harris and Hanson (2004), rates of sexual reoffending almost double when follow-up periods are extended from 5 to 15 years (e.g., 14–24 %), and offenders have greater “opportunity time” in the community to commit new sexual offenses. Given such measured base rates for detected sexual offense recidivism, it makes little sense to investigate the effect of psychotherapy on sexual offense recidivism for periods of less than 5 years; as the results of Prentky et al. (1997) showed that for a 5-year study, only ½ of the total number of cases of sexual offense recidivism would have been identified. Similarly, Craig et al. (2003) noted that with a 5-year follow-up to treatment studies, “only one-half of the total number of cases of sexual reoffending – would likely have been identified. Thus, extended periods of follow-up are necessary to determine if true, meaningful reductions in sexual offense recidivism occur.” Several studies have shown that sex offender recidivism increased to approximately a rate of 40 % by a 20-year follow-up, approximately triple the rate of 5 years sex offender recidivism rates (40 %, e.g., Hanson, Morton, & Harris, 2003; Harris & Hanson, 2004; Harris & Rice, 2007). Doren (2002) noted that there are no research studies of sex offender recidivism through the death of the entire sample (e.g., a lifetime rate of such recidivism). Sexual offenders demonstrate “first-time” sexual reoffending even 20–30 years after release from institutionalization (e.g., Hanson, Steffy, & Gauthier, 1993; Prentky et al., 1997); Hanson et al. (1993) reported that 23 % of sexual offender recidivists were reconvicted for more than 10 years after release.

Sample censorship is another issue relative to the accuracy of sexual offender recidivism rates. The common method of simply counting the percentage of individuals who sexually reoffend over a limited period of time for several reasons has several limitations that make it very likely to produce an underestimate of the true rate of such recidivism. First, persons with more severe histories of sexual offending may serve longer sentences of being indeterminate confined or detained and thus “unavailable” to sexually reoffend. Second, of those individuals released to the community, a significant number may only reside in the community for brief periods of time (e.g., due to re-incarceration of secondary to high general criminal recidivism rates or parole revocations) and will also be “unavailable” to sexually reoffend. Relative to this second point, Langan, Schmitt, and Durose (2003) found that 43 % of sexual offenders were rearrested for some crime (75 % of which

were felonies) within 3 years of their release from prison. Even more recently, Durose, Cooper, and Snyder (2014) showed that 71 % of violent offenders (including sexual offenders) were rearrested for some criminal offense within 5 years of release from prison. Given these results from the Department of Justice, a significant proportion of released sexual offenders are jailed or re-imprisoned during what would have been their “follow-up” time and, obviously, less “available” to commit another sexual offense. Epperson (2009) found that over 52 % of moderate-risk sexual offenders, 56 % of higher-risk sexual offenders, and 65 % of the highest-risk sexual offenders released from prison on conditional release experienced revocation that in a number of cases would have led to additional periods of jail or prison time. Thus, higher-risk sexual offenders were more likely to be out of the community during some portion of a potential “follow-up” period. Furthermore, since the late 1990s (e.g., Prentky et al., 1997; Rice, 1997), the scientifically endorsed method for follow-up studies of sexual offenders is survival analysis. This method takes into account not only whether members of the groups of sexual offenders commit subsequent sex offenses but also when the end of sexual offender treatment or release from incarceration occurs and the length of time “available” to each offender for sexual offending activity in the community (e.g., not or deceased and/or incarcerated or jailed for lengthy periods of time). That is, survival analysis only counts the time that an offender is, in fact, “available” to sexually reoffend; as a data analytic procedure, survival analysis provides a better estimate of sexual offense recidivism (relative to a point recidivism rate) as it takes into account the “opportunity time” for each offender who has been “in the community” and actually had the chance to sexually reoffend. Of note, when Olver, Beggs Christofferson, Grace, and Wong (2013) controlled for risk and individual differences in follow-up time using survival analyses over an 8-year fixed follow-up period, the overall group of treated sex offenders did not demonstrate significantly lower rates of sexual recidivism than a much smaller control group. Relative to this point, Langan, Schmitt, and Durose (2003) found that 43 % of sexual offenders were rearrested for some crime (75 % of which were felonies) within 3 years of their release from prison. Even more recently, Durose, Cooper, and Snyder (2014) showed that 71 % of violent offenders (including sexual offenders) were rearrested for some criminal offense within 5 years of release from prison. Since most sexual offender treatment studies to date have failed to employ survival analysis, it seems highly likely that existing treatment studies overstate any benefits of such studies, since they are likely “missing” a substantial number of sexual offenders in general—and higher-risk sexual offenders more specifically—during the follow-up period.



## Posttreatment Experiences of Treatment Participants

Another methodological issue concerns posttreatment experiences or services that treatment participants may have received. Following their experience of sexual offender treatment, some portion of “treated” sexual offenders remain in institutions and/or are followed in the community over time after their experimental treatment experience. During this period after sexual offender treatment, there are often further opportunities for exposure to many possible events that might have short- or long-term impact on their sex offense recidivism rates. That is, after the initial intervention hypothesized to be effective at reducing sex offense recidivism, it seems quite possible—and even likely—that treatment subjects and control subjects may have obtained additional treatment experiences, social services, and/or some degree of parole supervision, all of which might be significant factors related to lowering recidivism rates. It is commonly noted that sexual offenders in the Canadian correctional system may receive additional, often substantial, rehabilitation or pro-social programming (e.g., substance abuse treatment, criminal thinking interventions, reintegration services) while institutionalized and/or during probation, including additional specialized sexual offender treatment as they are placed sequentially at different institutions. Obviously, the nature (intensity of conditions) of post-release supervision or probation as well as varied types of post-release or posttreatment aftercare may have a significant and differential effect on those who did and did not participate in sexual offender treatment.

## Treatment Allegiance

Allegiance to a treatment approach refers to the degree to which a therapist providing the treatment believes that the psychotherapy is effective; in effect, this constitutes an expectancy effect and potential bias. Those who develop or are advocates for particular or general treatment programs may be relatively zealous about the likely benefits for their own proposed or endorsed interventions. Unlike medications studies (which can be administered in a blind or double-blind manner), allegiance effects in psychotherapy cannot easily be controlled. Wampold (2001) reported that early meta-analyses showed that treatment effects for which the clinician had an allegiance or expectancy produced an effect that was approximately 1/3 larger than the opposite condition. He noted that in one meta-analysis, the correlation between allegiance ratings and the effects of the study approached 0.60, while another similar study suggested that allegiance effects might be somewhat less. However, Wampold (2001) concluded:

...it is clear that allegiance of the therapies is a very strong determinant of outcome in clinical trials. That the effects due to the allegiance accounts for dramatically more of the variance in out-

come than does the particular type of treatment implies that therapist attitudes and expectancies about the results of psychotherapy are a critical component of effective therapy.... (p. 168)

As noted in a MA, Munder et al. (2013) found that research allegiance to the intervention itself showed a moderate effect size with treatment outcome; psychotherapy researchers are likely to “find” what they want or intend to “prove.” Not surprisingly, allegiance effects apply to sexual offender treatment as well. In their articles, Losel and Schmucker (2005, 2008) found that in more than 50 % of the primary research studies, the studies’ authors were affiliated with the treatment program that was implemented (suggesting allegiance issues). Not surprisingly then, they showed that for such treatment, studies (in which the study author(s) was in some way involved in the program delivery) showed clearly significant treatment effects. Yet in contrast, programs that were evaluated by *independent* researchers did *not* show positive treatment effects; this strongly suggests the so-called allegiance effects.

## The Lack of Randomized Controlled Studies: A Multitude of Problems

The primary methodological criticism of the existing literature on psychotherapy for sexual offenders concerns the almost uniform failure to utilize accepted standardized research designs for interventions (e.g., RCTs involving both random assignment of similar subjects to a psychotherapy condition and at least one control condition). Hanson et al. stated that a “strong” treatment outcome study would be one that involved “a well-implemented random assignment study (e.g., uncorrupted random assignment, 5 or more years of follow-up, sample size >100, < 20 % attrition, no preexisting differences between the groups found post hoc)” (p. 869). In contrast, the available sexual offender treatment research literature relies almost exclusively on experimental and control groups that are *each* biased in the direction of providing the appearance that sexual offender treatment has been demonstrated to be effective. While RCTs offer one perspective as part of an evidentiary hierarchy and of evidence-based practice and do not necessarily avoid some methodological issues themselves, they are the critical standard in providing key experimental findings that are more conclusive in establishing causal relations of treatment effects than results obtained utilizing other methods or approaches. As Kendall et al. (2004) articulated from a research perspective and the Cochrane Collaboration emphasized from a health policy/health economics perspective, RCTs provide the fundamental basis for evidence-based intervention research and resultant health-care treatment policies. Sackett et al. (1996) wrote “...we should avoid non-experimental approaches...since these routinely lead to false positive conclusions about efficacy...[so that]

the systematic review of several randomized trials... has become the 'gold standard' for judging whether a treatment does more good than harm" (p. 171). RCTs are studies that enable stakeholders a relatively unique opportunity to assess whether an intervention itself, as opposed to other factors, is responsible for observed outcomes in clients. RCTs are designed most likely to nullify unknown or hidden threats to internal validity or confounding factors. The failure to utilize random assignment of comparable and motivated sexual offenders to intervention or control conditions dramatically works to prevent reaching any meaningful conclusion that sex offense treatment might be effective at reducing sex offense recidivism. As noted previously, in the general psychotherapy literature, RCTs are considered the *sine qua non* of methodologically correct scientific study of treatment outcome.

McConaghy (1993) was one of the first authorities to emphasize the importance of RCTs and the limitations of uncontrolled sexual offender treatment studies. The unique significance of RCTs in sexual offender treatment specifically has been repeatedly emphasized by numerous individual authorities (e.g., Quincy et al., 1993; Rice & Harris, 1997; Quinsey, Khanna, & Malcolm, 1998, 2006; Rice & Harris, 2003; Seto et al., 2008). Seto et al. (2008) noted that primary health-care research and policy agencies, including the Cochrane Collaboration, the US Center for Disease Control and Prevention, and the US Food and Drug Administration, each identify effective interventions exclusively based on RCT results. As Seto et al. stated, from an experimental design perspective, RCTs "are the best at distributing [pretreatment] differences randomly, and only randomization can eliminate the subtle selection biases that affect even the best incident study designs" (p. 249). Similarly, the SBU stated:

The ideal study design is the randomized controlled trial (RCT), where offenders or people at higher risk of becoming offenders are randomly assigned to either a treatment (i.e. the studied intervention) or a control group (e.g. another intervention or no treatment)...[as a result of this procedure] we can be relatively confident that a difference in reoffending is a result of the treatment. (p. 16)

Since 2010, the Association for the Treatment of Sexual Abusers (ATSA) has been on record that:

[I]t recognizes randomized clinical trials (RCT's) as the preferred method of controlling for bias in treatment outcome evaluations. ATSA promotes the use of RCT to distinguish between interventions that decrease the recidivism risk of sexual offenders and those program that have no effect or are actually harmful...full RCTs are always preferable, and are unparalleled for determining causal relationships between treatment and outcome. (ATSA, 2010a)

RCTs provide for two factors that allow conclusions to be reached about the possible effectiveness of intervention. First, they require that an intervention condition be contrasted with

one or more control conditions; thus the RCT design provides a preliminary determination as to whether subjects who received the intervention may have received some specific positive benefits relative to the control conditions. In an early paper, Quinsey et al. (1993) (as cited in Rice & Harris, 2003) first advocated criteria that could provide useful scientific data on the effectiveness of treatment, stating "...unless a study measures officially recorded recidivism from at least two distinct groups of sex offenders (at least one of which receive treatment), and unless the groups are, except for treatment, comparable, that study has no scientific value in evaluating treatment" (p. 431). McConaghy (1993) noted that random allocation of subjects in sexual offender treatment is the only procedure that offers the possibility of controlling all relevant variables, known and unknown. Generally, some 30 years ago, Cook and Campbell (1979) pointed out that the main problem of quasi-experimental design is the differential selection of subjects that receive the program compared to the subjects that do not receive the program. If at the beginning of the program the groups are not equivalent for the relevant variables, then the posttest comparison of the two groups can produce a biased estimate of the effect size. More recently, regarding sexual offender treatment specifically, as Miner (1997) put it:

The major problem with uncontrolled designs is that they provide no means for assuring the internal validity of the study. The lack of control or comparison groups makes it plausible that any changes in subject status could be attributed to factors other than the intervention itself. This leaves the researcher unable to conclude much about the effectiveness of treatment. (p. 99)

As Schlank (2010) noted, several common psychological phenomena can affect intervention results. For example, she identified the Hawthorne effect, where a temporary change in measured behavior occurs as a result of subjects' awareness that they are being observed. In addition, she also noted the Pygmalion effect when a perceived "leader" or teacher's expectations affect the behavior of students (or clients), at least temporarily (a problem often related to allegiance effects).

A related methodological issue is the comparability of intervention and control groups; both groups must be relatively equivalent in key characteristics. As Miner (1997) stated: "The major problem with nonequivalent groups designs is an issues of the linkage between cause and effect," (p. 100) noting that differences in groups on variables as simple as motivation for treatment make it difficult to conclude that group differences may be related to an intervention condition. Thus, similar to standard psychotherapy outcome research, in order to best assure comparable treatment and control groups, it is necessary to start with subjects comparably interested in and motivated for treatment and then randomly assign them to treatment or control groups. Similar to Miner, Rice and Harris (2003) emphasized that investigators

generally agree that it is desirable to limit or control possible sources of measurement bias in the study groups and that the best and necessary means of accomplishing this is through random assignment via an RCT. With random assignment of subjects to intervention or control group(s), the allocation of similarly motivated subjects to either intervention or control groups is determined solely by chance (and not by personal preference or social mandate). [Such groups may differ by chance, as Rice and Harris noted that while “the gold standard is a random assignment study, but even with random assignment the treatment does not guarantee the groups are comparable: random assignment merely guarantees that differences are randomly distributed” (p. 429).] Thus, RCTs are a necessary but not necessarily sufficient condition to demonstrate that any differences found between experimental/treatment and control groups are most likely the result of the intervention and not simply the result of preexisting differences in the experimental and control groups.

Rice and Harris (2003, 2012) and Seto et al. (2008) have identified that, historically, several significant studies of medical and psychosocial interventions were initially conducted without random assignment of subjects to intervention and control groups and initially appeared to show that a particular treatment was effective (including studies of delinquency intervention, arthroscopic knee surgery, drug abuse prevention, and critical stress debriefing). However, when RCTs were utilized for these and other problems, either no or even negative effects were demonstrated for what had previously been regarded as theoretically sound interventions. Consequently, without the use of RCTs, inadequate and/or harmful interventions would have gone undetected. Similarly, regarding sexual offender treatment, Seto et al. (2008) commented on the possibility that “unproven treatment might have harmful effects, unintentionally increasing recidivism and thereby harming victims, offenders, and their respective families” (p. 250). They provided several examples of how current practices in sexual offender treatment might hypothetically lead to increased risk for sex offense recidivism. Other authorities have echoed these concerns (e.g., Corabian et al., 2010; Dennis et al., 2012).

Schmucker and Lösel (2008) acknowledged that 60 % of the studies they reviewed “used clearly non-equivalent control groups” (p. 16). In considering the sexual offender treatment outcome studies reviewed by Hanson et al. (2002) assigned to the category of random assignment of subjects (to either psychological treatment or no psychological treatment), Rice and Harris (2003) noted that Hanson et al. found only three studies in total that could be assigned to this category. Two of these studies indicated deleterious effects of treatment and one indicated reduced general but not sex offense recidivism. Only one RCT study reported positive treatment results for sex offense recidivism; Borduin et al.

(1990)<sup>6</sup> provided “multisystemic therapy” (MST—a model not easily applied to adults) for a small group ( $n=24$ ) of adolescent offenders with positive effects. Subsequently, Rice and Harris agreed with the conclusion by Hanson et al. “that no empirical support for treatment effectiveness can be drawn from the random assignment studies, especially not for sex offender specific treatment for adults” (p. 434). Of note, Lösel and Schmucker (2005) similarly found: “The size of the [treatment] effect is small to moderate...Restricting the analysis to a few randomized trials shows a comparable mean effect but it does not render it statistically significant.” Eggers et al. (2001) demonstrated how conclusions from a meta-analytic review based on a number of small-scale trials were subsequently contradicted by results from a single study containing a much larger sample; as McGuire has stated “The two most frequently repeated criticisms of meta-analysis, are loosely termed, those of ‘garbage in—garbage out’ and ‘apples and pears’.” With the absence of positive results from the very few RCTs for sexual offender treatment outcome, Rice and Harris wrote: “...weak inference evaluation leads to too many errors (and incorrectly accepting the existence of beneficial effects)...” (p. 429).

Further, another significant methodological issue makes even random assignment of potential sexual offender participants in treatment problematic. Typically, in RCTs for mood/anxiety and/or behavioral problems, the initial subject pool for treatment or control group assignment is persons who have volunteered to participate in such treatment. Consequently, most psychotherapy investigations start with persons who truly want—are motivated—to participate in such intervention to relieve personal distress or impairment. In fact, they are likely to be persons who may have tried other treatments without success and have elevated positive expectations and enthusiasm for treatment participation. Subsequently, that group of motivated help-seeking persons is typically randomly assigned to either treatment or control conditions. However, almost all extant sexual offender treatment outcome studies have not included motivated or help-seeking sexual offenders in comparison or control groups. Rather, these studies involve the biased preselection or composition of either or both the experimental/treatment group and the control group. More specifically, as will be seen, persons who end up participating in sexual offender treatment are generally likely to have lower sexual offense recidivism a priori, while those who decline such treatment are likely to have higher sexual offense recidivism rates a priori.

<sup>6</sup>Of note, per Westen et al. (2005), when MST was “transported” from research (“effectiveness” to community settings “efficacy,” with effect sizes diminished by as much as 1/3 when provided by community therapists (as opposed to carefully supervise graduate students); in addition, there were relatively small effects on individual psychopathology relative to family relations.

Thus, a significant issue is which sexual offenders are included in sexual offender treatment outcome studies. First, many or most sexual offenders appear to not even be offered treatment. As Marshall (Marshall & Marshall, 2007; Marshall, Marshall, Serran, & O'Brien, 2011) has pointed out, most RCT treatment studies involved exclusion criteria that are often quite extensive (e.g., not disruptive, no below average intellectual functioning, no comorbid psychiatric conditions, and so on), and as a result, those offenders who participate in treatment are much more likely to have lower recidivism rates even prior to treatment. In addition, other investigators have differentially excluded particular groups of sexual offenders from possible participation in treatment studies. Reviewers have identified that a number of treatment programs only included offenders deemed less "severe" (e.g., of only low or moderate risk) and *excluded* more high-risk or "severe" sexual offenders for participation in sexual offender treatment. For example, in Hall's meta-analysis, as many as 33 % of sex offenders eligible for treatment were "screened out" and *not offered* treatment; specifically, Hall noted that the more "severe" sexual offenders (e.g., those with more extensive sexual offense histories, with mental health problems, who denied their sexual offense history, perceived as management problems, etc.) were not offered treatment in the studies that he reviewed. Similarly, Jones, Pellissier, and Klein-Saffran (2006) reported that 16 % of persons who had volunteered for sexual offender treatment were refused because of psychological reasons including lower intellectual capacity, severe mental illness, low motivation, history of treatment failure, and nonacceptance of responsibility for sexual offending. An additional 22 % of sexual offenders were refused treatment after being accepted and assigned to treatment. In addition, Marques et al. (2005) excluded any sexual offender with more than two prior felonies; thus, the treatment candidates were a low- or moderate-risk group to begin with. They also excluded offenders who denied their crime. Further, Marques et al.'s study could be viewed as "incentive laden" in that it involved a transfer from a prison to a one specific hospital setting, further limiting potential candidates. After selection criteria in SOTEP, 68 % of participants were low or medium risk. As Marshall and Marshall (2007) have pointed out:

These exclusionary criteria would have biased the SOTEP *in favor* of finding a treatment effect... "they pointed out"...it seems reasonable to conclude that the nonvolunteers were among the most treatment resistant offenders and likely the ones in most in need of treatment. (p. 183, emphasis added)

[Despite this bias, of course, the SOTEP did not identify a positive treatment effect for CBT-RP and aftercare.] Thus, outside of mandated/coerced sexual offender treatment, as Harris, Rice, and Quinsey (1998) suggested, relying on persons who volunteer for and persist with treatment effectively screens out most high-risk sexual offenders and consequently

participation in "...treatment over the long terms serves as a filter for detecting those offenders who are relatively less likely to reoffend..." (p. 103).

Beyond higher risk, other factors also influence the inclusion of offenders into sexual offender treatment; these include acknowledgement of some history of sexual offending and self-reported motivation for intervention. Beyond eliminating offenders with more serious sexual offending history (e.g., per risk studies, higher-risk sexual offenders), studies have typically selected only those offenders who (1) (must) admit to their offenses and/or (2), to varying degrees, are willing to participate in sexual offender treatment, either because they believe it is beneficial or because they may view such participation as providing them some gain or advantage (e.g., early release). Tierney and McCabe (2002) noted that some treatment programs target only the most "motivated" sexual offenders because they are considered most likely to change their behavior. In the SOTEP (Marques et al., 2005), only 1/3 of sexual offenders invited to participate in sexual offender treatment were willing to enter the research intervention; that is, 2/3 of sexual offenders offered sexual offender treatment refused to even consider entering the sexual offender treatment study. Thus, in this unique modern RCT, there was a highly significant degree of self-selection relative to a willingness to pursue sexual offender treatment. Rice and Harris (2003) pointed out that, in general, offender self-selection for sexual offender treatment has been the norm. Losel and Schmucker (2005) found that only 16 % of sexual offender treatment participants could be characterized as "volunteers." Further, Larorchelle et al. (2011), in their review of 18 studies, found that between 15 and 86 % of sexual offenders who began sexual offender treatment dropped out (the most consistent predictor being antisocial personality disorder and other antisocial characteristics). Thus, outside of mandated/coerced sexual offender treatment, Rice and Harris (1998) suggested that the utilization of persons who volunteer for and persist with treatment effectively screens out the higher-risk sexual offenders. As a result, in most contexts, sexual offenders who do "volunteer" for and remain in sexual offender treatment appear to an extremely different group of sexual offenders (e.g., lower risk for sexual reoffending) from those who choose not to participate or those who are excluded from participating.

Yet another issue to be considered relative to those sexual offenders who "agree" to participate in sexual offender treatment is the degree to which entering into sexual offender treatment is truly voluntary. Marshall and Barbaree (1990) noted, "quite a number of patients are under judicial or administrative pressure to enter and remain in treatment" (p. 375). As Losel and Schmucker (2005) reported, only sexual offender treatment programs involving *voluntary* participation by offenders showed a significant effect; programs that involved "a more or less coerced treatment" did not

show a significant treatment effect. Other studies have also found that the degree of mandate or coercion is related to treatment outcome in offender populations.

In short, to date, those person who have been studied after receiving sexual offender treatment are a minority of sexual offenders, apparently lower-risk offenders, those without comorbid psychiatric disorders or intellectual disabilities, and those with mixed or uncertain motivation (e.g., some intrinsically motivated and others mandated and with external motivation). Consequently, the experimental or treatment groups in sexual offender treatment outcome research should be viewed skeptically as representatives of sexual offenders in general, relative to their apparently unique identification or interest in seeking treatment as well as their degree of risk and associated disorders.

In addition, the nature of the comparison or control group in sexual offender treatment outcome studies is another highly significant methodological issue that potentially contaminates the results of such studies. As Marshall and Marshall (2007) noted:

One problem with the incidental design, however, is that there may be a plethora of undetected but significantly influential differences between the treated and untreated subjects aside from the usual matching variables (i.e. some limited demographic and offense history features). Frustration with not begin given access to treatment, differential responses by the authorities to treated and untreated subjects (e.g. refusal to grant parole to untreated offenders, placement in a less attractive prison setting) may provoke responses in the untreated subjects that might confound the matching process. (p. 186)

The majority of sexual offender treatment outcome studies that utilize a nonrandom assignment control group are characterized as “incidental assignment” studies and constitute what is referred to as “quasi-experimental” designs. Since they are not RCTs, they have not randomly assigned comparable, motivated offenders to either treatment or a control condition. Thus, in the existing incidental assignment studies, researchers have resorted to utilizing various groups of offenders to serve as a “control” condition including identified treatment refusers; treatment dropouts; persons selected from a general group of sexual offenders (sometimes contemporaneous offenders and sometimes from a different time period); and/or general sexual offenders matched to a treatment group on one or more variables. However, it has been demonstrated that such comparison groups are each problematic for determining if the treatment condition for sexual offenders is actually effective. It can be demonstrated that for at least the first three potential control groups, their use of such types of sexual offender as a control group is compromised since each group—a priori or pretreatment—would almost certainly have a higher rate of sexual offense recidivism than those persons typically screened or volunteering for participation in sexual offender treatment.

The general group of sexual offenders, particularly after removing those selected as potential treatment candidates or who volunteer for sexual offender treatment, consists relatively of sexual offenders who would either refuse sexual offender treatment or would likely drop out of such treatment; both groups are known to be at higher risk of sexual offense recidivism than the “average” sexual offender. Using persons for control groups who have or would refuse sexual offender treatment will lead to a control group that is already characterized by an elevated risk for sexual offense recidivism. Thus, several years before Hall’s initial meta-analysis, Quinsey et al. (1993) had argued that treatment refusers should not be ignored in considering treatment efficacy because of their particularly high rate of recidivism. More generally, this recommendation was in line with the increased importance of intent-to-treat (ITT) analyses in the general psychotherapy outcome literature. “Intent to treat” is a strategy for the analysis of randomized controlled trials that compares all clients based on the groups to which they were originally randomly assigned. Thus, to meaningful measure how well a particular intervention works, all individuals randomly assigned to that treatment condition are followed and evaluated, regardless of whether they actually entered, dropped out, or completed that treatment. ITT analysis reflects the practical clinical scenario because it recognizes the meaning of treatment noncompliance, later treatment rejection and treatment protocol deviations. Clinical effectiveness may be overestimated if an intention to treat analysis is not done (e.g., Hollis & Campbell, 1999); for example, how effective is surgical castration if few or no persons are willing to consent to it? Of note, the FDA of the USA recommends ITT analyses, noting the results of a clinical trial should be assessed not only for the subset of patients who completed the treatment but also for the entire sample of individuals who were randomized to treatment or control conditions.

As Harris et al. (1998) initially pointed out for Hall’s (1995) meta-analysis in several studies, “all or most of the control group were men who refused or quit treatment...” (p. 102). Losel and Schmucker (2005) showed that in approximately 24 % of the psychosocial sexual offender treatment comparisons, the control group consisted of treatment refusers. Treatment refusers are clearly not characterized by significant motivation for sexual offender treatment. Marshall and Barbaree (1990) commented on earlier research by Abel, noting that almost 35 % of sexual offender entering his program withdrew or were terminated; “the highest rates of withdrawals from their program occurred in those patients who felt the greatest pressure to participate in therapy” (p. 375). Olver et al. (2013) noted that if dropout rates were not carefully managed and reduced, dropping out might act like a self-selection process, unwittingly resulting in the treatment of predominantly or exclusively lower-risk offenders. As

noted previously, in the SOTEP study by Marques et al. (2005), initially, 2/3 of sex offenders offered sexual offender treatment refused to consider such intervention. Later, an additional 21 % of that 1/3 that had previously volunteered for sexual offender treatment withdrew prior to the beginning of treatment. Given that 66 % of sexual offenders declined to participate in sexual offender treatment initially and then additional 21 % of those assigned to treatment refused, a basic issue raised regarding sexual offender treatment is the level of interest or motivation for participating in a particular intervention program. Such sexual offender treatment refusers have been identified as characterized by higher sex offense recidivism rates than persons who volunteer for sexual offender treatment (Abel, Becker, Cunningham-Rathner, Mittelman, & Rouleau, 1988, as cited in Quinsey et al., 1993). Utilizing a control group of sexual offenders who appeared to be untreated sexual offenders, Olver et al. (2013) found that “The importance of controlling for risk was underscored by the fact that untreated offenders scored significantly higher on [a static risk measure] and thus were higher risk for sexual and violent recidivism overall” (p. 415). That is, as with many studies that utilize a quasi-experimental design, those sexual offenders selected for control purposes were at elevated risk for sexual violence to begin with. As Rice and Harris (2003) wrote, “It is highly probable that, irrespective of the effects of treatment, those who refuse represent greater risk than those who volunteer for and completed...” (p. 432). More recently, Seager et al. (2004) found that treatment refusers had particularly high rates of sex offense recidivism relative to treatment completers (e.g., 42 %).

It is also problematic to utilize persons for control groups who have dropped out of or been terminated from treatment as that will also lead to a control group that is characterized by an elevated risk for sexual offense recidivism. As noted previously, for psychotherapy in general, treatment dropout or attrition from research studies has averaged 47 % and in actual clinical settings has been found to be even higher (e.g., Wierzbicki & Pekarrik, 1993). Quinsey et al. (1993) also pointed out that sexual offender treatment dropouts should not be ignored in considering treatment outcome because of their particularly high rate of recidivism. Beyko and Wong (2005) noted, “Unfortunately, attrition from many sexual offender treatment programs is high, up to 30–50 % in both residential and community programs” (p. 376). In SOTEP, Marques et al. (2007) found that 18 % of the small group of sexual offenders that had previously volunteered for and were assigned to sexual offender treatment did not complete the program (27 voluntarily withdrew and 10 were demitted because they presented as “severe management problems in the hospital”). Thus, dropout rates for persons placed in sexual offender treatment are very high.

More importantly, most available data suggests that treatment dropouts (or persons terminated from such inter-

ventions) are each characterized by higher recidivism rates than persons who volunteer for sexual offender treatment (Abel et al., 1988, as cited in Quinsey et al., 1993); failure to complete sexual offender treatment was a significant predictor of sex offense recidivism. As Olver et al. (2011) stated: “The clients who stands to benefit the most from treatment (i.e. high-risk, high-needs) are least likely to complete it” (p. 6). Marshall (1993) concluded “dropouts included a significant proportion of those sex offenders at greatest risk to offend” (p. 526). Specifically, Miner and Dwyer (1995) showed that treatment dropouts sexually reoffended at a rate three times that of treatment completers. Miner (1997) noted in his research that he “found higher reoffense rates in those [offenders] who terminated prematurely” (p. 101). As Seager et al. (2004) pointed out, Hanson and Bussiere (1998) found that there was a 17 % difference in sex offense recidivism rates between treatment dropouts and completers. Alexander (1999) found that dropouts were twice as likely to sexually reoffend. While Hanson et al. (2002) did not report the frequency with which control groups contained treatment dropouts, they did find that persons who eventually dropped out of treatment had consistently higher rates of sex offense recidivism. More recently, Seager et al. (2004) also found that treatment dropouts (as well as persons who were rated as failing to complete treatment) had particularly high rates of sex offense recidivism relative to treatment completers (e.g., six times greater). In this last investigation (albeit a small sample), the following rates of recidivism were found: 18 % for treatment dropouts and 100 % for those terminated from treatment. As noted, Losel and Schmucker (2005) reported, “Whether treatment was terminated regularly or prematurely had an impact on sexual recidivism. Where as regular completers showed better effects than the control groups, dropouts did significantly worse. Dropping out of treatment doubled the odds of relapse...” (p. 132). Langton, Barbaree, Hansen, Harkins, and Peacock (2007) also found that treatment dropouts showed the fastest failure rates.

Hanson (per a personal communication cited by Rice & Harris, 2003) agreed that, in fact, there were a priori reasons why treatment dropouts should be considered at higher risk to reoffend. Generally, as with sexual offender treatment refusers, dropouts are identified as likely to be more impulsive, show less self-control and other antisocial characteristics, and possess fewer social skills, all factors known to be associated with increased recidivism risk (e.g., Marques et al. 1994; McConaghy, 1999; Rice & Harris, 2003, 2012; Seager et al., 2004). Langton et al. (2006) also found treatment dropouts had significantly higher PCL-R scores (more psychopathic traits) than offenders who completed the same treatment program. Similarly, Beyko and Wong (2005) found that sexual offender treatment dropouts were characterized by two clusters of behaviors, one which they related to

criminogenic needs (e.g., aggression, rule-breaking behavior, longer offense histories, and more criminalized) and a second which they viewed as a responsivity issue (e.g., lack of motivation and denial). Olver and Wong (2009) found that 56 % of sexual offender treatment dropouts met study criteria for psychopathy. Nunes and Cortoni (2008) found that treatment dropouts were significantly associated with elevated general criminality characteristics. Olver et al. (2013) found that their untreated (but not randomized) control group scored as higher risk for sexual and violent offense recidivism than their treated group. In short, since sexual offenders who drop out of sexual offender treatment appear to be different and, most importantly, higher risk for reoffending than offenders who complete such treatment, “intent-to-treat” or treatment as assigned analyses appear imperative to rule out the significance of pretreatment differences.

A number of outcome studies for sexual offender treatment have utilized persons selected from some general group(s) of sexual offenders. Based on the research findings cited above, the majority or a large percentage of a general group of sexual offenders would refuse to participate in such interventions. In addition, such general groups of sexual offenders would include persons typically or historically excluded from treatment studies because they would be deemed high risk, mentally ill, intellectually limited, unmotivated, and so on, by virtue of what is known about treatment exclusion criteria. More importantly, as should be apparent from the just reviewed studies, since the great majority of sexual offenders either refuse or withdraw/drop out of sexual offender treatment, any group of general sexual offenders would almost certainly contain a substantial group of likely treatment refusers/dropouts if placed in or offered treatment. Thus, whether a control group included persons typically excluded from sexual offender treatment studies or treatment refusers/dropouts, such control groups would necessarily be composed of persons who, *prior* to any treatment being provided in a study, would very likely be sexual offenders at much higher risk to reoffend: high risk, unmotivated, more severely and comorbidity disordered, and likely to drop out if placed in treatment.

Another mechanism employed to create a control group relative to a treatment group is to attempt to “match” characteristics of the treatment group in the selection or creation of a control group, with the notion that such matching might result in equivalent groups for comparison. However, McConaghy (1993) noted that it is not possible to match offenders on all relevant variables as many offenders are not possible to assess accurately and many relevant variables were (are) not yet known. Similarly, relative to matching, as Seager et al. (2004) pointed out, studies rarely use more than three risk factors to match subjects, and they are often unable to match all treated subjects with untreated controls on the specified risk factors. In addition, while the comparison group may be “matched” on certain variables, the mem-

bers are still selected from the larger set of sexual offenders, which is to say that they include persons typically excluded from sexual offender treatment studies and/or treatment refusers/dropouts. Further, it is important to note that the most sophisticated sexual offender treatment outcome study, the SOTEP (prior to randomization), matched potential treatment candidates on age, criminal history, and type of offender (Marques et al., 2005) but still found no treatment effect. As the SBU indicated, when control groups are created for comparative purposes not by random assignment, there can be no certainty that any difference obtained between treatment and control group is the result of treatment; even with statistical attempts to control for variability between the groups, “since the differences between the groups cannot be attributed to chance, we can never be completely certain that the results are not due to some unmeasured, and perhaps unknown, risk factor that is more common in one of the groups” (pp. 17–18). Thus, Quinsey et al. (1998), when considering only the available studies from Hall’s MA that used a matching or randomization design, found that the effect size fell to 0; that is, the already “small” treatment effect in that MA was eliminated and no longer statistically significant.

As reviewed, the only basis for Hanson et al. (2002) concluding that there was *any* evidence of a treatment effect for psychosocial interventions applied to sexual offenders was their consideration of “incidental assignment” studies. Consequently, it is worth examining those results more closely. Rice and Harris (2003) reviewed the results of Hanson et al.’s (2002) “incidental assignment” treatment control group findings. As noted, in these studies, comparison groups were offenders selected as “matching” according to various methods including similar criminal records but who had been released before the implementation of the treatment program or who came from different geographical areas, who received an earlier version of the treatment, or who received no treatment or an alternative treatment due to such administrative reasons is too little time meeting their sentences. Hanson et al. (2002) labeled these seventeen studies as “incidental assignment” because they believed that there was no obvious *a priori* reason that the treated and untreated offenders would differ in risk and, thus, no “obvious” bias in group assignment. Of the 17 “incidental assignment” studies, only 11 were considered to be studies involving current treatments (those still being offered at the time of the meta-analysis).

Rice and Harris (2003) pointed out, “...with few exceptions, the studies included in this meta-analysis did not meet our criteria for minimally useful evaluation” (p. 433). They concluded that “the balance of available evidence suggests that various well-known threats to validity and the reliance on non-comparable groups are responsible for apparent beneficial treatment effects...” (p. 438). Rice and Harris (2003) specifically noted that 8 of the 11 “incidental assignment” studies clearly included sex offenders in the comparison group who

were not offered treatment, and thus, such studies appear to *include* likely treatment refusers or treatment dropouts in the control group who were not offered treatment. As noted previously, sex offenders selected for having completed treatment are not comparable to sex offenders who were not offered treatment because both refusal and dropping out are a priori risk factors for increased sex offense recidivism. As Seager et al. (2004) stated:

Within 'untreated' comparison samples a subset will be refusers and dropouts thus giving rise to concerns because refusers and dropouts reoffend at higher rates than completers. By failing to mathematically remove anticipated refusers and dropouts from untreated comparison groups, there is an inflationary effect for the treatment condition; that is, untreated comparison groups will have an exaggerated recidivism rate relative to the subgroup of untreated offenders who would have accepted treatment and remained till completion if offered the opportunity. (p. 601)

"Given that [a general group of sexual offenders] can be assumed to include a significant proportion who would have refused or quit treatment had it been offered to them and, therefore, are not appropriate comparison or control groups for the evaluation offender treatment" (Quinsey et al., 2006, p. 149) and would have dropped out of treatment had it been offered to them. Of the three remaining studies involving incidental design of current treatments reviewed by Hanson et al. (2002), Rice and Harris (2003) stated that each of them included significant methodological confounds that would neither meet their criteria for minimally useful evaluation nor even the Hanson et al. definition of "incidental assignment." Further, Rice and Harris (2003) and Seto et al. (2008) pointed out that in a number of studies included in "incidental assignment groups" in the Hanson et al. (2002) meta-analysis, a *double error* was found: offenders who refused or would have dropped out of treatment were not counted as part of the treatment condition but were counted as part of the control group. This procedure potentially reduced the measured sex offense recidivism of the treatment group *and* increased the sex offense recidivism of the control group, irrespective of the value of the intervention itself. As Rice and Harris (2003) concluded:

It is highly probable that, irrespective of the effects of treatment, those who refuse represent greater risk than those who volunteer for and completed...any study that does not track both refusers and dropouts cannot provide scientifically useful data in support of treatment effectiveness because there are clear a priori reasons to expect differences between the groups in recidivism... Samples of untreated sexual offenders will contain a substantial minority who would refuse treatment if offered, and another subset who, after beginning treatment, would quit or be ejected. (p. 432)

Similarly, Seto et al. (2008), responding to Marshall and Marshall (2007), pointed out that the design of the studies that provided the support for the conclusions of Hanson et al. (2002) emphasized that "...this decision creates a selection

bias, independent of any treatment effect, that increases the chances of finding newer offense among the treated sexual offenders..." (p. 252). They pointed out that "All of the incidental designs touted by Marshall and Marshall are even more vulnerable to the problem of inadvertent nonequivalence of groups, and all depend on some kind of statistical control from known risk factors" (p. 248). However, unknown risk confounding factors would not be subject to such a priori control. They also pointed out that Marshall and Marshall's rejection of RCTs would not take into account treatment motivation for sexual offender treatment (an issue on which Marshall himself has identified as a more systemic issue in providing interventions for sexual offenders). Barnett et al. echoed this concern stating "One problem with the incidental design, however, is that there may be a plethora of undetected but significantly influential differences between the treated and untreated subjects aside from the usual matching variables (i.e., some limited demographic and offense history features). Frustration with begin given access to treatment, differential responses by the authorities to treated and untreated subjects (e.g. refusal to grant parole to untreated offenders, placement in a less attractive prison setting) may provoke responses in the untreated subjects that might confound the matching process" (p. 186). Most recently, even Hanson (2014) has rejected the results of "incidental design," writing: "Comparisons between treated and untreated offenders from the same setting are usually biased because those who get treatment are systematically different from those that do not..." (p. 6).

Concerning both participants and refusers of sexual offender treatment, it seems clear that both typical treatment study participants and those excluded from inclusion, as treatment participants, are distinct and different groups of sexual offenders from one another. Both the exclusion of potential sexual offender participants by investigators and the self-selection by offenders relative to participation in treatment create meaningful differences in the pool of subjects who have composed treatment conditions. As Harris, Rice and Quinsey (1998) suggested years ago, volunteering for and persisting with treatment appears to effectively screen out most high-risk sexual offenders, writing "the data so far are consistent with the conclusion that agreeing to and persisting with treatment over the long term serves as a filter for detecting those offenders who are relatively less likely to reoffend..." (p. 103). That is, participation in sexual offender treatment does not appear to actually reduce recidivism rates for those who complied with treatment program but merely enables lower-risk, motivated sex offenders to demonstrate their commitment to not reoffend. Later, in 2012, Rice and Harris wrote, "...the predictors of treatment non-completion indicates that those who volunteer for and complete psychosocial treatment would, on average, exhibit a moderate to large difference in recidivism compared to those not offered



treatment, even if treatment had no effect” (p. 11). Effectively, in available sexual offender treatment outcome studies and the MAs and SRs of them, relatively lower-risk sexual offenders are being offered and accepting treatment participation, while higher-risk sexual offenders are both being excluded from or refusing participation in the sexual offender treatment that is the subject of study and often utilized as a comparison group. Further, both actual treatment refusers and dropouts appear similarly higher risk; relative to “intent-to-treat” principles, sexual offender treatment studies must identify and track both treatment refusers and dropouts because there are a priori reasons to expect differences between those groups and treatment volunteer/completers. Consequently, it is not at all surprising that group differences that *appear* to be treatment effects are identified when sexual offenders selecting and/or selected for treatment are compared to those sexual offenders who are not considered for or not volunteering for such treatment (because a significant proportion of who would likely refuse or drop out of such treatment) since a comparison group of non-volunteer sexual offender treatment individuals containing a relatively high proportion of both likely treatment refusers and actual treatment refusers will consist of a significant proportion of persons *already* at higher risk for sexual offense recidivism. Rice and Harris (2003) wrote: “In our opinion, few useful scientific data on effectiveness can come from studies contrasting complete treatment completers with sex offenders not offered treatment because such contrasts almost inevitably entail non-comparable groups” (p. 432). The most reasonable conclusion is that truly volunteering and being motivated for sexual offender treatment are among the most critical factors relative to outcome and that whatever intervention is offered makes little difference to the outcome; psychotherapy is irrelevant once client variables are accounted for.

In short, even prior to implementing treatment, in studies that find “small” differences between treated and untreated offenders, such differences would be expected based simply on the *likely preexisting differences in sexual offense recidivism rates* between persons selected and choosing treatment and other sexual offenders who are utilized as “control” groups. When differences in sexual offender treatment are found in non-RCT studies, the most reasonable conclusion is that sexual offender treatment *does not* lower the rate of sex offense recidivism below that of the average sexual offender; rather the most compelling conclusion is that rates of sex offense recidivism for *persons used as comparison groups are significantly higher than the average sexual offender*. Thus, the differences between treatment and control groups are not the result of treatment but a straightforward consequence of *preexisting risk status*. To this end, it is notable that per Table 1, the 5-year sex offense recidivism rates of sexual offenders who participated in treatment per the 2002 and 2005 treatment meta-analyses (10–12 %) are very simi-

lar to the 5-year sex offense recidivism rates of the very large groups of predominantly untreated sexual offenders identified in the risk-factor meta-analyses (13–14 %). Such a point is driven home even more so by the fact that when sexual offender treatment outcome studies utilize an RCT methodology (randomly assigned, comparably motivated, more equivalent treatment and comparison groups), no difference is found between those who participate in sexual offender treatment and those who do not as per Marques et al. (2005) (Table 1).

Thus, to date, scientific evidence has failed to demonstrate that sexual offender treatment completion per se reduces sex offense recidivism generally or for specific types of sexual offenders. Rice and Harris (2003) concluded, “The current empirical support suggesting beneficial effects of treatment rests on the use of non-comparable groups in which control subjects were of higher a priori risk” (437). Rice and Harris indicated, “Weak inference methods (as exemplified by almost all of the studies review by Hanson et al. (2002) ensure that the field of sexual offender treatment will continue to exhibit change without progress” (p. 438). Specifically, they concluded that the studies considered by Hanson et al. (2002), especially those in the so-called incidental category:

...cannot support even the tentative positive conclusions drawn...Indeed, the Hansen et al. (2002) analysis of incidental designs illustrates an important limitation of meta-analysis. The analysis of a set of uniformly weak designs cannot attribute variation of effect size to study quality. An overall effect size derived from studies of uniformly poor quality cannot obviate universal methodological weaknesses. Conclusions based on such a meta-analysis are no more justified inclusions based on the individual studies. (p. 437)

In fact, Rice and Harris (2003) found “the mean effect of treatment on sexual recidivism indicated a trend toward treatment having been detrimental...” (p. 437). It should also be pointed out that very little knowledge has accumulated about various matters critical to sexual offender treatment, including which aspects of treatment might produce reductions in recidivism or for what types of offenders might be most responsive to treatment. Rice and Harris reported that, “The literature provides almost no information about which treatment be most beneficial...” (p. 437). In fact, they pointed out “the mean effect of treatment on sexual recidivism indicated a trend toward treatment having been detrimental...” (p. 437). Hanson et al. (2009) concluded “Reviewers restricting themselves to the better quality, published studies... could reasonably conclude that there is no evidence that treatment reduces sex offense recidivism” (p. 881). The IHE (2010) stated, “Given the methodological problems of the available primary research it is difficult to draw strong conclusions about the effectiveness of sexual offender treatment programs using various CBT approaches for such a

**Table 1** Sexual offense recidivism rates in meta-analyses (MA) and SOTEP

Sex offense recidivism: general untreated sexual offenders+			
Hanson and Bussiere (1998) MA of risk factors	13 % SOR Sex offenders 23,000	19 % SOR Rapists	13 % SOR Child molesters
Hanson and Morton-Bourgon (2004) MA of risk factors	14 % SOR Sex offenders 31,000		
Sex offense recidivism: meta-analyses of sexual offender treatment			
Hanson et al. (2002) Treatment MA	12 % SOR General treatment	17 % SOR Comparison group	
	10 % SOR CBT	17 % SOR Comparison group	
Losel and Schmucker (2005) Treatment MA	12 % SOR Biological and psychological treatments	24 % SOR Comparison group	
	11 % SOR CBT	18 % SOR Comparison group	
Hanson et al. (2009) Treatment MA	11 % SOR	19 % SOR	
Sex offense recidivism rates for randomized control sex offender treatment study			
Marques et al. (2005) SOTEP	22 % CM SOR CBT+RP+ aftercare	17 % CM SOR Volunteer controls	21 % CM SOR Did not volunteer
Marques et al. (2005) SOTEP	20 % R SOR CBT+RP+ aftercare	29 % R SOR Volunteer controls	14 % R SOR Did not volunteer

CBT cognitive-behavioral therapy, SOTEP sex offender treatment evaluation project, RP relapse prevention, CM child molesters, R Rapists

heterogeneous population...Overall, the results reported by the selected SREs provide little direction regarding how to improve current treatment practices...There are still uncertainties reading the most useful elements and components of a sexual offender treatment program for convicted adult male sex offenders” (pp. iii–iv). As per the SBU in 2011, “For adults that have committed sexual offenses against children the scientific evidence is insufficient for determining which treatments that could reduce sexual reoffending.” Dennis et al. (2012) concluded: “The main finding of this systematic review is that there was no evidence from any of the trials in favour of the active intervention in a reduction of sexual recidivism—the primary outcome” (p. 25). Rice and Harris (2003) summarized:

In the end, we are obliged to conclude that the available data afford no convincing scientific evidence that psychosocial treatments have been effective for adult sex offenders... We conclude neither that treatment has been shown to be a waste of time nor that it has been demonstrated to be effective. (p. 427)

In 2010, the ATSA Executive Board endorsed the unique value of RCTs as the preferred method of demonstrating if sexual offender treatment is effective, writing: “ATSA

believes that RCT can and should be implemented in ways that respect the highest ethical standards. Community safety is better promoted by identifying treatments with strong evidence of effectiveness than by a proliferation of programs for which the efficacy is debatable.” There should be little disagreement with this point; no data from RCTs has yet to determine or establish that sexual offenders volunteering for (and not mandated for intervention) and who are randomly assigned to sexual offender treatment (as opposed to control conditions) exhibit lower rates of sexual offense recidivism during time spent in the community.

## Other Issues Regarding Outcome for Sexual Offender Treatment

### Alternate Outcome Methods and Results from Sexual Offender Treatment

Marshall (1993) has long disputed the notion that RCTs are demanded to make claims about the effectiveness of sexual offender treatment. Marshall and Marshall (2007) claimed

that while elegant, RCT studies “are fraught with all kinds of scientifically unacceptable problems when applied in a practical setting with sexual offenders” (p. 178). Similar to writers in the general psychotherapy field (Howard et al., 1996; Westen et al., 2004), Marshall and Marshall noted the limitations of external validity of RCTs, namely, that because they involve controlled variables but fail to control for all possible variables, and standardized implementation (reliance on manuals or other formal treatments which limit clinical flexibility), which could raise questions about their generalizability. Marshall and Pithers (1994) challenged the utility of carefully controlled investigations of treatment effectiveness, writing, “Highly structured outcomes studies requiring clients to take part in time-limited, inflexibly sequenced interventions are likely to underestimate the potential effectiveness of treatment” (p. 22). In particular, in these various writings, Marshall has argued that research that involves structured intervention programs (e.g., involving manuals, uniformity of treatment elements, prescribed (and limited) number of sessions, and duration of treatment) is problematic as such phenomena undermine the potential influence of the therapist. Marshall has argued that RCT designs are “not suitable for determining the effectiveness of sexual offender treatment” (e.g., Marshall & Marshall, 2007; Marshall et al., 2011); rather, he has suggested that treatment be optimized by largely individualizing treatment for offenders and allowing therapists freedom to be responsive to the particular presentations of specific clients. Similarly, Levenson and Prescott (2013), while calling for “accountability” in treatment outcome research, reject a reliance on methodological rigor as compromising “clinical validity,” suggesting that methodological approaches such as RCTs “rarely apply to practice in the field” because results are questionable in translating to therapeutic practice in the “real world.” On its surface, such claims are potentially appealing. However, evidence-based practice for any medical or psychological intervention requires some clear and consistent demonstrations of efficacy of particular treatment approaches, with select offenders under relatively controlled conditions and random assignment. Only once some substantial evidence of general treatment effects is demonstrated would it be appropriate to pursue subsequent investigation of whether, in fact, results of more individualized treatment elements (such as therapist variables; longer, more flexible, and intensive treatment programs; greater focus on personality issues and diatheses) be more systematically investigated.

Alexander (1999) while noting that research in sexual offender treatment outcome “remains in the formative stages” claimed, “Should offender treatment be abandoned until its efficacy is incontrovertibly established? While this course may be tempting from a scientific perspective, the public safety ramifications of withholding even relatively ineffective treatment from dangerous offenders cannot be

risked” (p. 112). More recently, Marshall (in various publications, e.g., Marshall & McGuire, 2003; Marshall et al., 2011) has argued that a consideration of effect sizes generally would indicate that even if an intervention has a small effect on outcome, it should be considered potentially useful in that it may lead to “harm reduction” (e.g., reduce sexual offense recidivism for some offenders and/or limit the number of victims among high-frequency offenders). He refers to the reported effect sizes of several SRs of sexual offender treatment and of the Hanson et al. (2002) meta-analysis of such studies as suggesting that the effects of studies utilizing incidental assignment allow the conclusion that sexual offender treatment is effective for some sexual offenders. He describes these results both as “encouraging” and as “convincingly demonstrate[ing]” that such interventions are effective. Certainly, the medical outcome literature has shown that when interventions truly show small effect sizes, they can have substantial practical value; however, in such cases (e.g., a daily aspirin is a common example), that might occur if a treatment is relatively inexpensive, is easy to execute, is politically feasible, and can be employed on a large scale so that it affects a large number of individuals. Assuming it was politically feasible, it is unlikely that sexual offender treatment can be delivered in a manner that is easy to execute, particularly in an inexpensive fashion to many or most sexual offenders. More importantly, Marshall’s argument on behalf of potential small effect sizes is predicated on what currently is an inaccurate or unproved presumption, namely, that the sexual offender treatment literature actually or “truly” shows a positive effect size, even a “small” one. However, as Hanson et al. (2002) and Losel and Schmucker (2005) demonstrated existing RCTs of sexual offender treatment have *not* shown positive effect sizes, let alone even “small” ones: if the mean recidivism score of the treatment group is significantly reduced by excluding higher-risk offenders and the mean recidivism score of the control group is significantly inflated by including excessive high-risk sexual offenders, then the resultant effect size becomes effectively zero. Consequently, if the effect size is minimal or nonexistent—reflective of the lack of significant differences in RCT comparisons—the harm reduction argument is significantly diminished or eliminated; it becomes moot. Seto et al. (2005) and Duggan and Dennis (2014) have effectively responded to all the concerns raised by Marshall regarding RCTs.

Alternately, arguments have been made that in correctional settings, RCTs for psychosocial interventions are not easily implemented and have been shown to make minimal differences in outcome results (e.g., Landenberger & Lipsey, 2005). However, they are clearly possible (e.g., Davidson et al. 2009; Cullen et al., 2011). In a Cochrane Review regarding CBT’s utility in reducing recidivism among general criminal offenders, Lipsey, et al. (2007) made the claim that there

was no difference in results of RCT versus quasi-experimental designs in interventions for criminal recidivism. Yet, regarding their meta-analysis of interventions for general criminality, they noted that only 6/19 RCTs were conducted on “real-world” CBT practice and that a different set of 6/19 RCT studies involved sufficiently high attrition that the validity of their results was compromised. In addition, Lipsey et al. noted that the mean length of the follow-up in most studies of criminal recidivism is rarely longer than 12 months and little information exists about the longer-term effectiveness of such interventions. As other writers (e.g., Sanchez-Meca, 1997) have noted, in the “corrections intervention literature,” most if not all of the studies comparing RCTs to “quasi-experimental” groups have relied on relatively short-term follow-ups and that effect sizes typically diminish with longer follow-up periods. In addition and more generally, Sanchez-Meca (1997) also noted other methodological issues in interpreting meta-analytic results of corrections interventions. First, he pointed out that different outcome measures lead to different effect sizes; recidivism as an outcome tends to have the lowest effect sizes, while “expert” [e.g., clinical] ratings produce the highest effect size. Second, studies with larger sample sizes typically evidence the lowest effect sizes. Third, pretest/posttest designs overestimate effect size in comparison with “between-group” designs; this is particularly problematic given the evidence that posttreatment measurements may be “faked” for purposes of impression management or distorted by ego-syntonic personality characteristics. Thus, studies of general psychosocial interventions in correctional settings have relied upon quasi-experimental control methods and are often characterized by factors that inflate their effect size relative to better-designed studies.

Ultimately, as emphasized previously, many or most investigators and research authorities agree that RCTs are the preferred method for evaluating any treatment’s effectiveness, including studies of correctional samples generally and sexual offenders specifically (e.g., Seto et al., 2008; Hanson et al., 2009) and specifically relied upon by the gatekeepers of approved interventions and funding stakeholders. However, while sex offense (and other criminal) recidivism has been the primary focus of existing treatment outcome studies, several other research methods have been suggested as an alternative means to evaluate treatment effectiveness. Several writers have noted that simply completing sexual offender treatment provides no guarantee that meaningful personal changes have occurred for treatment participants. Alternately, other writers have argued that reduced recidivism is a too absolute and stringent requirement to judge the potential success of sexual offender treatment; Levenson and Prescott (2013) have argued:

When measuring sex offender treatment, effectiveness studies have focused almost exclusively on measuring recidivism rates, while other measures of client improvement have been largely

ignored. Certainly, given the harm caused by sexual victimization, decreased recidivism is the salient goal of treatment. But dichotomous recidivism measures as the only outcome of importance limit our ability to define success. Traditionally, measurement of success in other types of psychotherapeutic interventions has included a reduction in the frequency, duration and intensity of distressing symptoms, or the increase of desirable behaviours. Such appraisals are relative measures. In contrast, sexual offender treatment outcomes evaluate only recidivism, which is an absolute measure. Recidivism as the only construct of improvement within sexual offender treatment almost surely sets everyone up for failure—sexual offenders, clinicians and the field as a whole. (p. 3)... By measuring only arrests and convictions as therapy outcomes, do we ignore information about other ways that an offender’s risk may diminish with treatment? Researchers should consider incorporating relative measures of behavioural change in addition to the absolute measure of recidivism.

One can certainly agree that dimensions of personal change have relevance to sexual offender treatment, particularly if and when a sexual offender is a sole or primary stakeholder in psychotherapy. However, to the extent that the public is a stakeholder, the likely victim of failed or inadequate sexual offender treatment, and the source of funding for such treatment, reducing sexual offense recidivism should be the principal aim of such psychosocial interventions. As Prentky et al. wrote (2011), “...the most compelling reason for treating sex offenders is reducing the likelihood that those offenders will reoffend and create additional victims. The primary goal of sex offender treatment is not to cure sexual offenders or to make them feel better but (a) to reduce the risk that they will reoffend, and (b) to assist with the optimal management of those sexual offenders who are in the community” (p. 117).

In fact, numerous investigations have attempted to examine relative change as a result of sexual offender treatment as means of obtaining perspective as to the efficacy of sexual offender treatment for select sexual offenders. One such approach to outcome research is the measurement of change of putative risk factors believed to be the mediators of sexual offending. Harkins and Beech (2007) reviewed different methodologies utilized to measure the effectiveness of sexual offender treatment suggest that multiple methods have both weaknesses and advantages. They questioned whether distal outcomes such as recidivism should be the only means of determining positive sexual offender treatment outcome. Harkins and Beech suggested that the examination of more proximate outcomes, such as apparent changes within treatment (e.g., intraindividual changes), might allow the comparison of those offenders apparently “successfully” and “unsuccessfully” treated; Hanson (1997) had previously noted that a potential indicator of treatment effectiveness might be to assess within-treatment changes on the typical elements that sexual offender treatment therapists presumably target in their work. Participation in a focused sexual offender treatment (such as one incorporating CBT principles

and techniques) is theorized to produce changes in a sex offender's cognitions, behavior, and affective experiences. If that were the case, it would be presumed that treatment would produce valid proximal changes in treatment targets (which in turn would be associated with more distal changes in a more global outcome measure, namely, sex offense recidivism). Change on treatment targets is typically measured by comparing difference between the treatment and the control group (via a "difference" score created between mean pre-treatment scores of variables of interest and mean posttreatment scores). In particular, the so-called risk principle would be expected to be particularly operative; as Olver et al. explained, "as would be predicted by the risk principle, higher-risk individuals, that is, those who have more 'room' to lower their risk, are expected to show more risk reductions in treatment when compared to lower risk individuals, whose potential for risk reduction would be limited by the 'floor effect'" (p. 114). An additional step would be to examine the possible association of differences in intraindividual pre- and posttreatment measures and sex offense recidivism. Historically, CBT was initially studied by determining if specific techniques did, in fact, modify particular targets of intervention in treatment outcome participants. In addition, to serving as another important outcome measure, such assessments might also shed light on what targets of treatment might be mediators of intervention and associated with larger positive treatment outcomes.

However, there are several issues with this proposed method. Hanson (1997) pointed out the primary behavior of interest (sexual offending) would not be expected to occur in most treatment settings; consequently, potential within-treatment changes on primary behavior of interest would be difficult to measure in institutional settings (e.g., with no or limited contact with children or adolescent and adult females). In addition, Harkins and Beech (2007) also noted that the meaning of any identified change would be dependent upon the sensitivity and validity of the measures of such change. As noted earlier for psychotherapy in general, Gregerson et al. (2001) looked at ratings of treatment made pre- and posttreatment. They found that the difference in the size of treatments of pre- and posttreatment suggests that retrospective (post) evaluations of treatment change "overestimated treatment effects" by a *factor of two* compared to actual pre-/post-measurements. Further, there are several issues with regard to the validity of measurement of potential change for sexual offenders. First, in general, Kelly (2000) showed treatment participants generally tend to present themselves to therapists in a socially desirable manner; from a forensic therapy perspective (e.g., with potential sanctions for perceived noncompliance), this would seem likely to be substantially more characteristic of treatment client/offenders. In addition, many of the test or measures for potential outcome or change utilized by extant studies are extremely

face valid, such that it is likely clear to an offender what the socially desirable or even expected response might be from the perspective of a therapist or treatment program; Marshall and Eccles (1991) pointed out that the majority of instruments used in measuring select aspects of sexual offenders are relatively transparent and it is relatively obvious in identifying the socially acceptable responses. Gannon and Polaschek (2005) hypothesized that relative to self-report measures, "It may be naïve to assume that offenders will not fake good following treatment. A compelling argument can be made that after (post) treatment, offenders have even more incentive to fake good than they did previously. After all, if they don't demonstrate change after treatment then maybe they are not ready for release, or perhaps the therapist, with whom they may have developed strong bonds will be displeased with lack of change" (p. 196). Gannon and Polaschek (2005) found that evidence for this phenomenon was supported. In a later study, these authors found that when sexual offenders believed they were subject to a polygraph, they admitted to increase offense-supportive cognitive distortions relative to their previous reports and those of a control group, thus suggesting that their report of change was little more than impression management of their clinicians (Gannon et al., 2007).

A particularly critical question has to do with whether relative change as measured by pre-post results of testing is, in fact, associated with sexual offense recidivism and might indicate potential mediators of personal change. As Olver et al. (2013) noted: "Aside from a small collection of studies (e.g., Beggs & Grace, 2011; Olver, Wong, Nicholaichuk, & Gordon, 2007, 2013; Wakeling et al., 2013), remarkably little research has explicitly examined linkages between treatment-related changes in important sexual offender risk-need domains and possible reductions in recidivism." Further, what research that is available has found that pretreatment scores are more predictive than posttreatment or change measures (posttreatment scores-pretreatment scores). That is, almost universally, pretreatment information is more strongly associated with the degree of sexual reoffending after treatment. Marshall and Barbaree (1990) found that clients demonstrated reduced deviant sexual arousals (DSA) at the end of treatment but found that neither pre-, posttreatment, nor change scores of DSA were associated with sexual offense recidivism for either rapists or child molesters. Beggs and Grace (2011) reported that for a group of low-risk child molesters treated with CBT, several measures of treatment gain were associated with small reductions in recidivism for up to a 12-year follow-up (even controlling for pretreatment scores). However, they pointed out that correlations between change scores and recidivism were near zero and stated "Given the transparent nature of the tests and incentives for the men to show improvement, it is likely that much, if not most of the self reported gains were due to impression

management” (p. 9). As Beggs (2010) noted, the significance of secondary gain for sexual offender treatment participants (e.g., early release from institutions or favorable parole boards) and the potential lack of intrinsic motivation for change or treatment may obscure any potential true treatment effects on individuals. Beggs and Grace also noted that their results could mean that offenders who performed better in the program might have actually been at lower risk to begin with. Further and more broadly, as Miner (1997) noted, “There is a tendency for test scores to regress from the extremes to the mean. Thus, changes in measures from beginning of treatment to end may be simply an indication of regression to the mean rather than actual change in the construct being measure” (p. 98). Perhaps, an even more important issue in utilizing pre-post changes in proposed outcome measures is that persons who drop out or are removed from sexual offender treatment are not available to provide posttest outcome measures; this differential availability of offenders is likely to inflate positive results from interventions.

Marques et al. (2005) found that self-reported cognitive distortions and self-reported sexual arousal to children and rape were significantly lower after treatment than before treatment. Williams et al. (2007) found that sexual offenders showed significant improvement on almost all self-reported measures of treatment change including denial, minimization, cognitive distortions, empathy, relapse prevention strategies, and self-esteem. Of note, the largest effect size was for relapse prevention strategies, followed by empathy for victims. Williams et al. also examined the association between risk and change and found that no risk group showed significantly more or less improvement than other risk groups. However, there was no control group, and there was no determination as to what level of these measures individual offenders endorsed prior to exposure to treatment. In combination, then, sexual offenders’ responses on self-report measure “appear” to improve, regardless of determined risk level; however, without accounting for individual pretreatment scores, having a control group to determine in what ways offenders’ responses change at a second assessment point, or demonstrating an association with decreased sexual offense recidivism, such studies provide little information about the “meaning” of reported “improvement” in self-report of sexual offenders. Thus, it remains unclear if reported change in self-report measures, particularly victim empathy and relapse prevention, is simply about impression management of a clinical team or other public agents.

McGrath et al. (2012) reported that ratings on the Sex Offender Treatment Intervention and Progress Scale (SOTIPS) made at 1, 7, and 13 months after community-based treatment began predicted sexual recidivism at the follow-up (after starting sexual offender treatment) for a group of predominantly (87 %) first-time sexual offenders

(e.g., mean Static-99R score=0.2, SD 2.4). In a repeated measure design, group SOTIPS ratings by therapists and supervision officers, on their own, were predictive of sexual offense recidivism during the short follow-up period; however, offenders were generally rated as showing improvement over time on the measure, and no change scores were apparently utilized. The results were found for sexual offenders against children but not for those with adult victims.

Further, as Nunes et al. (2011) pointed out, such group-level analyses of treatment change are not sensitive to the presence of non-dysfunctional posttreatment status, specifically clinical significance (e.g., did the client reach some target level of function as a result of treatment and whether the amount of improvement found was large than what would be expected by chance alone). Nunes et al. studied treatment change both in the group and individual level. They found, generally, that the results from group-level analyses were more supportive of “change” than those from individual-level analyses. Thus, measurements of the changes for specific individuals indicated more modest gains, with approximately one-third of participants showing reliable change and reaching functional levels posttreatment on specific measures. Nunes et al. also showed that group-level findings of presumed treatment differences were not always consistent with individual-level findings. They also noted a number of methodological issues that might qualify their results.

Change on treatment targets may also be measured in terms of *clinical significance* (the degree to which self-reported measures fall in the “normal” range for a particular variable or measure) to determine if a client is characterized by meaningful improvement during treatment. Mandeville-Norden et al. (2008) examined pre- and posttreatment measures of cognitive distortion, emotional identification with children, victim empathy, self-esteem, loneliness, underassertiveness, ability to cope with negative feelings, and locus of control. They compared used norms on those measures based on correctional officers to compare treated sexual offenders. They found that between 51 and 71 % of sexual offenders (depending on the particular measure) had scores in the “functional” range after treatment. Mandeville-Norden et al. (2008) also tested to determine if this was a reliable change (e.g., not due to chance); they found clinically significant improvement had been achieved by 7–26 % of offenders (depending on the particular measure). However, these investigators failed to separate offenders who self-reported already functional scores at pretreatment from those who were dysfunctional at pretreatment. Thus, the proportion of participants who were in the functional range posttreatment would overestimate the effectiveness of treatment since a significant number were reporting “functional self-reports prior to treatment. More recently, Barnett et al. (2012) also found in a large sample of sexual offenders who received

a mean of 14 months of community-based sexual offender treatment that posttreatment psychometric test scores were less discriminative and less predictive of reconviction than were pretreatment scores; further, when tests were grouped into dynamic risk domains, only the pretreatment scores of the domain-labeled socioaffective function predicted recidivism. They concluded “the poor performance of these measures posttreatment suggests that treatment providers should rely less on these scores as way of assessing risk after treatment” (p. 23). Similarly, based on a similar study of potential measures of treatment change, Olver et al. (2014) concluded “The results from the present sample generally do not support using most of these self-report psychometric measures to assess sexual offender risk or predict recidivism” (p. 13).

Olver et al. (2007) included measures of possible treatment change rated by therapists from records and indicated that the dynamic change measure added incrementally to a static measure of risk of sexual offense recidivism. Similarly, Beggs and Grace (2010) also found that the same “dynamic” scale made independent contributions to risk assessment beyond that of static factors. However, Beggs and Grace noted that the greater association of the dynamic scale might simply reflect its increased breadth and comprehensiveness (e.g., more than twice as many individual items). Beggs (2010) provided a review of within-treatment outcome among sexual offenders. She noted that there was relatively little research yet conducted on possible proximal treatment outcome among sexual offenders. She pointed out that such outcome if based on self-report might be problematic given the transparency of self-report measures and their openness to social desirability bias responding. Beggs concluded that “Overall, it can be seen that as yet there is a lack of reliable and consistent findings linking within-treatment dynamic change (measured psychometrically) with decreases in recidivism...” (p. 375). She further concluded that evidence for the validity of guided clinical judgment was poor regarding within-treatment outcome. In a later study that showed that suggested that measures of change in treatment were associated with sexual offense recidivism, Beggs and Grace (2011) noted that the association between treatment change and sexual offense recidivism was “relatively modest” and that an explanation for those results might be that offenders who were lower risk to begin with performed better in the program. Finally, Beggs (2010) noted that the results of within-treatment outcome change as measured by idiosyncratic systems of clinical rating were varied, including results within the same studies using multiple operationalization of such outcome. She also pointed out that none of the available studies linked specific treatment changes or individual treatment targets with recidivism so their results did not provide insight into potential mechanisms of change related to sexual

offender treatment. Most recently, Olver et al. (2013) again reported that record-based “change” scores from pre- to posttreatment added incremental value to static variables and showed good predictive accuracy; after sexual offender treatment, they found significant pre-post changes on rated dynamic factors, ranging from small to moderate in magnitude ( $d=0.22-0.62$ ) across various intensity programs. These change scores, in turn, were associated with decreases in sexual offense recidivism; the majority of relationships examined attained significance even after partialing out of pretreatment scores. Thus, there is now recent evidence from one research group utilizing a particular measure that rated treatment change is associated with sexual offense recidivism. Yet as the authors noted, there was no control group, and more importantly, risk scores indicate that it was a predominantly moderate- to low-risk cohort of sexual offenders with a lower base rate of sexual offense recidivism relative to other Canadian samples for similar follow-up periods.

Currently, little evidence currently exists that provides reliable empirical support linking proximal changes in treatment targets with distal changes in sex offense recidivism. To date, treatment progress as measured by difference scores between pre- and posttreatment measures has been found to be a poor predictor of sex offense recidivism (e.g., Hanson & Morton-Bourgon, 2004; Marques et al., 2005). Few studies have found a link between treatment changes and sex offense recidivism (e.g., Beech & Ford, 2006). Langton et al. (2006) found that a rating sexual offense response to treatment failed to predict either serious or sexual recidivism. Similarly, Looman et al. (2005) studied whether an offender’s risk to reoffend was reduced during treatment based on an overall rating of treatment performance (including performance not only in groups and homework assignments but also on the client’s behavior outside of the formal treatment program). However, the performance ratings showed *no* association with posttreatment sexual offense recidivism. Hanson et al. (2008) stated “... much less is known about the processes by which sexual offenders change. Studies frequently find that improvements on factors presumed to be criminogenic have no effect on sexual recidivism rates” (p. 887). In effect, only Olver et al. (2013) reported that after controlling for risk, change scores (total and sexual deviance) were associated with decreases in sexual offense recidivism. However, they noted: “There was no untreated control group with pre- and posttreatment VRS-SO ratings in order to compare change over the passage of time with that made with treatment services. As such, there is some possibility that other change agents, aside from treatment, contributed to the changes... we cannot rule out the influence of other change agents (e.g., participation in other programs, aging) that may have contributed to changes in risk” (p. 12). In a recent review,

Wakeling and Barnett (2014) reviewed the relationship between psychometric test scores and reconviction in sexual offenders participating in sexual offender treatment in the UK. They concluded:

We believe that these results suggest that it may be unwise to rely on large batteries of psychometric tests to determine change in treatment and that further research is required before we can be sure of the relationship of psychometric tests to recidivism outcome...it is very unfortunate that we are not yet in a position to make reliable estimates of the extent to which such programs have benefited individual participants. The use of psychometric tests may not be so promising as we once thought. Additionally, *the evidence so far suggests that to use change on psychometric test scores for program evaluation (i.e. as a proxy measure of reconviction outcome) is not warranted.* (p. 143; emphasis added)

Thus, despite the desire to determine “relative change” in sexual offenders as a means of showing treatment effectiveness, little evidence exists that sexual offender treatment outcome can be meaningfully assessed or demonstrated by the use of within-program, pre-post self-report tests and/or questionnaires.

Further, a number of studies have found that posttreatment measures are either not or are *less* predictive of sex offense recidivism than pretreatment measures. In a key early study, Quinsey (1983) first reported that neither changes in deviant sexual interest indices nor posttreatment deviance measure was associated with subsequent recidivism in treatment sexual offenders. Subsequently, Rice et al. (1991) found that pretreatment measures of deviant sexual arousal were better predictors of sex offense recidivism than posttreatment measures, raising questions about what those posttreatment measures actually assessed. Marshall and Barbaree (1990) also reported that neither pretreatment, posttreatment, nor apparent changes in reducing deviant sexual interests were related to treatment outcome. Thus, it remains unclear if sexual offenders actually change as a result of sexual offender treatment. Langton et al. (2006) noted:

For some sex offenders, ratings of treatment progress in later, follow-up programs may prove unreliable indicators of any gains made...they may obscure the validity of ratings made for participation in earlier/initial treatment programs as offenders become familiar with program content and expectations and strive to appear compliant and ‘treated’ in order to be found eligible for parole or relaxation of supervision intensity...The challenge is, in part, one of measurement. Because sex offenders learn, indeed are expected to learn, the terms and concepts of CBT and relapse prevention, determining the veracity of their presentations will be difficult. (p. 116)

Consequently, while sexual offenders who participate in sexual offender treatment may learn the information related to and terms of sexual offender treatment such that they can answer self-report and even interview questions to reflect such information acquisition, their “internalization” or intent to use that information may remain unchanged.

In a recent paper, Rice et al. (2013) concluded:

While research on this issue is preliminary, evidence suggests that, given a comprehensive set of valid static, historical factors, pre-release difference scores afford minimal incremental validity (Olver & Wong, 2011; Olver et al., 2007). Again, we conclude this is due to pre-release risk-relevant change on these constructs indexing the same aspects of temperament and personality that are reflected by established static, historical risk factors.... (p. 10)

For example, Rice et al. (2013) noted that the results of Olver et al. (2007) showed that static risk scores were more predictive than either dynamic pre- or posttreatment scores. Olver et al. (2013) were the first and only group to report that “change” scores from pre- to posttreatment added incremental value to static variables and showed good predictive accuracy. After sexual offender treatment, Olver et al. (2013) found significant pre-post changes on select observer-rated dynamic factors, ranging from small to moderate in magnitude ( $d=0.22$  to  $0.62$ ) across various intensity programs. These change scores, in turn, were associated with decreases in sexual offense recidivism; the majority of relationships examined attained significance even after partialing out of pretreatment scores. Thus, there is now recent evidence from one research group utilizing a particular measure. However, more broadly, Serin et al. (2013) reported in a review of intraindividual changes in criminal offenders following interventions: “It is apparent within this review that therapeutic change does not consistently lead to reduced likelihood of future crime” (p. 50). They stated:

It is especially difficult to defend programs when apparent successful adoption of treatment skills does not translate into a definitive lower risk to reoffend. However, it is also difficult to defend successful programs when it is unclear which treatment elements are responsible for presumed or “perceived” change and which offenders might have changed. (p. 50)

As Wakeling and Barnett (2014) also concluded: “Pretreatment psychometric scores appear to have a better relationship [with sexual offense recidivism] than those gained post-treatment, suggesting the former should be preferred to the latter when assessing risk of recidivism outcome...it may be that the [posttreatment] results are negatively impacted by desirable responding...” (p. 143). They recommended that future efforts be directed at developing reliable and valid measures of “risk domains” as opposed to specific risk factors. In short, the fact that pretreatment measures are more predictive of sexual offense recidivism than posttreatment self-report and clinician ratings provides further evidence that even apparent change reported by sexual offenders or perceived by their treatment providers may well not be genuine and that as the SOTEP identified (e.g., Marques, Nelson, Alarcon, & Day, 2000; Marques et al., 2005), sexual offenders can learn the language and “display” motivation while in a sexual offender



treatment program, but fail to demonstrate that motivation or enact purportedly learned skills once returned to the community even with aftercare and supervision.

Of particular interest is the ability of sexual offender treatment therapists to offer a meaningful or valid perspective on the relative progress of their sexual offender clients. Unfortunately, most available studies indicate that sexual offender treatment clinicians' opinions about their clients are not informative. In more general psychotherapy literature, clinicians have typically been found to be poor judges of treatment progress. Thus, research has shown that therapists' ratings of clients' progress are significantly greater than what their clients report or what is reported by client's significant others (e.g., Hill & Lambert, 2004); this is likely to be even more pronounced in forensic therapy settings, where there is significant potential secondary gain for sexual offenders who present as reflecting positive treatment behavior and apparent treatment gains. Further, as noted previously, in their review, Hill and Lambert concluded that therapist ratings of treatment outcome and global ratings of change are associated with the "perception of greater effectiveness" of treatment compared to more specific measures and more distal measures. In their review, Hill and Lambert also pointed out that data from therapists or expert judges who are aware of the treatment status of clients produce larger positive ratings than those from virtually all other sources. Walfish et al. (2012) showed that clinicians providing psychotherapy tended to overestimate the rates of their client improvement relative to their own perceived clinical skills. The same seems to be particularly true for therapists in sexual offender treatment programs; this is a finding that has been replicated over time. Quinsey et al. (1998) first found that therapists' judgments about treatment progress were unrelated or negatively related to recidivism. Similar results were found by Seto and Barbaree (1999) in the initial analyses of a research sample but not found in a somewhat expanded sample from the same source (e.g., Barbaree, 2006). Marshall and Eccles (1991) also opined that generally clinicians' judgments of treatment effectiveness (as well as those of offender clients) were considered unreliable. Citing a variety of earlier studies, Hanson and Harris (2000) noted: "Experienced clinicians are frequently unable to differentiate between sexual offenders who benefited from treatment and those who did not..." (p. 7). In a short-term prospective study, Seager et al. (2004) showed that clinical judgments of treatment (even guided by specific clinical criteria) were unrelated to recidivism failure. Specifically, they found that positive evaluations of treatment changes in posttreatment assessments (e.g., such as quality of disclosure and perceived enhanced victim empathy) showed no correlation with sex offense recidivism. They found that "...clinical judgments of treatment change, although guided by specific clinical criteria, were unrelated to recidivism failure...Narrative commentary

on treatment participation appears superfluous in the context of predicting recidivism of rates. Quality of participation appears unrelated to recidivism" (p. 610). Thus, they found that positive evaluations of treatment change such as the quality of disclosure and perceived increased victim empathy found in posttreatment assessments did not correlate with recidivism. Seager et al. (2004) concluded, "...sex offender programs are not changing psychological characteristics that affect recidivism" (p. 610). Hanson and Bussiere (1998) found that most clinical measures of treatment progress were unrelated to sex offense recidivism as did the updated meta-analysis for risk factors for sex offense recidivism by Hanson and Morton-Bourgon (2005). They found that poor progress in sexual offender treatment, measured at the end of such intervention, was unrelated to sexual reoffending. In addition, the aforementioned general and specific findings regarding that allegiance to a treatment model was associated with more positive finding (e.g., Losel and Schmucker, 2005) suggests that clinicians' belief in that model may well account for more variance in outcome than any specific interventions or changes by treatment participants. This corresponds to the consistent finding in the general treatment outcome literature that researcher/therapist allegiance accounts for a significant amount of the outcome in treatment studies that find particular interventions effective. Consequently, there are both empirical and theoretical reasons to view therapist ratings of personal change and individualized risk reduction as non-empirically supported and not particularly useful in gauging the outcome or psychotherapies for sexual offenders.

It should be noted that some writers (e.g., Levenson & Prescott, 2013) call attention to a specific observation reported by Marques et al. (2005), relative to sexual offenders who "got it" or were seen as benefiting from treatment provided as having lower sexual offense recidivism. Several points are worth noting. First, "When the Got It scores of sexual recidivists were compared to those of non-recidivists, no significant differences were found..." When they the investigators employed a median split, the trend was still not significant. No differences in reoffending were found between low and medium treated sexual offenders. However, they reported that "high-risk offenders," "largely accounted for by child molesters" who "got it," showed lower sexual reoffending after treatment. However, per their results, there was only *one* "high" risk sexual offender (a total of 7 treated sexual offenders who got it), who was responsible for their claim of decreased sexual offense recidivism.

In summary, alternative ways of assessing outcome for the efficacy of sexual offender treatment (in contrast to reduced recidivism) also do not provide support for psychotherapy, particularly as a means to measure individual change. While some change is evident on self-report measures in certain instances, most of those measures are quite

transparent, and to date, no consistent, replicated association between pretreatment or “change” scores and recidivism has been demonstrated. Further, clinician-rated improvement in sexual offenders as a function of treatment appears to provide an overly positive view of change. Rather, pretreatment and essentially static constructs show the strongest association with sexual offense recidivism.

### **The Efficacy of Sexual Offender Treatment for Higher-Risk Sex Offenders**

There is a profound lack of information about the RCT-based effectiveness of sexual offender treatment with higher-risk/high-need sexual offender, both from an actuarial perspective and those with a greater degree of criminogenic needs. Not unexpectedly, no RCT of any psychosocial treatment exists at this date for such a subset of sexual offenders. Hall (1995) noted that the most severe sexual offenders were typically excluded from treatment in the studies he reviewed. In the most comprehensive RCT to date, Marques et al. (2005) excluded any sexual offender with more than two prior felonies; thus, the treatment candidates were generally a low- or moderate-risk group to begin with (77 % feel into that category and only 22 % were deemed “high risk.” They excluded sexual offenders with more than two prior felonies, major mental disorders, and lower IQ and/or those who had displayed severe management problems while in prison (they also excluded any offenders who denied their sexual offense from the “volunteer” group). The SBU review from 2011 stated “Unfortunately, no studies have assessed the effects of treating high-risk individuals who have not sexually offended against children” (p. 22). It is notable that numerous studies of sexual offender treatment have systematically excluded high-risk/high-need sexual offenders. That is, the sexual offender treatment outcome literature is marked by “sample censorship” or exclusion for higher-risk sex offenders. Thus, Hall (1995) noted that many more severe sexual offenders were not even offered sexual offender treatment, while the SOTEP study did not include a significantly large group of “high-risk” sex offenders. In SOTEP, most treatment subjects were first-time sex offenders, of low- or moderate-risk groups that per RNR criminological models of intervention should have responded best to intervention and shown decreased sex offender recidivism rates. In short, most of the existing treatment outcome literature relates to low- or moderate-risk sexual offenders; thus, that evidence indicates that sexual offender treatment has not been demonstrated via RCT to be effective with such offenders.

However, what literature does exist indicates that sexual offender treatment is not effective or, at best, is substantially less effective with higher-risk sexual offenders. Both the Hanson et al. (2002) and Losel and Schmucker’s (2005)

meta-analyses found offenders referred to treatment based on *perceived need* had significantly higher sexual recidivism rates compared to offenders considered not to need treatment. It will be recalled that Hanson et al. (2002) found that “Offenders referred to treatment based on perceived need had significantly higher sexual recidivism rates than the offenders considered not to need treatment” (p. 182). The odds ratio was 3.4 (with an outlier study removed), and there was no significant variability, indicating that this was a robust phenomenon; thus, sex offenders viewed as high need provided sexual offender treatment reoffended over three times the rate of untreated sex offenders. Olver et al. (2011) identified that high-risk/high-need offenders are those persons least likely to complete treatment, presenting with a number of specific responsivity issues such as low motivation, poor engagement, and disruptive behavior. Over just a 2-year follow-up, Friendship et al. (2003) found that high-risk offenders were six times more likely to be reconvicted of a new sexual and/or violent offense than low-risk offenders. Stirpe, Wilson, and Long (2001) found that higher-risk sex offenders who received sexual offender treatment did not maintain motivation over time in the community. In their review, Rice and Harris (2003) also considered Hanson et al.’s (2002) group of studies involving “assignment based on need” and emphasized that the overall odds ratio was 3.0 for sexual recidivism; studies of these offenders “indicated that those [of greater perceived need] who were treated reoffended over three times the rate of the untreated” (p. 434). Olver et al. (2001) applied MA and found that general criminal offender treatment non-completers (e.g., those who started but dropped out) were higher-risk offenders and rates increased when pretreatment attrition was also included. Olver et al. (2013) also did not find a risk by treatment interaction. Rice and Harris also noted that some “assignment based on need” studies, in effect, controlled for static factors before examining whether the treatment added anything to the assessment of outcome (a methodological plus) but still showed no recidivism lowering effect of sexual offender treatment for higher need sexual offenders. Stirpe et al. (2011) reported that RP-related treatment components showed a steady increase from pretreatment throughout follow-up in the community for low- or moderate-risk offenders, but *not* for high-risk offenders. Both groups improved substantially in level of motivation from pretreatment to posttreatment; however, only those in the low- or moderate-risk group maintained their motivation levels once released to the community; that is, higher-risk sexual offenders did not maintain motivation once released to the community.

Similar to all other presenting problems, it must be the case that some sexual offenders are characterized by sufficient severity, chronicity, and/or a large number of risk factors (as predisposing or maintaining factors). As the larger psychotherapy literature clearly indicates, more severe,

chronic problems are quite resistant to the effects of psychotherapy generally; more typically, more minor changes, at best, result from such interventions and may not be retained. Fifteen years ago, Harris et al. (1998) wrote, "The idea that a high-risk sex offender can be converted into a low-risk offender through the application of treatment or through progress in treatment simply has no empirical support from the literature taken as a whole" (p. 106). Further, as suggested by various writers (e.g., Rice et al., 1999), there may be some offenders whose risk level is sufficiently high that no psychotherapy could reasonably be expected to reduce it to a level at which release to the community could be recommended and that "The idea that a high-risk offender...can be changed into a low risk offender through treatment or through progress in treatment simply has no empirical support from the literature taken as a whole" (p. 305). To date, little data has accrued that undermines that contention. Thus, as some writers have stated, "...it is important to note that there are some sex offenders whose risk level is so high that no treatment could reasonably be expected to lower it to a level where release to the community could be recommended" (e.g., Harris et al., 1998, p. 106). At the risk of repetition, regarding the Hanson et al. meta-analysis of 2002, Berliner (2002) pointed out: "It is not at all clear that these results can be generalized to the highest risk offenders. Even if they could be applied to these offenders, a moderate effect size reduction would still mean that high-risk offenders continue to be dangerous" (p. 196). In fact, as Hanson et al. (2008, 2009) reported, the risk principle of the RNR model was not confirmed by their meta-analysis; the most high-risk sexual offenders did not respond significantly better to sexual offender treatment relative to lower-risk sexual offenders. [This is actually similar to a work with criminal recidivism where the risk principle showed the smallest effect of the RNR dimensions in relapse prevention programs for criminal offenders (Dowden, Antonowicz, & Andrews, 2003)]. Ten years ago, Rice, Harris, and Quinsey (2001) concluded:

The idea that a high-risk offender (especially focus serious offenders as serial sexual murderers) can be changed into a low-risk offender through treatment or through progress in treatment simply has no empirical support from the literature taken as a whole...it is also important to point out that there may be some offenders whose risk level is so high that no treatment could reasonably be expected to lower it to a level at which release to the community could be recommended. (pp. 305–306)

At the present time, nothing in the empirical or scientific literature has emerged that would support the belief that such conclusions would be different for the general high-risk/high-need sexual offender. Rather, each individual's experiences in psychosocial and other adjunctive treatments would need to be carefully considered to offer a well-documented, person-specific opinion that a particular high-risk/high-need

sexual offender has changed substantively as a result of such psychotherapeutic efforts.

The treatment of sexual offenders with higher levels of psychopathy, as a specific subset of likely higher-risk sexual offenders, has received research attention. Generally, there has been a pessimistic view that persons with a higher degree of psychopathy can be successfully treated to reduce their potential for future violence, including sexual offending. Ogloff et al. (2013) reported that psychopathic traits are associated with negative behaviors in treatment. More recently, both Langton et al. (2006) and Looman et al. (2005) found that, despite sexual offender treatment, more psychopathic sexual offenders (e.g., with PCL-R scores  $\geq 25$ ) reoffended in sexually and/or violently at significantly higher rates than those with lower scores; however, they suggested that there may be a subset of psychopathic sexual offenders who may respond to some interventions. An early study of violent offenders by Rice et al. (1992) found that persons with elevated levels of psychopathic traits who participated in a therapeutic community while incarcerated subsequently had higher rates of violent recidivism than similarly psychopathic persons who did not participate in such an intervention. A similar finding was made by Seto and Barbaree (1999); later studies by Langton et al. (2006) and Looman et al. (2005) did not find an interaction between psychopathy and treatment for increased recidivism. Thornton and Blud reported that both of the aforementioned studies showed that "...offenders in whom higher levels of psychopathy were combined with 'good' treatment performance had worse rates of serious recidivism" (p. 517). Olver and Wong (2009) maintained that with appropriate treatment interventions, sex offenders with significant psychopathic traits can be retained in correctional treatment program and those showing therapeutic improvement can reduce their risk of both sexual and violent recidivism. Doren and Yates (2008) reviewed the effectiveness of sexual offender treatment for psychopathic sexual offenders. They concluded that (1) sexual offender treatment does not appear effective in lowering serious recidivism and (2) sexual offense recidivism rates were variable for treated psychopaths, but there were indications that some psychopaths did show decreased recidivism after treatment. However, the available research did not indicate which psychopathic sexual offenders benefited from sexual offender treatment and which did not. Doren and Yates (2008) also noted: "The present qualitative analysis also clearly found a consistent absence of untreated comparisons groups in all studies. Hence, no conclusion can be drawn from existing research about the degree to which psychopath offenders benefit from sexual offender treatment" (p. 354); thus, again methodological factors precluded drawing absolute conclusions. Several writers have suggested that the most successful interventions for psychopathic offenders are likely to be characterized by high structure, high intensity, and extended duration, a high degree of

involvement by mental health professionals and increased attention to responsivity and, particularly, to maintain such offenders in the treatment process (e.g., Salekin, 2002; Olver and Wong, 2009; Thornton & Blud, 2007). Currently, the most appropriate perspective appears to be that perhaps some sexual offenders with psychopathic traits may respond somewhat differentially to CBT, with some more psychopathic offenders showing a more positive response to such intervention. Thornton and Blud (2007) both review a significant set of factors that would likely lead to the poor outcomes typically found in treating psychopathic offenders; they also offer suggestions about possible aspects of intervention that might lead to more positive outcomes with more psychopathic sexual offenders. In a more pessimistic vein, Harris and Rice (2006) stated:

We believe there is no evidence that any treatments yet applied to psychopaths have been shown to be effective in reducing violence or crime... We believe that the reason for these findings is that psychopaths are fundamentally different from other offenders and that there is nothing 'wrong' in the manner of a deficit or impairment that therapy can 'fix.' (p. 568)

Ultimately, whether more psychopathic sexual offenders are amenable to psychosocial treatment is an empirical question; since no RCTs of relatively psychopathic sexual offenders have yet been conducted since the review by Doren and Yates (2008), their conclusion remains the same: the available evidence is not generally positive, but no absolute conclusions can be drawn in the absence of scientifically valid research.

If sexual offender treatment cannot be demonstrated to be effective in RCTs with more motivated sexual offenders with fewer comorbid conditions and lower severity of "problems" (e.g., fewer victims, lower density of risk factors and/or criminogenic needs), then even more serious questions are raised about its potential utility for "higher"-risk sexual offenders characterized by entrenched maladaptive behavior patterns maintained by a greater number and severity of risk factors and predisposing conditions. To date, of the various government programs in various jurisdictions that have detained violent sexual offenders (e.g., civil commitment of the so-called Sexually Violent Predators in the USA, Dangerous Offender Programs in Canada, and the Dangerous and Severe Personality Disorder Program in the UK), no data are available—no studies have been published—as to whether more intensive and long-term treatment of high-risk/high-need sexual offenders show reductions in sex offense recidivism as a specific result of treatment received while detained. Consequently, little useful information exists to establish the efficacy of psychosocial interventions in the management of moderate and high-risk/high-need sexual offenders; there are no RCTs of high-risk/high-need sexual offenders; empirically, it is simply unknown as to what the components, other treatment

factors, and the length and density of psychosocial treatment are necessary to reduce such offenders' likelihood for sexual offense recidivism. Further, even with studies of the treatment of high-risk/high-need sexual offenders while detained, given that many of these individuals will only be released back to the community under terms of intensive and long-term supervision, it may not be possible to isolate the effects of sexual offender treatment generally or its components for such individuals.

### Longer-Term Outcomes for Sexual Offender Treatment

Most sexual offender treatment outcome studies have not followed subjects for lengthy periods of time. Consequently, little is known about the longer-term effectiveness of sexual offender treatment. However, all authorities have stated that sex offense recidivism increases with the length of follow-up (e.g., Hanson et al., 2003; Harris & Hanson, 2004; Harris & Rice, 2007). Other studies demonstrate that even some treated sexual offender reoffends after lengthy periods without a detected sexual offense (e.g., Prentky et al. 1998). Given at best small effects for psychosocial treatments for sexual offender, it is important to know to what degree any positive outcomes may persist for such persons. This is particularly important given the general psychotherapy results regarding the diminishing persistence of treatment-related changes. As noted, Barrett, Wilson, and Long (2003) found that treated sexual offenders showed a significant decrease in rated motivation after release to the community. They wrote:

The results of this study clearly show that clinicians in community settings should expect to have difficulty re-engaging offenders in the treatment process and should not assume that a positive institutional report will be reflected in a client' attitude and behavior in the community. (p. 279)

Similarly, Stirpe et al. (2001) found that higher-risk sex offenders who received treatment did not maintain motivation for sexual offender treatment practices when released to the community, and within 3 months after release, apparent treatment gains had diminished for higher-risk sexual offenders when returned to the community (even with the benefit of 3 months of additional treatment in the community). Marques et al. (2005) revealed: "We learned from interviews with the offenders that a number of our treatment failures did not use the self-management skills they acquired in the program, and some did not even accept the basic goals of self-control and relapse avoidance..." (p. 100). In summary, then the available literature would suggest that even if some degree of recidivism-related change initially results from sexual offender treatment, that effect may diminish or is eliminated once treated offenders return to the community just as it does for most mental health problems.

## Conclusions and Future Directions

Reducing or eliminating sexual offender recidivism is an important and desirable goal and one shared by all stakeholders relative to sexual offending. While short-term recidivism for adult sexual offenders consistently appears to be approximately 12–15 %, perspectives on long-term sexual offense recidivism indicate that sexual reoffending increases over longer follow-up periods to perhaps 40 % detected offenses (e.g., Hanson et al., 2003; Harris & Hanson, 2004; Harris & Rice, 2007). For persons already sanctioned at least once previously for sexual offending, this is a very high rate of violent criminal offending. There can be no question that given the severe consequences of sexual victimization, effective and enduring management of sexual offenders is of critical importance. Psychosocial treatments have long been considered a central component of accepted and implemented management strategies—for many practitioners in the field of sexual offender treatment, they have been perceived as *the* critical element of management. All stakeholders agree on the importance of effective management of sexual offenders for community safety; however, the degree to which psychosocial interventions matter by “working” (as well as other management mechanisms) necessarily must be demonstrated. Consequently, the determination of whether psychosocial treatments for sexual offenders have been empirically established as a mechanism to reduce future sexual offending is of critical importance.

Early in 2010, R. Karl Hanson sent an email stating: “I, for one, have done enough meta-analyses of barely acceptable studies. It is time to counter the political resistance to random assignment studies by getting ATSA to endorse a position statement supporting their use” (cited in, Rice et al., 2013). Subsequently, at Dr. Hanson’s recommendation, the Executive Board of the Association for the Treatment of Sexual Abusers (ATSA) proclaimed “After 50 years, the field of sex offender treatment cannot, using generally accepted scientific standards, demonstrate conclusively that effective treatment are available for adult sex offenders” (ATSA, 2010b). More recently, in an editorial, Ho and Ross (2012) criticized public representations regarding the Sex Offender Treatment Programme in the UK and claims that the program “worked;” they wrote:

Twenty years since the SOTP [in the U.K.] was launched, its efficacy has yet to be convincingly demonstrated... Interventions such as the SOTP are too important in terms of financial cost and cost to society for them not to perform as they are claimed to perform. They are too important for the participant men themselves than for anything other than the highest standards of evidence underpin them. In the absence of an enormous effects size, encouraging pilot work and open studies should lead to independently conducted RCTs. (p. 5)

Clearly, at present, there can be little argument with that conclusion. The current review of the scientific evidence of sexual offender treatment is that, at best, minimal evidence currently exists to demonstrate that psychotherapy is effective at reducing sex offense recidivism or at changing sexual offenders in meaningful or substantive ways. The most optimistic perspective that could be gleaned from the existing studies is that the sexual recidivism of select, lower-risk sexual offenders may be lowered when they are treated in community settings; at the same time, as others have written, results could also be interpreted to mean that regardless of treatment, lower-risk sexual offenders (not surprisingly) generally have lower sexual offense rates. In contrast, for higher-risk sexual offenders who are treated in correctional settings, there is no data from controlled trials to suggest a desired reduction in sex offense recidivism can be attributed to psychosocial interventions. In addition, it remains unclear if sexual offender treatment is effective for different types of sexual offender and what, if any, elements of psychotherapy are particularly useful in impacting sexual offenders and sex offense recidivism rates. In actuality, this appears to be the increasing consensus among the experts in the field of sexual offender research.

Perhaps not unexpectedly, the conclusions of this chapter echo those that of several other reviewers of the sexual offender treatment literature. As Furby et al. (1989) stated in their early review of sexual reoffending in both treated and untreated sex offenders, “Many of recidivism studies reviewed here were, unfortunately, not very informative...” (p. 28). In 1999, Gallagher et al. wrote, “The literature on the efficacy of sexual offender treatment programs is inconclusive. The more exhaustive narrative reviews tend to conclude that little current is known due to the considerable methodological weaknesses of the individual evaluations” (p. 19). Hanson et al. (2002) concluded, “we believe that the balance of available evidence suggests that current treatments reduce recidivism, but that firm conclusions wait more and better research” (p. 187). Rice and Harris (2003) have written, “We suspect all would agree that very little knowledge has accumulated about several crucial matters... The dearth of knowledge about sexual offender treatment contrast sharply with the rapid expansion of knowledge in other areas” (p. 437). Noting that methodological factors had a significant effect on their results, Losel and Schmucker (2005) stated: “Bearing the methodological factors in mind, one should draw very cautious conclusions from our meta-analysis... We need more high-quality outcome studies that address specific subgroups of sex offenders as well as more detailed process evaluations on various treatment characteristic and components” (p. 138). Abracen and Looman (2004) opined:

With reference to sex offender recidivism research, more generally it is quite clear that the quality of many research studies has, to date, been relatively poor... Regardless of the difficulty

associated with finding statistical significance, there is little rational for poorly conducted studies. (p. 16)

Harkins and Beech (2007) wrote, “The effectiveness of sex offender treatment has been studied and reviewed extensively...in spite of great effort and numerous studies, this has yet to be conclusively demonstrated” (p. 37). Seto et al. (2008), in their consideration of RCT methodology for sexual offender treatment outcome studies, wrote: “It is possible that some adult sex offender treatments currently being offered are effective; it is also likely that some treatments are ineffective, or worse” (p. 253). Schmucker and Losel (2008) wrote that the field should remain critical of existing research results and that “In order to reach a more definitive answer on the questions’ Does sexual offender treatment work?’ we need more high quality studies” (p. 16). Hanson et al. (2008) wrote “Readers sympathetic to sexual offender rehabilitation may be content with the encouraging findings from weak research designs; however, skeptics will only be compelled to change their opinions by the strongest possible evidence” (p. 887). Over 20 years after Furby et al.’s (1989) SR, the IHE SR in 2010 concluded:

...research on the efficacy/effectiveness of SOT interventions and programs has been slow to mature, and the results have been contradictory...the perceived efficacy/effectiveness and value of SOT program and the views on how best to manage adult male sex offenders have been inconsistent. (p. 32)

They reported that the current evidence showed a small statistically significant treatment effect but that “a lack of high-quality primary research studies...raise uncertainty about which of the available approaches work for adult male sex offenders” (p. 32). Rice and Harris (2013) summarized the current status of the sexual offender treatment outcome literature, stating “The most parsimonious interpretation of findings from weaker designs is that pretreatment differences and other forms of selection bias are responsible for apparent treatment effects” (p. 23). As a form of criminal recidivism, sexual reoffending appears to be a substantially more difficult problem to successfully address than general criminal behavior. Depending on how the data is viewed, to a certain extent, limited scientific data suggests that the reduction of future nonsexual criminal behavior via psychosocial interventions is somewhat more successful than the ability to show reductions in sexual criminal behavior by similar means. [However, most of those empirical results are based on largely CBT models following the RNR approach but are largely dependent on quasi-experimental findings and follow-up periods of 2 years or less (e.g., Bonta & Andrews, 2006; Landenberger & Lipsey, 2005; Latessa & Lowenkamp, 2006).] More generally, an increasing number of writers have raised concern about the indirect harm that can result from inaccurate conclusions drawn about treatment efficacy, noting that an ineffective treatment that is

falsely assumed to be beneficial exacts various costs in terms of both expense and other resources and expectations of clients and other involved or affected parties.

In addition, a number of other key questions have yet to be answered and in some cases have yet to even be addressed. As previous reviews have pointed out, it is unclear what characteristics identify sexual offenders who might respond to psychotherapy initially and maintain any apparent gains at the end of such interventions. As Nunes et al. demonstrated, there are differences between group results and those for individuals. It is likely that particular individual sexual offenders are responsive to psychosocial interventions and do change as a result of them; the question is can those who do change be reliably identified and by what means. Certainly, on an individual level, any sexual offender who has participated in sexual offender treatment should be carefully and comprehensively evaluated to determine if there are substantive grounds in determining a measurable basis for judging treatment progress and potential treatment success. Are there differences in how paraphilic, psychopathic, and/or otherwise personality disordered sexual offenders respond to sexual offender treatment? In addition, it is unclear what elements of sexual offender treatment programs may make significant contributions to potentially positive outcomes for sexual offenders or subgroups of such offenders. To what degree do the intensity, length, and site of treatment, general client characteristics, general therapist characteristics, experience of therapists, the interaction of therapists and clients (and approaches), severity and psychosocial impairment of clients, as well as similar variables relate to the outcome of sexual offender treatment? Beyond what is currently known, are there other identifiers of sexual offenders that distinguish those who refuse, dropout of, or are terminated from sexual offender treatment?

With some distance and a dispassionate perspective, the lack of demonstrated efficacy for sexual offender treatment should not be surprising. Sexual offending lacks a detailed and scientifically defined etiology for either the initiation or maintenance of this particular type of criminal behavior. Both the integrated, multidimensional models of sexual offending lack specificity, and the available empirical studies of risk factors indicate relatively small contributions of multiple, cumulative risk factors. Further, as Hoberman (2013c) notes, the specific treatment elements and delivery of sexual offender treatment lack both an established theoretical and empirical basis; largely, the elements of sexual offender treatment were borrowed from treatments for other presenting problems, and few if any of the elements or delivery issues have been demonstrated to be effective in affecting robust behavioral change. In the context of the general literature on the effectiveness and efficacy of psychotherapy, a number of factors would strongly suggest that psychosocial interventions might have little effect on sexual offense recidivism.

The evidence suggests that psychotherapy is most effective at relieving personal distress, which may be lacking among many sexual offenders. If client variables contribute the most variance to psychotherapy outcome, the multiplicity of personal risk factors acting convergently and cumulatively would make it difficult to affect sexual reoffending. Further, given the likelihood of problem severity and interpersonal difficulties, including the strong association of maladaptive personality traits and disorders with such offending, would suggest that numerous therapy-interfering effects and the general difficulties limiting the change of core signs and symptoms would impose significant limits to potential positive outcomes from psychotherapy. The effects of mandated treatment and a generalized lack of motivation for behavioral change would also impact negatively on the potential “success” of psychosocial interventions. As several authors have noted, perceived high levels of criminogenic needs or risk factors are associated with poor treatment outcomes for sexual offenders. Conversely, the general lack of evidence for outcome effects from specific techniques or strategies would also direct that the varied programs of such techniques and strategies would not be key factors in leading to individual change for many sexual offenders.

The lack of available evidence for the efficacy of sexual offender treatment has a variety of implications for offenders, forensic/clinical practitioners, public safety, and the scientific field of sexual offender management. As Marques et al. (2005) indicated, “Questions about whether and when sex offenders can be treated are extremely important, not just to our field but to victims, policy makers and the public” (p. 104). In turn, this leads to a variety of potential overlapping decision points and courses of action that might be taken relative to the role of psychotherapy in the management of sexual offenders.

Unless evidence of psychosocial treatment effectiveness (defined by accepted scientific practices) is produced, the sexual offender treatment outcome field increasingly runs the risk of being further marginalized and discredited. Lilienfeld (2011) reviewed data that suggests that a large percentage of the public regard the general field of psychology’s scientific status with considerable skepticism. Nasrallah (2013), the editor of *Current Psychiatry*, similarly suggests that “...psychotherapy has never been able to shrug off an unwarranted aura of fuzziness as a legitimate medical intervention...Psychotherapy is sometimes perceived as a scam –that is, a placebo packaged and propagated as treatment” (p. 18). In the face of a lack of a strong scientific basis, ongoing affirmation of the tenuous position that sexual offender treatment is effective for most sexual offenders runs the risk that the sexual offender field will become a poster child for “pseudoscience.” It is notable that, in the 1980s, it was largely the American Psychiatric Association’s decision or belief that sexual offender treatment

was not effective that led to the dismantling of earlier attempts to intervene with such offenders therapeutically and a shift to a relatively exclusive emphasis on correctional management for sexual offenders. As several writers have argued and has been demonstrated by the decreasing research and policy support for sexual offender treatment, there is great danger in promising results that may not be obtainable; the credibility of both practitioners and their professional organizations may suffer long-lasting damage in terms of public mistrust. Marques et al. (2005) stated, “Questions about whether and when sex offenders can be treated are extremely important, not just to our field but to victims, policy makers and the public” (p. 104). In 1999, Alexander wrote: “...public funding for sexual offender research and treatment has declined in the last decades. A poverty of research funds has hampered improved understanding of the effectiveness of various offender treatment interventions, perhaps due to the belief that no treatment is effective” (p. 102).

In a plenary address to the membership of the ATSA in 2005, James Breiling (a branch director of NIMH) strongly encouraged persons in the sexual offender field to confront the lack of empirical evidence regarding sexual offender treatment outcome research or risk losing further credibility with funding agencies and the general public. Dennis et al. (2012) were adamant that without consistent, methodologically sound research to demonstrate that sexual offender treatment is effective, there is a risk “that society [may be] lured into a false sense of security in the belief that once the individual has been treated, then their risk of reoffending is reduced. Currently, the evidence does not support this belief” (p. 28). ATSA too, in 2010, stated: “Community safety is better promoted by identifying treatments with strong evidence of effectiveness than by a proliferation of programs for which the efficacy is debatable” (ATSA, 2010a). Thus, currently, the question, “Does sexual offender treatment work?” cannot technically be answered, because the research base for psychosocial interventions for sexual offenders consists almost exclusively of methodologically limited studies. Moreover, the few available RCTs for such offenders have failed to show that sexual offender treatment decreases sexual offense recidivism rates or affects personal change related to such reoffending.

The present limitations of the available scientific study regarding the potential effectiveness of sexual offender treatment raise a number of practical questions. On the one hand, it might be considered a reasonable option to simply accept the lack of knowledge and the limitations of multiple aspects of individual change as a result of psychotherapy, at least for a select but large groups of such offenders and simply acknowledge that, to date, the efficacy of psychotherapy for sexual offenders has yet to be demonstrated; no robust scientific evidence yet exists that sexual offender treatment

“works.” This reasoned conclusion suggests that, at this time, the most appropriate response to the lack of empirical evidence of sexual offender treatment effectiveness should be a greater emphasis on other aspects of social management, perhaps a containment strategy (e.g., see English et al., this volume) that does not accord psychotherapy a primary or central role in the management of sexual offenders. From this perspective, a decision that the extant failure to demonstrate a robust treatment effect for sexual offenders means that other management approaches should be accorded a larger role and appropriately funded as the primary mechanisms to manage sexual offenders. As Harris et al. (1998) opined, “to the extent that treatment fails to reduce recidivism, supervision (including denial of community access) has to take its place” (p. 104). This perspective would emphasize a need for further scientific study of and continued development of mechanisms for differentiating among sexual offenders to provide for the most accurate appraisals of risk and need for sexual offense recidivism. In turn, as risk and other assessments became more refined, they might provide that the basis for more comprehensive alternative management strategies could be directed at those sexual offenders with the highest risk and the greatest needs, without the assumption that systematic means existed to motivate, promote insight and understanding, and accompany psychological change for sexual offenders via formal psychosocial interventions.

Another option would also seem quite reasonable, namely, that new methodologically adequate studies of sexual offender treatments should be developed, funded, and studied toward the end of potentially establishing more definitive and positive results that psychosocial interventions can be effective—either by substantially reducing the future risk of sexual offense recidivism or by a more delimited goal (e.g., clear evidence of individual changes and some relatively rigorous form of harm reduction that would still be socially acceptable). This second alternative would advocate that the most likely means toward scientific and public credibility regarding the potential utility of sexual offender treatment are through applied science, as is the case with other severe presenting problems that threaten the well-being of self and others. Harris et al. (1998) wrote, “to say that treatments have not thus far been conclusively evaluated is not to say that they do not work” (p. 104)—or cannot work for that matter. Langstrom et al. (2013) offered the same observation as well as suggested that professional opinion (e.g., clinicians’ judgments based on hope or wishes) was no substitute for evidence. However, Harris et al. also stated:

It behooves those who provide treatment and supervision, especially when directly or indirectly publicly funded, to reduce the existing uncertainty about the effects of these interventions by conducting scientifically useful evaluations of the services provided. We believe that such evaluations should be mandatory for publicly funded offender treatment. (p. 107)

Thus, one appropriate practical conclusion of this, as well as most other reviews, is that the challenge remains for the field of sex offender research to empirically demonstrate that sexual offender treatment works (particularly for higher-risk sexual offenders) through multiple and repeated RCTs as for other significant social and medical problems. As Hanson (1997) pointed out over 15 years ago, “Independent replication is a foundation of scientific knowledge. It is only through the accumulation of consistent results from diverse studies that skeptics either become convinced or lose their own credibility within the scientific community” (p. 133). Similarly, Miner (1997) wrote, “Science is a process of replication, since any study has flaws. Knowledge is thus advanced through a body of research that builds on what preceded it, correcting the flaws of previous studies, while raising additional questions... (p. 103). It is ultimate, this quantitative accumulation of research findings that will provide scientific evidence of sexual offender treatment effectiveness” (p. 108). More recently, Schlank (2010) also stressed the importance of replication: “Any professional field that is based on scientific research must stress the importance of replication of studies, which can provide either verification or disconfirmation” (pp. 22–23) and noted that in many fields, researchers (and practitioners) will not even consider a study complete until it has been replicated several times.

As pessimistic as the available data are regarding the status of current sexual offender treatment outcome, the path to a more empirically based understanding of such interventions seems fairly obvious. Within the field, there is an increasing consensus as to what should characterize future research efforts. In fact, over the past 30 years, most researchers have consistently spoken of the ways to enhance the understanding and potential credibility of sexual offender treatment. Furby et al. (1989) stated in their early review of sexual reoffending in both treated and untreated sex offenders: “...Progress in our knowledge about sex offense recidivism will continually elude us until adequate resources of time, money, and research expertise are devoted to this issue... It is time that we give this issue the resources and attention it deserves” (p. 28). In 2005, Marques et al. indicated:

Questions about whether and when sex offenders can be treated are extremely important, not just to our field but to victims, policy makers and the public. The only way to provide answers with confidence is to build a knowledge base on thoughtful and well-controlled studies of treatment effectiveness. (p. 104)

Craig et al. (2003) concluded: “Treatment studies should adopt well matched and randomized controls using appropriate and universal measures of recidivism” (p. 86). Seto et al. (2008):

Only methodologically rigorous research will allow us to determine which is which...[we] want to identify and disseminate treatments that can effectively reduce the likelihood that sex



offenders will do further harm, but we believe that only good science can inform good clinical practice and lead to the advancement of sex offender treatment. (p. 253)

Hanson (2014) has stated: "...we know very little about the effectiveness of methods used to rehabilitate sexual offenders...it is hard to make any strong conclusions about whether treatment works at all...This is a depressingly similar conclusion to that of ...more than 20 years earlier. Knowing which treatment works for which type of sexual offender remains a distant dream" (p. 5). He states, "Although we, as service providers must believe in what we do in order to do it, we also need the humility to admit that we could be fundamentally mistaken. Consequently sexual offender treatment needs rigorous scientific scrutiny..." (p. 6). Referencing the large number of studies of risk assessment of sexual offenders and the hundreds of treatment outcome for general criminal offenders, Hanson (2014) stated: "What we need now are hundreds of new studies of sexual offender treatment outcome..." (p. 7). He advocated for the significance of evidence-based practice, noting that the growing interest in evidence-based practice in the large mental health field should be viewed as a sign of progress and for those who want science to influence sexual offender practice and should be viewed as a genuine force for the good. In short, further scientific study of sexual offender treatment is a necessary step for the field to advocate that such interventions should be an essential component of sexual offender management. The onus rests solely on the field of sexual offender research and management to establish that sexual offender treatment is clearly effective, particularly for the most high-risk sexual offenders.

Almost all credible researchers agree that the initial step to determining if sexual offender treatment can be effective in reducing sex offense recidivism is to systematically develop a body of methodologically sound RCTs of such interventions, preferably involving a relatively large number of subjects. Langstrom et al. (2013) concluded, "Based on the meagre results from our extensive systematic review, we concluded that there is an urgent need for well designed and well executed trials of treatment for adults who commit sexual offences against children" (p. 4). In their 2912 Cochrane Review, per Dennis et al. (2012), "We concluded that further randomised controlled trials are urgently needed in this area..." (p. 28). Rice and Harris (2003) stated, "...it is abundantly clear that any conclusions about the effectiveness of psychological therapy await many more random assignment studies" (p. 427). Similarly, Seto et al. (2008) declared, "In our view, RCTs are both ethical and necessary in order to prevent more victims of sexual violence and abuse" (p. 254). In addition, Hanson et al. (2009) wrote that:

...strong studies are needed...we believe that an important requirement of strong research design is the experimenter's ability to determine participant assignment based on a procedure

that controls for both measured and unmeasured features of the offenders (i.e. Random assignment)...Random assignment studies remain the best available alternative for minimizing participant election bias. Random selection is also one of the most ethically defensible methods of assigning individuals to treatment when demand exceeds supply or the relative superiority of alternate treatment is unknown. (p. 887)

Shortly after the publication of his 2009 review, Hanson indicated: "I, for one, have done enough meta-analyses of barely acceptable studies. It is time to counter the political resistance to random assignment studies by getting ATSA to endorse a position statement supporting their use" (Hanson, cited in Rice & Harris, 2013). In yet another recent review of sexual offender treatment, Kaplan and Krueger (2012) wrote: "...large, well-conducted randomized trials of long duration are essential if the effectiveness or otherwise of these treatments is to be established. Most of the studies upon which the knowledge base of the treatment of sexual offenders is based are seriously flawed. Overall, however, the evidence base for cognitive-behavioral treatment is extremely limited and empirical research focusing on effective treatment for this population is critically needed" (p. 295). Thus, investigations of the efficacy of sexual offender treatment must begin with RCTs, with largely similar offenders randomly assigned to one or more interventions and control groups of similarly motivated persons. Such studies must involve repeated, multi-method measures of likely psychologically meaningful risk factors. Offender subjects must be followed via survival analysis (with attention to attrition, reincarceration, and other types of reoffending) with any additional relevant experiences (e.g., additional treatment, correctional supervision) that must be accounted for as well. As part of its commitment to promoting evidence-based practices and high-quality research, ATSA (2010) has stated: "ATSA recognizes randomized clinical trials (RCT's) as the preferred method of controlling for bias in treatment outcome evaluations. ATSA promotes the use of RCT to distinguish between interventions that decrease recidivism risk of sexual offenders and those programs that have no effect or are actually harmful." There should be little doubt that RCTs are feasible for sexual offenders. Rice and Harris (2012) noted that there were 267 existing RCTs in the field of criminal justice in 1993 and 87 RCTs in correctional research alone as of 2005. They also noted that there have been several RCTs for adolescent sexual offenders (and of note, studies which have demonstrated the efficacy of one particular model of intervention). Consequently, Rice and Harris (2003) concluded: "It is abundantly clear that RCTs are feasible both ethically and practically in crime and justice fields in general, and in corrections, specifically" (p. 18). A number of such studies have been conducted such as those by Cullen et al. (2011) and Davidson et al. (2009). Further, Hanson et al. (2008) point out that other improvements to research study quality

could be implemented at relatively low cost, including reporting intent-to-treat analyses, using equal and fixed follow-up periods, scoring actuarial risk measures on the treatment and comparison groups, using statistical controls, and matching on risk-relevant variables. They also pointed out that “much less is known about the processes by which sexual offenders change. Studies frequently find that improvements on factors presumed to be criminogenic have no effect on sexual recidivism rates...” (p. 887). Such a perspective echoes that of Borkovec and Castonguay (1998) who wrote: “Creating increasingly effective therapies through between-group designs is best done by controlled trials specifically aimed at basic questions about the nature of psychological problems and the nature of therapeutic change mechanisms. Naturalistic research is important for external validity but is valuable only if it uses scientifically valid methods to address basic knowledge questions” (p. 1). Thus, sexual offender treatment approaches should be rooted in evidence-supported theories of sexual offending and initially determine if sexual offender treatment can be demonstrated through RCTs and several theory-supported models should be evaluated. Such methodologically rigorous studies are only the beginning by establishing internal validity. Subsequently, the heterogeneity of sexual offenders in relation to sexual offender treatment outcome as well as sets of components and parameters of sexual offender treatment can be rigorously evaluated. As with other disorders, methodologically sound investigations should target those criminogenic needs (or in the language of the larger treatment field, risk factors) in the context of evidence-based elements of effective psychosocial interventions. In addition to outcome studies, Hanson et al. suggested that researchers also focus on short- and medium-term changes on intermediate treatment targets and criminogenic needs; they noted that outcome research should help advance knowledge of the change process by examining the relationship between changes on more proximal treatment targets and more distal sex offense recidivism. Following this point, it becomes critical for further investigation into the effectiveness of different targets of sexual offender treatment and comparisons of different methods and approaches for affecting those targets. As with other recurrent, multidimensional problems targeted for change, multiple high-quality progression of methodologically refined studies of sexual offender treatment outcome will be necessary over time to best understand first whether treatment “works,” and if so, what aspects of the therapist, client, and treatment program components contribute to effectiveness.

Despite the discouraging findings from scientific evaluations of sexual offender treatment programs at present, updated and innovative perspectives on treatment and on potentially effective approaches to the treatment of sexual offenders have continued to develop, many of which are reviewed by Yates (2015, this book). Such developments

include the Self-Regulation (SR) Model (Ward & Hudson, 1998), Good Lives Model (GLM; e.g., Ward & Stewart, 2003), the integrated Good Lives/Self-Regulation Model (Yates & Ward, 2008), a “Strength-Based” Model of psychotherapy (e.g., Marshall et al., 2011), and the Recidivism Risk Reduction Treatment approach (3RT; Wheeler & Covell, 2013). Such approaches uniformly suggest that sexual offender treatment will be most successful when it is comprehensive and incorporates the management of predisposing, risk-related characteristics as well as encouraging the development of positive personal goals and healthy lifestyles. Particularly, given their “positive” approaches to the nature of offenders as individuals and to the goals and methods of therapeutic work, these developments appear to be heartening and inspiring to forensic/clinical practitioners. However, these newer perspectives on sexual offender treatment have and should be met with some significant degree of skepticism by others; 15 years ago, Quinsey et al. (1998) remarked “Overall, it seems clear that the field of sex offender treatment is changing without progressing” (p. 150). The promise of novel or presumed innovative approaches to sexual offender treatment must be put to the test of empirical investigations prior to unquestioned excitement and wide adoption. Both Quinsey et al. (1998) and Hanson (2003) have noted that these and previous novel sexual offender treatment models have been sequentially proposed (and others recommended for rejection) on exclusively or predominantly *nonempirical* or *theoretical bases*, in contrast to models of interventions for this group advancing progressively on the basis of scientifically sound appraisals. Most recently, Hanson (2014), in commenting on the lack of a scientific foundation for sexual offender, stated “The development of the [sexual offender treatment] field cannot be attributed to strong empirical evidence that such treatment is effective ...*the changes in our treatment practices during my professional career have had only the lowest inspiration from research findings...It is hard to argue that we switched from aversive conditioning to relapse prevention (RP) and from RP to Good Lives because of any deep commitment to evidence-based practice*” (p. 3, emphasis added). Thus, to date, there are no RCTs of the Self-Regulation Model (SRM; Ward & Hudson, 1998), the Good Lives Model (GLM; Ward & Gannon, 2006; Ward & Stewart, 2003), the combined SR/GLM (Ward & Gannon, 2006; Yates, Prescott, & Ward, 2010; Yates & Ward, 2008), and Marshall et al.’s Strength-Based Approach model (SBA; Marshall et al., 2011) or Recidivism Reduction Therapy (3RT) (Wheeler & Covell, 2013); similarly, there is incomplete evidence that each of the principles of RNR treatment for general criminal offenders applies to specifically sexual criminal offenders, particularly the risk principle. However, as noted earlier, an increasing number of sexual offender treatment programs in North America and the UK employ aspects of GLM, a

change in practice that is not based on scientific evidence. In part, this reflects a larger issue, a professional reliance on “unstructured clinical judgment,” where clinicians overwhelmingly rely on their own beliefs and experiences in their clinical practice as opposed to research findings (e.g., Ogilvie, Abreu, & Safran, 2005; Stewart & Chambless, 2007). This is despite the caution of Langstrom et al. that “Professional *opinion* is no substitute for evidence” (p. 4, emphasis added). Consequently, there continues to be an increasing divergence between what has been (or has not been) scientifically demonstrated regarding sexual offender treatment and what forensic/clinical practitioners actually do with clients who have sexual offended. While intuitively appealing, newer treatment models and interventions (such as the GLM, SR/GLM, SBA, 3-RT, and other positive approaches, motivational interviewing, and treatment preparation) need to be subject to empirical testing to determine their relationship to sexual offender treatment outcome and individual change. Psychosocial interventions should not be abandoned at this point. Rather, as with other complex psychologically based presenting problems, the lack of scientifically demonstrated treatment outcome results should be the urgent impetus for increased study of psychosocial interventions for sexual offenders, informed by existing data and the principles of effective psychotherapy.

In the future, outcome studies of sexual offender treatment must be derived from more evidence-based principles and practices, including the value of careful identification and comprehensive, systematic evaluation of clinical expertise and patient characteristics, preferences, and circumstances (e.g., Spring, 2007). Primarily, treatment outcome studies of sexual offender psychotherapies must rely on RCTs, with similar offenders randomly assigned to one or more interventions and control groups of similarly characterized and motivated sexual offenders. Experimental interventions must be bona fide interventions, and given the current status of the field, it makes sense to study multiple psychosocial approaches in treating sexual offenders. RCT studies of treatments for sexual offenders should involve repeated, multi-method measures of likely psychologically meaningful risk factors. Subjects must be followed via survival analysis, with minimal attrition, overtime, and any additional relevant experiences (e.g., additional treatment, correctional supervision) that must be accounted for as well. Potential moderators and mediators of clinical change need to be further identified, monitored, and refined. Sexual offenders of differing levels of risk, particularly high-risk offenders, need to be the focus of outcome research. As well, it seems timely for the sexual offender field to reconsider what the essential focus/content of treatment approaches should consist of. The criminological literature has suggested that a focus on the so-called criminological needs as the most meaningful treatment targets; this parallels the broader psychotherapy field’s

increasing focus on personality issues or predisposing condition diatheses which seem to underlie the presenting problems of more psychotherapy-refractory clients; while these are similar, the former relies exclusively on the results of science, while the latter reflects both empirical research and clinical theory. Hanson et al. (2009) pointed out: “Studies frequently find that improvements on factors presumed to be criminogenic have no effect on sexual recidivism rates...” (p. 886). As a result, they advocated for changes in substance of sexual offender treatment programs, stating:

...it would be beneficial for treatment providers to carefully review their programs to ensure that the treatment targets emphasized are those empirically linked to sexual offense recidivism. Examples of promising criminogenic needs include sexual deviancy, sexual preoccupation, low self-control, grievance thinking and lack of meaningful intimate relationships with adults...Outstanding questions remain, however, concerning potential gains from matching interventions to the needs of individual offenders and whether recidivism can be most effectively reduced by addressing certain combinations of needs. (p. 886)

As noted, both models of other presenting problems/mental disorders as well as sexual offending (e.g., the ITSO; Ward & Beech, 2006) increasingly highlight the potential significance of implicit psychological experiences and intra-individual content and process issues. Potential changes in such needs, implicit theories and issues, and other theory-based treatment-related factors must be tested for reliable, clinically significant, and valid change.

In addition to the focus of sexual offender treatment, treatment delivery issues also need to be investigated in a controlled systematic manner. Per the RNR model of correctional intervention, the importance of the relative intensity and duration of sexual offender treatment as well as responsivity dimensions for offenders with different levels of needs and risk must be carefully examined. Truly effective methods of changing offenders’ interconnected thoughts, feelings/motivations, and behaviors need to be identified; the relative value of psychoeducational and experiential treatment tactics needs to be identified and refined; dismantling and recreating truly effective methods and strategies of sexual offender treatment should occur. While intuitively appealing, the so-called positive treatment models and interventions (such as the Good Lives approach, Strengths Approach, 3RT, motivational interviewing, and treatment preparation) must be subject to rigorous empirical testing to determine their relationship to the central outcomes in the treatment of sexual offenders: personal change and decreased sexual offense recidivism rates. Therapist, client, and process variables need to be carefully studied; particularly, for high-risk/high-need sexual offenders, it seems likely that they would benefit from particularly well-trained and experienced therapists and that this would likely be a cost-effective practice. In this vein, Marshall’s (e.g., 2005) work regarding

the significance of clinician qualities in impacting sexual offenders seems increasingly important. However, first, the delineation and cultivation of effective psychotherapist qualities and knowledge specific to working with sexual offenders seem essential. Second, multiple investigators must systematically and empirically examine the empirical value of therapist qualities and knowledge. To the extent that the evidence that the therapeutic relationship or alliance is believed to be critical to treatment success, then controlled outcome studies involving enhanced therapist characteristics or therapist–client matching should certainly be conducted. At the same time, issues of treatment fidelity and the value of parameters of clinical supervision remain to be examined in relationship to the effectiveness of sexual offender treatment approaches. The relative contribution of individual and group psychotherapy separately and in combination (and the types of groups such as closed or “rolling”) needs to be evaluated particularly in relationship to specific collections of presenting problems and other client dimension; as Hanson et al. suggested, it may be time to more explicitly match treatment approaches to the particular needs of specific offender clients. Ultimately, for the problem of sexual offending, as it would be any type of presenting problem, this is what Paul (1967) wrote over 40 years ago: “...the question towards which all outcome research should ultimately be directed is the following: *What* treatment, by *whom*, is most effective for *this* individual with *that* specific problem and under *which* set of circumstances?” (p. 111). However, the answers to those questions in regard to sexual offender treatment outcome will only be determined by the development of research programs of sexual offender treatment with diverse samples of sexual offenders that are controlled, comprehensive, and able to be replicated across settings.

A key issue for mental health professionals involved in the psychosocial treatment and management of sexual offenders is to come to terms with the role of scientific investigation and results regarding such treatment and the larger field of psychotherapy and mental health interventions. As noted earlier, there is tremendous public skepticism of the mental health field, particularly psychology; as Stanovich (2009) observed, “Most judgments about the field and its accomplishments are resoundingly negative” (p. 175). Lilienfeld (2011) points out that we ignore such skepticism at our peril, in terms of potential client’s expectancies about possible improvement, third-party reimbursement, and government funding of significant research questions. Gaudiano and Miller (2013) point out that “Many psychotherapists are opposed to the idea of the specification of evidence-based treatments in principle, viewing psychotherapy at least as much art as science and preferring to rely on clinical intuition and experience instead of scientific evidence...” (p. 815). They note that most clinicians do not base their

treatment decisions on “state-of-the-art” clinical research and that approximately 50 % reject the use of more formal, evidence-based treatment approaches and rely primarily on their own subjective clinical experiences. Lilienfeld et al. (2013) provides a wide-ranging analysis of the reasons why mental health professionals have been resistant to evidence-based practice and remedies to those issues. In effect, all of these apply to MHPs practicing in the sexual offender treatment field. As Dennis et al. argued, “...this weaker evidence (and the conclusions drawn from it) often leads to a more optimistic conclusion about efficacy than is warranted, and unfortunately becomes embedded in clinicians’ consciousness. This may result in a belief that current approaches are more effective than the evidence suggests” (p. 28). They even noted that previous conclusions of earlier reviews of the limitations relative to the demonstrated effectiveness of sexual offender treatment and the repeated call for further research have typically been cushioned by misleading phrases such as the results are nonetheless “promising. More recently, Duggan and Dennis (2014) noted that there are over 2,900 RCTs of psychosocial interventions for Schizophrenia. Regarding the place of evidence in the treatment of sex offenders, they concluded”:

Although RCTs in any area of healthcare are difficult to conduct, other specialties have overcome the challenges that they present...It is clear that high quality evidence can be produced in most areas of healthcare, if there is the will to do so. For this to happen with respect to treatment for sex offenders, spurious impediments...must be set aside. Those who enter sex offending programmes, together with their past and potential future victims, should expect to be provided with treatments with a strong evidence base. Acquisition of this evidence must be a process, which includes, although is not confined to, RCTs. (p. 160)

Relative to this issue, it is striking that in the sexual offender field, mental health professionals readily accepted and utilized various structured risk instruments based on the finding that experimentally derived statistical information outperforms [pure] clinical judgment. Yet, in marked contrast, given a more striking lack of empirical data and justification for any psychosocial sexual offender treatment, mental health professionals in the sexual offender field have consistently defended their belief or faith in psychotherapy as a viable component of management to reduce sexual offending.

As an antidote to these issues, Lilienfeld (2011) argues that the mental health fields must “police themselves”; he specifically states that while thoughtful debates about the best means of operationalizing evidence-based practice should continue, “practioners with the applied fields of psychology (e.g. clinical, counseling, school) would be well advised to become less tolerant of pseudoscience and more willing to ground their practices in replicated research evidence” (p. 14). Actually, Andrews and Bonta (2006) offered a similar perspective regarding the study of criminal

behavior: “Unsparring criticism is a major source of advancement...all criticism, including criticism of theoretical and research-based assertions, is best combined with respect for evidence...” (p. 3). Placed in the larger national and international context of health economics and management, Baker et al. (2009) stated: “The current context of health care in American (and beyond) demands a higher level of accountability than in the past...the future of clinical psychology will be dictated largely by what data show regarding the relative cost-effectiveness of psychosocial and behavioral interventions compared with other competing intervention options in mental health care...Clinical psychologists must offer compelling evidence relating [to the criteria of such comparisons] if they expect their psychosocial and behavioral interventions to have a fair chance of gaining widespread support, to be adopted in the health delivery system, and to be funded via health coverage mechanisms...” (p. 69). Baker et al. reviewed the history of the progress of medical care in the USA and offered a convincing argument that the increased, nearly universal acceptance of medical treatment is based on three sociopolitical changes: (1) the scientific grounding of medical practice in experimental study, primarily RCTs; (2) a greatly expanded body of science accompanied by increasingly rigorous training of physicians in evidence-based procedures and standards of practice; and (3) higher standards in training and licensure. Baker et al. note that physicians have almost exclusively positive views regarding experimental evidence such that it constitutes a touchstone regarding practice, and as a result, practice studies show that a very high percentage of medical patients receive interventions that are evidence based. In contrast, they demonstrate that psychologists and other non-medical mental health providers view science and research as having very little relevance to their practice activities and decisions. Moreover, they note that “Clinical psychologists often practice in a manner that *conflicts* with considerable research evidence or at least is not clearly supported by research evidence...practitioners often say they do not care, because they consider the available scientific evidence to be relatively uninformative or irrelevant to their practice decisions...” (p. 80). They argued that unless significant changes occur in the mental health profession’s acceptance of a scientific approach to the treatment of mental health problems, MHPs risk being even more devalued and even further reduce in their roles in both the practice and policymaking about the utility of psychosocial interventions.

Thus, to continue to argue—and more importantly to *act*—as if psychotherapy has been empirically demonstrated to be effective at reducing future sexual offending is ultimately to risk the exclusion of such intervention modalities or treatment practitioners as one element in a broad approach to managing sexual offenders. Until strong empirical evidence exists that sexual offender treatment does significantly

and differentially reduce sexual offense recidivism, several issues remain for the various participants and stakeholders. This is quite similar to the related field of psychosocial interventions for persons with ASPDa. Duggan (2008), an author of several Cochrane and related reviews of this Personality Disorder, concluded: “The implication is clear: that there is an imperative for scientists and clinicians to provide decision makers with the appropriate evidence to allow the latter to arrive at the best decision...we are in a weak position to influence the political process in the allocation of funds so that unless and until these areas are addressed, interventions for [criminal offenders] with ASPD are likely to continue to remain in a scientific limbo” (p. 2610). Several writers have suggested that the sexual offender field effectively becomes more accurate and honest in representing what sexual offender treatment might offer some sexual offenders. Another perspective to take regarding the lack of evidence of sexual offender treatment is to simply consider if one would recommend to others or choose for oneself a medical intervention that lacked one, let alone replicated, empirically demonstrated trials of its relative effectiveness. Over 15 years ago, advocating for a harm reduction approach to psychosocial interventions for sexual offenders, given the “not particularly optimistic” evidence for treatment success for sexual offenders, Laws (1996) stated:

The domain of treatment provision is an imperfect one and we should openly acknowledge that...I believe we should stop using the words sexual offender *treatment* to characterize our work and substitute sex offender *management* instead, since it is actually more accurate...*Treatment* suggests sexual deviance may remit or be cured and so, like a treatment for a disease, establish expectations for success which are quite unrealistic... At bottom, our job in managing sexual offenders and reducing harm is, in reality, a sort of social policing.... (p. 246)

However, as Harris et al. (1998) wrote: “It behooves those who provide treatment and supervision, especially when directly or indirectly publicly funded, to reduce the existing uncertainty about the effect of these interventions by conducting scientifically and useful evaluations of the services provided. We believe that such evaluations should be mandatory for publicly funded offender treatment” (p. 107).

Recently, in the ATSA Forum, Pake (2010) opined that the state of science relating to the management of human behavior is not yet at a point when one can proclaim treatment success with any certainty, saying “It is currently impossible to support such a statement scientifically.” Further, he notes that whether a sexual offender who has participated in treatment chooses to utilize understanding and learned skills necessarily remains at the discretion of the offender and no therapist can account for a particular offender’s choice in any given circumstance. Pake concluded by stating: “By portraying treatment as successful, we offer a false sense of security. Portraying treatment as successful encourages non-clinical

partners in community risk management to perceive our efforts as having eradicated the potential for reoffending on the part of the treated sexual abuser. This is misleading. It leads one to question the profession's intellectual honesty."

Clinician's practicing non-forensic psychotherapy with clients who are independently choosing to engage in and pay for such treatment—for whom the only stakeholders are the client and the therapist—should be relatively free to engage in whatever procedures they mutually believe are in the client's best interest. However, the process of forensic clinicians providing forensic psychotherapy to anticipated or actual forensic clients involves other considerations, particularly the interests and support of other stakeholders—in most cases, the treatment is being funded by third parties (e.g., an agency acting on behalf of society) and the purpose of that psychotherapy is primarily public safety (which is the basis for the agency and/or public funding the treatment). For forensic/clinical psychotherapists in the community and institutions, what does one do in sexual offender treatment? Certainly, there are some sexual offenders who do and will benefit from psychotherapy such as those that are low-risk and presumably one-incident offenders. Psychosocial treatments of some type are likely effective for some specific offenders, but, at present, group data does not support this conclusion. Consequently, the degree to which a particular sexual offender will or does benefit from sexual offender-specific treatment, with or without additional psychosocial interventions, must be carefully considered. [Similarly, whether a particular sexual offender has benefited from psychotherapy cannot meaningfully rely on group outcome data, therapist ratings, or self-reported change but rather must be determined on some individual basis, with its own extensive set of "measurement" issues.] Given the apparent failure to demonstrate effectiveness for general CBT-RP-type sexual offender treatments—and by implication the component "modules" (e.g., Hoberman, 2015)—it appears critical for those who intend to or must provide psychosocial interventions to sexual offenders that they critically examine the components and implementation of their treatment. Clearly, for presenting problems like eating disorders and drug abuse, empirical evidence for the qualified efficacy of existing treatments exists, and both help-seeking and resistant clients are offered psychosocial interventions. Such interventions are necessarily based on demonstrated or hypothesized harm reduction for the individual client and not necessarily as a "cure" (albeit the risk or harm associated with eating disorders and most other presenting problems is largely to the client and not others). However, the "harm" dimension of those disorders relates to the client and not to others/society; harm reduction may be a useful concept for disorders that pose issues of self-harm. However, harm reduction may not be a sufficient outcome for sexual offender treatments as with other violent offenders. It is reasonable to ask that psycho-

logical treatments of persons with a demonstrated history and propensity for violent sexual offending against others be demonstrated to be substantively effective if they are to be accorded a primary place in the management of such offenders and/or funded by the public.

To the extent a sexual offender is motivated or can be genuinely influenced to engage in treatment (whether it be intrinsically or extrinsically), several practices seem reasonable. To begin with, there is an ethical and practical issue as to what type and degree of expectancy can and should be communicated to offenders who express interest in or are mandated for sexual offender treatment; an emphasis on collaboration; the relevance of the offender's motivation to be open, to learn, and to enact life changes; and an agreement by the psychotherapist to work empathically, respectfully, and collaboratively with the offender should provide an appropriate framework for potentially effective treatment. However, these practices need to be guided by scientifically informed data and then the clinical needs and responsivity issues of particular offender clients. In the absence of scientifically informed forensic/clinical practice, several questions exist for practicing clinicians who provide psychotherapy for sexual offenders. Harris et al. (1998) advocated as follows:

The best option in these circumstances of relative ignorance is to adopt treatments that (a) fit with what is known about the treatment of offenders in general, (b) have a convincing theoretical rationale in that they are motivated by what we know about the characteristics of sex offenders, (c) have been demonstrated to produce proximal changes in theoretically relevant measures, (d) are feasible in terms of acceptability to offenders and clinicians, cost, and ethical standards, (e) are described in sufficient detail that program integrity can be measured, and (f) can be integrated into existing institutional regimens and supervisory procedures. (p. 104)

Similarly, Langstrom et al. (2013) wrote "Without specific guidelines for treating individuals at risk, the most ethically defensible position would be to assess the presence of treatable risk factors for sexual abuse of children, including concurrent psychiatric disorder, and offer individualised treatment" (p. 4). As with other presenting problems that lack of demonstrated effective interventions, offering psychotherapy should continue to be offered to offenders who appear genuinely and intrinsically motivated for such interventions. However, as forensic therapy, with the community as a significant "client" or "interested party," honest and accurate representations about the existing empirical evidence for such psychotherapies must be acknowledged; related concerns exist about who should bear the cost of unproven interventions. A related practical and ethical question concerns what practitioners can and do communicate to sexual offenders about the possible benefits of sexual offender treatment. Should offenders be provided with an accurate "likely no effect" or an "optimized" perspective on the likely effectiveness of such psychosocial interventions relative to their

expectancies of potential change? Is it ethical to induce a heightened positive expectancy for sexual offender treatment via motivational interviewing or other preparation in light of both the failure to demonstrate treatment effectiveness and the lack of empirical evidence that such theoretical notions themselves actually affect sexual offender treatment outcome? As with other presenting problems lacking demonstrated effective interventions, offering psychotherapy should continue to be offered to offenders who appear genuinely and intrinsically motivated for such interventions and for whom resources are available to fund their treatment. However, as forensic therapy, with the community as a significant “client” or “interested party,” honest and accurate representations about the existing empirical evidence for such interventions must be provided; the failure to demonstrate efficacy of psychotherapies for enacting personal change in sexual offenders and decreases in sexual offense recidivism must be acknowledged so that policy makers and the public are informed about the potential value of resource allocation and the degree of community safety such resources might provide.

From a public policy perspective, the lack of an empirical demonstration of the efficacy of sexual offender treatment, particularly as forensic psychotherapy, raises several significant questions. Should treatment be mandated in the absence of clear demonstrations that sexual offender treatment “works?” Regarding the study of criminal behavior, Andrews and Bonta (2006) wrote: “...it views a reduction of the costs of both crime and criminal justice processing as highly desirable. We are particularly interested in reducing the costs of crime by reducing criminal victimization in the first place” (p. 3). To the extent that much of sexual offender treatment is provided by way of public funding for institutionalized offenders or social service benefits (and to a lesser degree by insurance funding), should demonstrated effectiveness be necessary to justify such funding? In the absence of objective evidence, should offenders themselves bear the costs of funding sexual offender treatment, given providers belief that such intervention is hypothesized to impact their lives in a positive manner? In addition, if, at best, such interventions can only offer some small degree of “harm reduction” in reducing the frequency and severity of sexual offending, is that a sufficient goal for public safety and for justification of the use of public funds to provide such interventions? To what degree does society have a right to demand evidence of a large effect for sexual offender treatment—a high degree of empirically demonstrated persisting change on the part of sexual offenders? Another perspective would suggest that if sexual offender treatment cannot currently be strongly relied on to clearly and consistently reduce sexual offense recidivism, then the management of sexual offenders should shift to other alternative practices. Over a decade ago, Harris et al. (1998) noted that, to the extent that treatment fails to reduce recidivism (the current state of science), “...supervision

(including denial of community access) has to take its place” (p. 104). In addition to more intensive and long-lasting community management, other options to a reliance on psychotherapy as a primary management tool for sexual offenders might include utilizing insurance programs for sexual offenders to obtain coverage for liability relative to their risk for future sexual offending (e.g., similar to motor vehicle or malpractice insurance) or extend or indeterminate sentences for offenders with prior sexual offending history and so on.

All can agree that for all stakeholders, the prioritization for the prevention of sexual violence requires, even demands, increased time and resources be devoted to studying and innovating programs for reducing future acts of sexual offending by identified sexual offenders. Given the degree of public concern about sexual offending expressed by society and political entities, there should be no question that substantially increased funding of psychotherapy outcome studies for sexual offenders should occur, just as such expanded funding has increased for other identified public health problems which effect far fewer members of the community. The *potential* for psychosocial interventions to play a central role in facilitating understanding and change in sexual offenders clearly exists. However, only by *accepting* the reality of the current status of the field of sexual offender treatment can the scientific and larger public community *commit* to a reasonable process prioritizing and funding theorizing, testing, and refining treatment models, strategies, and tactics that might be shown to effectively assist sexual offenders in modifying their personal characteristics and social contexts in such ways that their risk for future sexual offending is eliminated or substantially reduced. Following the principles of EBP, it is critical that researchers and policy makers collaborate to develop, test, and assign resources to sexual offender treatment and that researchers and sexual offender treatment programs work collectively to execute standardized research studies that clarify the role that such interventions can play in the management of sexual offenders. Sexual offender treatment clinicians and program managers must be fully informed on the existing and evolving scientific research regarding the outcome and implementation of sexual offender treatment and be educated, committed to, and supervised in implementing best practices in clinical work with sexual offender clients. At the same, as with other presenting problems lacking demonstrated effective interventions, it is reasonable to continue to provide psychotherapy to offenders who appear genuinely and intrinsically motivated for such interventions; however, without demonstrated efficacy, funding responsibility may and perhaps should shift to sexual offenders themselves. For sexual offender treatment as forensic psychotherapy, with the community as a significant “client” or “interested party,” honest and accurate representations about the existing empirical evidence for such interventions must be provided. The failure, to date, to demonstrate efficacy of

psychotherapies for enacting personal change in sexual offenders and decreases in sexual offense recidivism must be acknowledged so that policy makers and the large community are informed about the potential value of resource allocation. Finally, given its role as almost exclusively forensic psychotherapy, advocates of sexual offender treatment must be transparent about what is known about its efficacy so that realistic notions of its role in public safety (as well as personal change) can be taken into consideration relative to its role in the management of sexual offenders.

## References

- Abel, G. G., Becker, J. V., Cunningham-Rathner, J., Mittelman, M., & Rouleau, J. L. (1988). Multiple paraphilic diagnoses among sex offenders. *Bulletin of the American Academy of Psychiatry and the Law, 16*, 153–168.
- Abel, G. G., Blanchard, E. B., & Becker, J. V. (1978). An integrated treatment program for rapists. In R. T. Tada (Ed.), *Clinical aspects of the rapist*. New York: Grune & Stratton.
- Abrcen, J., & Looman, J. (2004). Issues in the treatment of sexual offenders: Recent developments and directions for future research. *Aggression and Violent Behavior, 9*, 229–246.
- Ahlmeyer, S., Heil, P., McKee, B., & English, K. (2000). The impact of polygraph on admissions of victims and offenses in adult sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 12*, 123–138.
- Alexander, M. A. (1999). Sexual offender treatment efficacy revisited. *Sexual Abuse: A Journal of Research and Treatment, 11*(2), 101–116.
- Arkowitz, H., & Lilienfeld, S. (2006). Psychotherapy on trial. *Scientific American Mind*, (April/May), 42–49.
- Association for the Treatment of Sexual Abusers (ATSA). (2008). *Sex offender treatment for adult males: Public policy position paper*. Retrieved January 5, 2013, from the ATSA website.
- Association for the Treatment of Sexual Abusers (ATSA). (2010a). *Statement supporting the use of randomized control trials for the evaluation of sexual offender treatment*. Adopted by the ATSA Executive Board of Directors on August 17, 2010. Retrieved from the ATSA website, 5/1/13.
- Association for the Treatment of Sexual Abusers (ATSA). (2010). *Recommendations for implementation of randomized clinical trial design*. Adopted by the ATSA Executive Board of Directors on August 17, 2010. Retrieved from the ATSA website, 5/1/13.
- Baker, T. B., McFall, R. M., & Shoham, V. (2009). Current status and future prospects of clinical psychology. *Psychological Science in the Public Interest, 9*, 67–103.
- Barnett, G. D., Wakeling, H. C., Mandeville-Norden, R., & Rakestraw, J. (2012). How useful are psychometric scores in predicting recidivism for treated sex offenders? *International Journal of Offender Therapy and Comparative Criminology, 56*, 420–446.
- Barrett, M., Wilson, R. J., & Long, C. (2003). Measuring motivation to change in sexual offenders from institutional intake to community treatment. *Sexual Abuse: A Journal of Research and Treatment, 15*, 269–283.
- Beggs, S. (2010). Within-treatment outcome among sexual offenders: A review. *Aggression and Violent Behavior, 15*(2010), 369–379.
- Beggs, S. M., & Grace, R. C. (2010). Assessment of dynamic risk factors: An independent validation study of the violence risk scale: Sexual offender version. *Sexual Abuse: A Journal of Research and Treatment, 22*, 234–251.
- Beggs, S. M., & Grace, R. C. (2011). Treatment gain for sexual offenders against children predicts reduced recidivism: A comparison validity study. *Journal of Consulting and Clinical Psychology, 79*, 182–192.
- Bell, E. C., Marcus, D. K., & Goodlad, J. K. (2013). Are the parts as good as the whole? A meta-analysis of component treatment studies. *Journal of Consulting and Clinical Psychology, 81*, 722–736.
- Bergin, A. E., & Garfield, S. L. (Eds.). (1971). *Handbook of psychotherapy and behavior change* (1st ed.). New York: Wiley.
- Bergin, A. E., & Garfield, S. L. (Eds.). (1978). *Handbook of psychotherapy and behavior change* (2nd ed.). New York: Wiley.
- Bergin, A. E., & Garfield, S. L. (Eds.). (1986). *Handbook of psychotherapy and behavior change* (3rd ed.). New York: Wiley.
- Bergin, A. E., & Garfield, S. L. (Eds.). (1994). *Handbook of psychotherapy and behavior change* (4th ed.). New York: Wiley.
- Berliner, L. (2002). Commentary. *Sexual Abuse: A Journal of Research and Treatment, 14*, 195–197.
- Beuko, M., & Wong, S. (2005). Predictors of treatment attrition as indicators for program improvement not offender shortcomings: A study of sex offender treatment attrition. *Sexual Abuse: A Journal of Research and Treatment, 17*(4), 375–389.
- Bilby, C., Brooks-Gordon, B., & Wells, H. (2006). A systematic review of psychological interventions for sexual offenders: II quasi-experimental and qualitative data. *Journal of Forensic Psychiatry and Psychology, 17*, 467–484.
- Binks, C. A., Fenton, M., McCarthy, L., Lee, T., Adams, C. E., & Duggan, C. (2006). Psychological therapies for people with borderline personality disorder. *Cochrane Database Systematic Review*, 25, CD005652.
- Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. J. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative, 34*, 105–113.
- Borkovec, T. D., & Castonguay, L. G. (1998). What is the scientific meaning of empirically supported therapy? *Journal of Consulting and Clinical Psychology, 66*, 136–142.
- British Psychological Society & Royal College of Psychiatrists. (2009a). *Borderline personality disorder treatment and management*. London: National Institute for Health & Clinical Excellence (NICE).
- British Psychological Society & Royal College of Psychiatrists. (2009b). *Antisocial personality disorder treatment, management, and prevention*. London: National Institute for Health & Clinical Excellence (NICE).
- Brooks-Gordon, B., Bilby, C., & Wells, H. (2006). A Systematic review of psychological interventions for sexual offenders: I random controlled trials. *Journal of Forensic Psychiatry and Psychology, 17*, 442–466.
- Center for Sex Offender Management. (2008). *The comprehensive assessment protocol (CAP): A systemwide review of adult and juvenile sex offender management strategies*. Retrieved from <http://www.csom.org/pubs/cap/>
- Clarkin, J. F., & Levy, K. N. (2003). The influence of client variables on psychotherapy. In M. J. Lambert (Ed.), (2004) *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 194–226). New York: Wiley.
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*, 7–18.
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.
- Collaborative Outcome Data Committee (CODC). (2007a). *Sexual offender treatment outcome research: CODC guidelines for evaluation. Part 1: Introduction and overview*. Ottawa: Public Safety Canada.
- Collaborative Outcome Data Committee (CODC). (2007b). *The collaborative outcome data committee's guidelines for the evaluation*



- of sexual offender treatment outcome research. Part 2: CODC guidelines. Ottawa: Public Safety Canada.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation design & analysis issues for field settings*. Boston: Houghton Mifflin.
- Corabian, P., Ospina, M., & Harstall, C. (2010a). *Treatment for convicted adult male sex offenders*. Edmonton: Institute of Health Economics.
- Corabian, P., Dennett, L., & Harstall, C. (2010b). Treatment for convicted adult male sex offenders: An overview of systematic reviews. *Sexual Offender Treatment, 6*, 1–18.
- Craig, L. A., Browne, K. D., & Stringer, I. (2003). Treatment and sexual offense recidivism. *Trauma, Violence, & Abuse: A Review Journal, 4*, 70–89.
- Crits-Christoph, P., Johnson, J. E., Connolly Gibbons, M. B., & Gallop, R. (2013). Process predictors of the outcome of group drug counseling. *Journal of Consulting and Clinical Psychology, 81*(1), 23–34.
- Cullen, F. T., Agnew, R. A., & Wilcox, P. (Eds.). (2011). *Criminological theory: Past to present—Essential readings* (4th ed.). New York: Oxford University Press.
- Davidson, K. W., Goldstein, M., Kaplan, R. M., Kaufmann, P. G., Knatterud, G. L., & Orleans, C. T. (2003). Evidence-based behavioral medicine: What is it and how do we achieve it? *Annals of Behavioral Medicine, 26*, 161–171.
- Davidson, K. M., Tyrer, P., Tata, P., Cooke, D., Gumley, A., Ford, I., et al. (2009). Cognitive behaviour therapy for violent men with antisocial personality disorder in the community: An exploratory randomized controlled trial. *Psychological Medicine, 39*, 569–577.
- Dennis, J., Huband, N., Khan, O., Ferriter, M., Jones, H., & Powney, M. (2012). Psychological interventions for adults who have sexually offended or at a risk of sexually offending. *Cochrane Database of Systematic Reviews*, (12), CD007507.
- Doren, D. M. (2002). *Evaluating sex offenders: A manual for civil commitments and beyond*. Thousand Oaks, CA: Sage.
- Doren, D. M., & Yates, P. M. (2008). Effectiveness of sex offender treatment for psychopathic sexual offenders. *International Journal of Offender Therapy and Comparative Criminology, 52*, 234–245.
- Douglas, K. S., & Ogloff, J. R. P. (2003). Violence by psychiatric patients: The impact of archival measurement source on violence prevalence and risk assessment accuracy. *Canadian Journal of Psychiatry, 48*, 734–740.
- Douglas, K. S., Vincent, G. M., & Edens, J. F. (2006). Risk for criminal recidivism: The role of psychopathy. In C. Patrick (Ed.), *Handbook of psychopathy*. New York: Guilford.
- Dowden, C., Antonowicz, D., & Andrews, D. A. (2003). The effectiveness of relapse prevention with offenders: A meta-analysis. *International Journal of Offender Therapy and Comparative Criminology, 47*, 516–528.
- Duggan, C. (2008). Why are programmes for offenders with personality disorder not informed by the relevant scientific findings? *Philosophical Transactions of the Royal Society, 363*, 2599–2612.
- Duggan, C., & Dennis, J. (2014). The place of evidence in the treatment of sex offenders. *Criminal Behaviour and Mental Health, 24*, 153–162.
- Duggan, C., Huband, N., Smailagic, N., Ferriter, M., & Adams, C. (2007). The use of psychological treatments for people with personality disorder: A systematic review of randomized controlled trials. *Personality and Mental Health, 1*, 95–125.
- Durose, M. R., Cooper, A. D., & Snyder, H. N. (2014). *Recidivism of prisoners released in 30 states in 2005: Patterns from 2005 to 2010*. Department of Justice: Bureau of Justice Statistics.
- Egger, M., Davey Smith, G., & Altman, D. (2001). *Systematic reviews in health care meta-analysis in context*. London: BMJ Publishing.
- English, K., Jones, L., Pasini-Hill, D., Patrick, D., & Cooley-Towell, S. (2000). *The value of polygraph testing in sex offender management* (Final Research Report Submitted to the National Institute of Justice for Grant NO. D97LBVX0034). Denver, CO: Colorado Division of Criminal Justice, Office of Research and Statistics.
- Epperson, D. L. (2009). *Actuarial risk assessment with the MnSOST-remember: Updated research, scoring guidelines, and training*. Sacramento, CA.
- Eysenck, H. J. (1994). Meta-analysis and its problems. *British Medical Journal, 309*, 789–792.
- Fortney, T., Baker, J. N., & Levenson, J. (2009). A look in the mirror: Sexual abuse professionals' perceptions about sex offenders. *Victims and Offenders, 4*, 42–57.
- Friendship, C., Mann, R. E., & Beech, A. R. (2003). Evaluation of a national prison-based treatment program for sexual offenders in England and Wales. *Journal of Interpersonal Violence, 18*, 744–759.
- Furby, L., Weinrott, M., & Blackshaw, L. (1989). Sexual offender recidivism: A review. *Psychological Bulletin, 105*, 3–30.
- Gallagher, C. A., Wilson, D. B., Hirschfield, P., Coggeshall, M. B., & MacKenzie, D. L. (1999). A quantitative review of the effects of sex offender treatment on sexual reoffending. *Psychological Management Quarterly, 3*, 19–29.
- Gannon, T. A., Keown, K., & Polaschek, D. L. L. (2007). Increasing honest responding on cognitive distortions in child molesters: The bogus pipeline revisited. *Sexual Abuse: A Journal of Research and Treatment, 19*, 5–22.
- Gannon, T. A., & Polaschek, D. L. L. (2005). Do child molesters deliberately fake good on cognitive distortion questionnaires? An information processing-based investigation. *Sexual Abuse: A Journal of Research and Treatment, 17*, 183–200.
- Gaudiano, B. A., & Miller, I. W. (2013). The evidence-based practice of psychotherapy: Facing the challenges that lie ahead. *Clinical Psychology Review, 33*, 813–824.
- Gibbon, S., Duggan, C., Stoffers, J., Huband, N., Völlm, B. A., Ferriter, M., et al. (2010). Psychological interventions for antisocial personality disorder. *Cochrane Database of Systematic Reviews*, (6), CD007668.
- Gregerson, A., Nielsen, S. L., Lamber, M. J., Isakson, R., Smart, D., & Worthen, V. (2001). *Testing the consumer reports effectiveness scale: What can consumers report?* Paper presented at the Western Psychological Association, Honolulu, HI.
- Grossman, L. S., Martis, B., & Fichtner, C. G. (1999). Are sex offenders treatable? A research review. *Psychiatric Services, 50*, 349–361.
- Hall, G. C. N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology, 63*, 802–809.
- Hanson, R. K. (1997). How to know what works with sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 9*, 129–145.
- Hanson, R. K. (2014). Treating sexual offenders: How did we get here and where are we headed? *Journal of Sexual Aggression, 20*, 3–8.
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2008). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Public Safety Canada, 1–48*, 865–891.
- Hanson, R. K., & Bussiere, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology, 66*, 348–362.
- Hanson, R. K., Gordon, A., Harris, A. J. R., Marques, J. K., Murphy, W., Quinsey, V. L., et al. (2002). First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 14*, 169–194.
- Hanson, R. K., & Harris, A. J. R. (2000). Where should we intervene? Dynamic predictors of sex offense recidivism. *Criminal Justice and Behavior, 27*, 6–35.
- Hanson, R. K., Morton, K. E., & Harris, A. J. R. (2003). Sexual offender recidivism risk: What we know and what we need to know. In R. A. Prentky, E. S. Janus, & M. C. Seto (Eds.), *Sexually coercive behavior: Understanding and management* (pp. 154–166). New York: Annals of the New York Academy of Sciences.
- Hanson, R. K., & Morton-Bourgon, K. (2004). *Predictors of sexual recidivism: An updated meta-analysis 2004–02*. Ottawa: Public Safety and Emergency Preparedness.

- Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73*, 1154–1163.
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009a). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior, 36*, 865–891.
- Hanson, R. K., Phenix, A., & Helmus, L. (2009b). *Static-99(R) and static-to interpret and report in light of recent research*. Pre-conference workshop presented at the Annual Meeting of the Association for the Treatment of Sexual Abusers. Dallas, TX, September, 2009.
- Hanson, R. K., Steffy, R. A., & Gauthier, R. (1993). Long-term recidivism of child molesters. *Journal of Consulting and Clinical Psychology, 61*, 646–652.
- Hanson, R. K., & Thornton, D. (2003, October). *Models of real reoffenses rates: Clinical implications*. Paper presented at the annual meeting of the Association for the Treatment of Sexual Abusers, St. Louis, MO.
- Harkins, L., & Beech, A. (2007). A review of the factors that can influence the effectiveness of sexual offender treatment: Risk, need, responsivity, and process issues. *Aggression and Violent Behavior, 12*, 615–627.
- Harkins, L., Flak, V., Beech, A., & Woodhams, J. (2012). Evaluation of a community-based sex offender treatment program using a good lives model approach. *Sexual Abuse: A Journal of Research and Treatment, XX(X)*, 1–25.
- Harris, A., & Hanson, R. K. (2004). *Sex offender recidivism: A simple question (2004–03)*. Ottawa: Public Safety and Emergency Preparedness.
- Harris, G., & Rice, M. (2006). Treatment of psychopathy: A review of empirical findings. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 555–572). New York: Guilford.
- Harris, G. T., & Rice, M. E. (2007). Characterizing the value of actuarial violence risk assessments. *Criminal Justice and Behavior, 34*, 1638–1658.
- Harris, G. T., Rice, M., & Quinsey, V. (1998). Appraisal and management of risk in sexual aggressors: Implications for criminal justice policy. *Psychology, Public Policy, and Law, 4*, 73–115.
- Heil, P., Simons, D., & Ahlmeyer, S. (2003). Crossover sexual offenses. *Sexual Abuse: A Journal of Research and Treatment, 15*, 221–236.
- Helmus, L. (2009). *Re-norming Static-99 recidivism estimates: Exploring base rate variability across sex offender samples*. Masters thesis submitted to Carleton University.
- Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist, 58*, 78–79.
- Hill, C. E., & Lambert, M. J. (2004). Methodological issues in studying psychotherapy processes and outcomes. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 84–136). New York: Wiley.
- Hindman, J., & Peters, J. (2001). Polygraph testing leads to better understanding adult and juvenile sex offenders. *Federal Probation, 65*, 8–15.
- Hindman, J., & Peters, J. M. (2010). Shedding light on the histories of sexual offenders using clinical polygraphy. In A. Schlank (Ed.), *The sexual predator* (Vol. IV, pp. 1–24). Kingston, NJ: Civic Research Institute. Chapter 20.
- Ho, D. K., & Ross, C. C. (2012). Cognitive behaviour therapy for sex offenders: Too good to be true? *Criminal Behaviour Mental Health, 22*, 1–6.
- Hoberman, H. M. (2015). Forensic psychotherapy for sexual offenders: Possible factors contributing to its apparent ineffectiveness. In A. Phenix & H. M. Hoberman (Eds.), *Sexual offending: Predisposing conditions, assessments and management*. New York: Springer.
- Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *British Medical Journal, 319*, 670–674.
- Hollon, S. D., & Beck, A. T. (2004). Cognitive and cognitive-behavioral therapies. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 447–492). New York: Wiley.
- Hollon, S. D., & Beck, A. T. (2013). Cognitive and cognitive-behavioral therapies. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 393–442). New York: Wiley.
- Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). The evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51*, 1059–1064.
- Jones, N., Pellissier, B., & Klein-Saffran, J. (2006). Predicting sex offender treatment entry among individuals convicted of sexual offense crimes. *Sexual Abuse: A Journal of Research and Treatment, 18*(1), 83–98.
- Kaplan, M. D., & Krueger, R. B. (2012). Cognitive-behavioral treatment of the paraphilias. *Israeli Journal of Psychiatric and Related Sciences, 49*, 291–296.
- Kazdin, A. E. (1986). The evaluation of psychotherapy: Research designs and methodology. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 26–68). New York: Wiley.
- Kazdin, A. E. (1994). Methodology, design, and evaluation in psychotherapy research. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 19–71). New York: Wiley.
- Kelly, A. E. (2000). Helping construct desirable identified: Aronson self-presentational view of psychotherapy. *Psychological Bulletin, 126*, 475–494.
- Kendall, P. C., Holmbeck, G., & Verduin, T. (2004). Methodology, design, and evaluation in psychotherapy research. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 16–43). New York: Wiley.
- Kenworthy, T., Adams, C. E., Bilby, C., Brooks-Gordon, B., & Fenton, M. (2003). Psychological interventions for those who have sexually offended or at risk of sexually offending. *Cochrane Database of Systematic Reviews, (4)*, CD004858.
- Kenworthy, T., Adams, C. E., Bilby, C., Brooks-Gordon, B., & Fenton, M. (2004). Psychological interventions for those who have sexually offended or are at risk of offending (Cochrane Review). *The Cochrane Library, (3)*. Chichester, UK: Wiley.
- Kingston, D. A., Yates, P. M., & Firestone, P. (2011). The self-regulation model of sexual offending: Relationship to risk and need. *Law and Human Behavior, 36*, 215–244.
- Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology, 62*, 1009–1016.
- Lambert, M. J. (1992). Psychotherapy outcome research: Implications for integrative and eclectic therapists. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration*. New York: Basic Books.
- Lambert, M. J. (Ed.). (2012). *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.). New York: Wiley.
- Lambert, M. J. (2013). Outcome in psychotherapy: The past and important advances. *Psychotherapy, 50*(1), 42–51.
- Lambert, M. F., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139–193). New York: Wiley.
- Landenberger, N. A., & Lipsey, M. W. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology, 1*–26.
- Langan, P. A., Schmitt, E. L., & Durose, M. R. (2003). *Recidivism of sex offenders released from prison in 1994*. Department of Justice: Bureau of Justice Statistics.

- Langstrom, N., Enebrink, P., Lauren, E.-M., Lindblom, J., Werko, S., & Hanson, R. K. (2013). Preventing sexual abusers of children from reoffending: Systematic review of medical and psychological interventions. *British Medical Journal*, *347*, 1–11.
- Langton, C. M., Barbaree, H. E., Harkins, L., & Peacock, E. J. (2006a). Sex offenders' response to treatment and its association with recidivism as a function of psychopathy. *Sexual Abuse: A Journal of Research and Treatment*, *18*(1), 99–120.
- Langton, C. M., Barbaree, H. E., Harkins, L., & Peacock, E. J. (2006b). Sex offenders' response to treatment and its association with recidivism as a function of psychopathy. *Sexual Abuse: A Journal of Research and Treatment*, *18*, 99–120.
- Langton, C. M., Barbaree, H. E., Seto, M. C., Peacock, E. J., Harkins, L., & Hansen, K. T. (2007). Actuarial assessment of risk for reoffense among adult sex offenders: Evaluating the predictive accuracy of the Static-2002 and five other instruments. *Criminal Justice and Behavior*, *34*, 37–59.
- Latessa, E. J., & Lowenkamp, C. T. (2006). What works in reducing recidivism. *University of St. Thomas Law Journal*, *13*, 521–535.
- Laws, D. R. (1996). Relapse prevention or harm reduction? *Sexual Abuse*, *8*, 243–247.
- Levenson, J., & D'Amora, D. (2005). An ethical paradigm for sex offender treatment: Response to Glaser. *Western Criminology Review*, *6*(1), 145–153.
- Levenson, J. S., & Prescott, D. S. (2013). Déjà vu: From Furby to Langstrom and the evaluation of sexual offender treatment effectiveness. *Journal of Sexual Aggression*, 1–10.
- Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science*, *2*, 53–70.
- Lilienfeld, S. O. (2011). Public skepticism of psychology: Why many people perceive the study of human behavior as unscientific. *American Psychologist*, *67*, 1–19.
- Lipsey, M. W., & Cullen, F. T. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science*, *3*, 297–320.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, *48*, 1181–1209.
- Longmore, R. J., & Worrell, M. (2007). Do we need to challenge thoughts in cognitive behavior therapy? *Clinical Psychology Review*, *27*, 173–187.
- Looman, J., Abracen, J., Serin, R., & Marquis, P. (2005). Psychopathy, treatment change and recidivism in high risk high need sexual offenders. *Journal of Interpersonal Violence*, *20*, 549–568.
- Losel, F., & Schmucker, M. (2005). The effectiveness of treatment for sex offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology*, *1*, 117–146.
- Lynch, D., Laws, K. R., & McKenna, P. J. (2010). Cognitive behavioural therapy for major psychiatric disorder: Does it really work? A meta-analytical review of well-controlled trials. *Psychological Medicine*, *40*, 9–24.
- Mann, R. (2009). Sex offender treatment: The case for manualization. *Journal of Sexual Aggression*, *15*, 121–131.
- Mann, R., & Marshall, W. L. (2009). Advances in the treatment of adult incarcerated sexual offenders. In A. Beech, L. Craig, & K. Browne (Eds.), *Assessment and treatment of sex offenders: A handbook* (pp. 329–347). Chichester: Wiley.
- Mann, R., & Shingler, J. (2009). Schema-driven cognition in sexual offenders: There, assessment and treatment. In W. Marshall, Y. Fernandez, L. Marshall, & G. Serran (Eds.), *Sexual offender treatment: Controversial issues* (pp. 173–185). Chichester: Wiley.
- Mann, R. E., Webster, S. D., Schofield, C., & Marshall, W. L. (2004). Approach versus avoidance goals in relapse prevention with sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, *16*(1), 65–75.
- Marlatt, G. A. (1982). Research and treatment with alcohol problems: Frontiers for the eighties. *Bulletin of the Society of Psychologists in Substance Abuse*, *1*, 145–150.
- Marques, J. K., Day, D. M., Nelson, C., & West, M. A. (1994). Effects of cognitive-behaviour treatment on sex offender recidivism: Preliminary results of a longitudinal study. *Criminal Justice and Behaviour*, *21*, 28–54.
- Marques, J., Nelson, C., Alarcon, J., & Day, D. (2000). Preventing relapse in sex offenders: What we learned from SOTEP's experimental treatment program. In D. Laws, S. Hudson, & T. Ward (Eds.), *Remaking relapse prevention with sex offenders* (pp. 321–340). Thousand Oaks, CA: Sage.
- Marques, J. K., Wiederanders, M., Day, D. M., Nelson, C., & Ommeren, A. V. (2005). Effects of a relapse prevention program on sexual recidivism: Final results from California's sex offender treatment and evaluation project (SOTEP). *Sexual Abuse: A Journal of Research and Treatment*, *17*, 79–107.
- Marquis, J. (1970). Orgasmic reconditioning: Changing sexual object choice through controlling masturbation fantasies. *Journal of Behavior Therapy and Experimental Psychiatry*, *1*, 263–271.
- Marshall, W. L. (1993). The treatment of sex offenders: What does the outcome data tell us? A reply to Quinsey, Harris, Rice, and Lalumière. *Journal of Interpersonal Violence*, *8*, 524–530.
- Marshall, W. L. (1994). Treatment effects on denial and minimization in incarcerated sex offenders. *Behaviour Research and Therapy*, *32*, 559–564.
- Marshall, W. L. (1996). Assessment, treatment, and theorizing about sex offenders: Developments during the past twenty years and future directions. *Criminal Justice and Behavior*, *23*(1), 162–199.
- Marshall, W. L. (1999). Current status of North American assessment and treatment programs for sexual offenders. *Journal of Interpersonal Violence*, *14*, 221–239.
- Marshall, W. L. (2005). Therapist style in sexual offender treatment: Influence on indices of change. *Sexual Abuse: A Journal of Research and Treatment*, *17*(2), 109–116.
- Marshall, W. L. (2009). Manualization: A blessing or a curse? *Journal of Sexual Aggression*, *15*, 109–120.
- Marshall, W. L., & Barbaree, H. E. (1990). An integrated theory of the etiology of sexual offending. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 257–275). New York: Plenum Press.
- Marshall, L. E., & Marshall, W. L. (2007). Preparatory programs for sexual offender treatment. In D. S. Prescott (Ed.), *Knowledge & practice* (pp. 108–123). Oklahoma: Wood 'N' Barnes.
- Marshall, W. L., Marshall, L., Serran, G., & O'Brien, M. (2009). Psychological treatment of sexual offenders. In F. Saleh, A. Grudzinskas Jr., J. Bradford, & D. Brodsky (Eds.), *Sex offenders: Identification, risk assessment, treatment, and legal issues* (pp. 159–170). New York: Oxford University Press.
- Marshall, W. L., Marshall, L. E., Serran, G. A., & O'Brien, M. D. (2011). *Rehabilitating sexual offenders: A strength-based approach*. Washington, DC: American Psychological Association Press.
- Marshall, W. L., & McGuire, J. (2003). Effect sizes in treatment of sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, *46*, 653–663.
- Marshall, W. L., O'Brien, M., & Marshall, L. (2009). Modifying sexual preferences. In A. Beech, L. Craig, & K. Browne (Eds.), *Assessment and treatment of sex offenders: A handbook* (pp. 311–327). Chichester: Wiley.
- Marshall, W. L., & Pithers, W. D. (1994). A reconsideration of treatment outcome with sex offenders. *Criminal Justice and Behavior*, *21*, 10–27.
- Marshall, W. L., Serran, G. A., Fernandez, Y. M., Mulloy, R., Mann, R. E., & Thornton, D. (2003). Therapist features in sex offender treatment: Their reliable identification and influence on behaviour change. *Clinical Psychology and Psychotherapy*, *9*, 395–405.

- Maruna, S., & Mann, R. E. (2006). A fundamental attribution error? Rethinking cognitive distortions. *Legal and Criminological Psychology, 11*, 155–177.
- Matt, G. E. (1989). Decision rules for selecting effect sizes in meta-analysis: A review and reanalysis of psychotherapy outcome studies. *Psychological Bulletin, 105*, 106–115.
- McConaghy, N. (1993). *Sexual behavior: Problems and management*. New York: Plenum Press.
- McConaghy, N. (1999). Methodological issues concerning evaluation of treatment for sexual offenders: Randomization, treatment dropouts, untreated controls, and within-treatment studies. *Sexual Abuse: A Journal of Research and Treatment, 11*, 183–193.
- McGrath, R. J., Lasher, M. P., & Cumming, G. F. (2012). The sex offender treatment intervention and progress scale (SOTIPS): Psychometric properties and incremental predictive validity with Static-99R. *Sexual Abuse: A Journal of Research and Treatment, 24*, 431–458.
- Miner, M. H. (1997). How can we conduct treatment outcome research? *Sexual Abuse: A Journal of Research and Treatment, 9*, 95–110.
- Miner, M. H., & Dwyer, S. M. (1995). Analysis of dropouts from outpatient sex offender treatment. *Journal of Psychology and Human Sexuality, 7*, 77–93.
- Mitchell, D., Simourd, D. J., & Tafra, R. C. (2014). Introduction: Critical issues and challenges facing forensic CBT practitioners. In R. C. Tafra & D. Mitchell (Eds.), *Forensic CBT: A handbook for clinical practice* (pp. 302–326). Chichester, UK: Wiley Blackwell.
- Moher, D., Tetzlaff, J., Tricco, A. C., Sampson, M., & Altman, D. G. (2007). Epidemiology and reporting characteristics of systematic reviews. *PLoS Medicine, 4*(3), e78. Retrieved May 15, 2014.
- Monahan, J., Steadman, H., Silver, E., Appelbaum, P., Robbins, P. C., Mulvey, E., et al. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York: Oxford University Press.
- Munder, T., Brüttsch, O., Leonhart, R., Gerger, H., & Barth, J. (2013). Researcher allegiance in psychotherapy outcome research: An overview of reviews. *Clinical Psychology Review, 33*, 501–511.
- Nasrallah, H. A. (2013). Repositioning psychotherapy as a neurobiological intervention. *Current Psychiatry, 12*, 18–19.
- Nicholson, R. A., & Berman, J. S. (1983). Is follow-up necessary in evaluating psychotherapy? *Psychological Bulletin, 93*, 261–278.
- Nunes, K. L., Babchishin, K. M., & Cortoni, F. (2011). Measuring treatment change in sex offenders: Clinical and statistical significance. *Criminal Justice and Behavior, 38*, 157–173.
- Nunes, K. L., & Cortoni, F. (2008). Dropout from sex-offender treatment and dimensions of risk of sexual recidivism. *Criminal Justice and Behavior, 35*, 24–33.
- Ogilvie, A. J., Abreu, I., & Safran, J. D. (2005). What findings do psychotherapy researchers use in their own practice? A survey of the society for psychotherapy research. *The New School Psychology Bulletin, 3*(2), 17–34.
- Ogloff, J. R. P., Wong, S., & Greenwood, A. (2003). Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences & the Law, 8*, 181–190.
- Olver, M. E., Beggs Christofferson, S. M., Grace, R. C., & Wong, S. C. P. (2013). Incorporating change information into sexual offender risk assessments using the violence risk scale-sexual offender version. *Sexual Abuse: A Journal of Research and Treatment, 26*, 472–499.
- Olver, M. E., Kingston, D. A., Nicholaichuk, T., & Wong, S. C. P. (2014). A psychometric examination of treatment change in a multisite sample of treated Canadian Federal sexual offenders. *Law and Human Behavior, 38*, 544–559.
- Olver, M. E., Nicholaichuk, T. P., Gu, D., & Wong, S. C. P. (2013). Sex offender treatment outcome, actuarial risk, and the aging sex offender in Canadian corrections: A long-term follow-up. *Sexual Abuse: A Journal of Research and Treatment, 25*, 396–422.
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2011). A meta-analysis of predictors of offender treatment attrition and its relationship to recidivism. *Journal of Consulting and Clinical Psychology, 79*, 6–21.
- Olver, M. E., Wong, S. C. P., Nicholaichuk, T., & Gordon, A. (2007). The validity and reliability of the violence risk scale-sexual offender version: assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment, 19*, 318–329.
- Olver, M. E., & Wong, S. C. P. (2009). Therapeutic responses of psychopathic sexual offenders: Treatment attrition, therapeutic change, and long term recidivism. *Journal of Consulting and Clinical Psychology, 77*, 328–336.
- Olver, M. E., & Wong, S. C. P. (2011). A comparison of static and dynamic assessment of sexual offender risk and need in a treatment context. *Criminal Justice and Behavior, 38*, 113–126.
- Pake, D. R. (2010). There's no such thing as success in treatment. *ATSA Forum, XXII*, 9–12.
- Paul, G. L. (1967). Strategy of outcome research in psychotherapy. *Journal of Consulting Psychology, 31*, 109–118.
- Pekarik, G., & Guidry, L. L. (1999). Relationship of satisfaction to symptom change, followup adjustment, and clinical significance in private practice. *Professional Psychology: Research and Practice, 30*, 474–478.
- Polizzi, D. M., MacKenzie, D. L., & Hickman, L. J. (1999). What works in adult sex offender treatment? A review of prison- and non-prison-based treatment programs. *International Journal of Offender Therapy and Comparative Criminology, 43*, 357–374.
- Prentky, R. A., Gabriel, A. M., & Coward, A. I. (2011). Treatment for perpetrators of sexual violence. In M. P. Koss, J. W. White, & A. E. Kazdin (Eds.), *Violence against women and children* (Navigating solutions, Vol. 2, pp. 115–135). Washington, DC: American Psychological Association.
- Prentky, R. A., Lee, A. F. S., Knight, R. A., & Cerce, D. (1997). Recidivism rates among child molesters and rapists: A methodological analysis. *Law and Human Behavior, 21*, 635–659.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk* (2nd ed.). Washington, DC: American Psychological Association.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Lalumière, M. L. (1993). Assessing treatment efficacy in outcome studies of sex offenders. *Journal of Interpersonal Violence, 8*, 512–523.
- Quinsey, V. L., Khanna, A., & Malcolm, B. (1998). A retrospective evaluation of the regional treatment center sex offender treatment program. *Journal of Interpersonal Violence, 13*, 621–644.
- Quinsey, V. L., Lalumière, M. L., Rice, M. E., & Harris, G. T. (1995). Predicting sexual offenses. In J. C. Campbell (Ed.), *Assessing dangerousness: Violence by sexual offenders, batterers, and child abusers* (pp. 114–137). Thousand Oaks, CA: Sage.
- Rice, M. E. (1997). Violent offender research and implications for the criminal justice system. *American Psychologist, 52*, 414–423.
- Rice, M. E., & Harris, G. T. (1997). Cross-validation and extension of the violence risk appraisal guide for child molesters and rapists. *Law and Human Behavior, 21*, 435–448.
- Rice, M. E., & Harris, G. T. (2003). The size and sign of treatment effects in therapy for sex offenders. In R. A. Prentky, E. Janus, & M. C. Seto (Eds.), *Understanding and managing sexually coercive behavior* (pp. 428–440). New York: Annals of the New York Academy of Science.
- Rice, M. E., & Harris, G. T. (2012). Treatment for adult sexual offenders: May we reject the null hypothesis? In K. Harrison & B. Bradley (Eds.), *Handbook of legal and ethical aspects of sexual offender treatment and management* (pp. 219–235). London: Wiley-Blackwell.

- Rice, M. E., Harris, G. T., & Quinsey, V. L. (2001). Research on the treatment of adult sex offenders. In J. B. Ashford, B. D. Sales, & W. H. Reid (Eds.), *Treating adult and juvenile offenders with special needs* (pp. 291–312). Washington, DC: American Psychological Association.
- Rice, M. E., Quinsey, V. L., & Harris, G. T. (1991). Sexual recidivism among child molesters released from a maximum security psychiatric institution. *Journal of Consulting and Clinical Psychology, 59*, 381–386.
- Romero, J. J., & Williams, L. M. (1983). Group psychotherapy and intensive probation supervision with sex offenders. *Federal Probation, 47*(4), 36–42.
- Rosenblatt, A., & Rosenblatt, J. A. (2002). Assessing the effectiveness of care for youth with severe emotional disturbances: Is there agreement between popular outcomes measures? *Journal of Behavioral Health Services and Research, 29*, 259–273.
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., & Richardson, W. S. (1996). Evidence based medicine; what it is and it isn't. *British Medical Journal, 312*, 71–72.
- Salekin, R. T. (2002). Psychopathy and therapeutic pessimism: Clinical lore or clinical reality? *Clinical Psychology Review, 22*, 79–112.
- Sanchez-Meca, J. (1997). Methodological issues in the meta-evaluation of correctional treatment. In S. Redondo, V. Garrido, J. Perez, & R. Barberet (Eds.), *Advances in psychology and law: International contributions* (pp. 486–498). New York: Walter de Gruyter.
- Schlack, A. (2010). The baby, the bathwater, and pendulum swings—The need to slow down and evaluate research critically. In A. Schlack (Ed.), *The sexual predator* (Vol. IV, pp. 1–9). Kingston, NJ: Civic Research Institute. Chapter 22.
- Schmucker, M., & Lösel, F. (2008). Does sexual offender treatment work? A systematic review of outcome evaluations. *Psicothema, 20*, 10–19.
- Seager, J. A., Jellicoe, D., & Dhaliwal, G. K. (2004). Refusers, drop-outs, and completers: Measuring sex offender treatment efficacy. *International Journal of Offender Therapy and Comparative Criminology, 48*, 600–612.
- Serin, R. C., Lloyd, C. D., Helmus, L., Derkzen, D. M., & Luong, D. (2013). Does intra-individual change predict offender recidivism? Searching for the holy grail in assessing offender change. *Aggression and Violent Behavior, 18*, 32–53.
- Seto, M. C., & Barbaree, H. E. (1999). Psychopathy, treatment behavior, and sex offender recidivism. *Journal of Interpersonal Violence, 14*, 1235–1248.
- Seto, M. C., Marques, J. K., Harris, G. T., Chaffin, M., Lalumière, M. L., Miner, M. H., et al. (2008). Good science and progress in sex offender treatment are intertwined: A response to Marshall and Marshall (2007). *Sexual Abuse: A Journal of Research and Treatment, 20*(3), 247–255.
- Shadish, W. R., Matt, G., Novaro, A., Siegle, G., Crits-Christoph, P., Hazelrigg, M., et al. (1997). Evidence that therapy works under clinically representative conditions. *Journal of Consulting and Clinical Psychology, 65*, 355–365.
- Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review, 17*, 881–901.
- Smith, M. L., Glass, G. V., & Miller, T. (1980). *The benefits of psychotherapy*. Baltimore, MD: Johns Hopkins University Press.
- Spring, B. (2007). Evidence-based practice in clinical psychology: What it is, why it matters; what you need to know. *Journal of Clinical Psychology, 63*, 611–631.
- Stanovich, K. (2009). *How to think straight about psychology* (9th ed.). Boston: Pearson.
- Stewart, R. E., & Chambless, D. L. (2007). Does psychotherapy research inform treatment decisions in private practice? *Journal of Clinical Psychology, 63*, 267–281.
- Stirpe, T. S., Wilson, R. J., & Long, C. (2001). Goal attainment scaling with sexual offenders: A measure of clinical impact at posttreatment and at community follow-up. *Sexual Abuse: A Journal of Research and Treatment, 13*, 65–77.
- Stoffers, J. M., Völlm, B. A., Rucker, G., Timmer, A., Huband, N., & Lieb, K. (2012). Psychological therapies for people with borderline personality disorder. *Cochrane Database of Systematic Reviews*, Issue 8, Art. No.: CD005652.
- Tafra, R. C., & Mitchell, D. (2013). *Forensic CBT: A handbook for clinical practice* (pp. 302–326). Chichester, UK: Wiley Blackwell.
- Terry, K. J., & Mitchell, E. W. (2001). Motivation and sex offender treatment efficacy: Leading a horse to water and making it drink? *International Journal of Offender Therapy and Comparative Criminology, 45*, 663–672.
- Thornton, D., & Blud, L. (2007). Psychopathy and deception. In H. Hervé & J. C. Yuille (Eds.), *The psychopath: Theory, research, and practice* (pp. 505–539). Mahwah, NJ: Lawrence Erlbaum Associates.
- Tierney, D. W., & McCabe, M. P. (2002). Motivation for behavior change among sex offenders. A review of the literature. *Clinical Psychology Review, 22*, 113–129.
- Tyrer, P., & Johnson, T. (1996). Establishing the severity of personality disorder. *American Journal of Psychiatry, 153*, 1593–1597.
- Wakeling, H., & Barnett, G. D. (2014). The relationship between psychometric scores and reconviction in sexual offenders undertaking treatment. *Aggression and Violent Behavior, 19*, 138–145.
- Walsh, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Report, 110*, 639–644.
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Erlbaum.
- Wampold, B. E., Minami, T., Baskin, T. W., & Tierney, S. C. (2002). A meta-(re)analysis of the effects of cognitive therapy versus 'other therapies' for depression. *Journal of Affective Disorders, 68*, 159–165.
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, E., Benson, K., & Hyun-nie, A. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes." *Psychological Bulletin, 122*, 203–215.
- Ward, T., & Beech, A. (2006). An integrated theory of sexual offending. *Aggression and Violent Behavior, 11*, 44–63.
- Ward, T., Beech, A., & Fischer, S. (2015). An integrated theory of sexual offending. In A. Phenix & H. M. Hoberman (Eds.), *Sexual offending: Predisposing antecedents, assessments and management*. New York: Springer.
- Ward, T., & Hudson, S. M. (1998). A self-regulation model of the offense process. *Sexual Abuse: A Journal of Research and Treatment, 10*, 141–157.
- Ward, T., & Gannon, T. (2006). Rehabilitation, etiology, and self-regulation: The good lives model of sexual offender treatment. *Aggression and Violent Behavior, 11*, 77–94.
- Weiss, I., Rabinowitz, J., & Spiro, S. (1996). Agreement between therapists and clients in evaluating therapy and its outcomes: Literature review. *Administration and Policy in Mental Health, 23*, 493–511.
- Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology, 69*, 875–899.
- Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin, 130*, 631–663.
- Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2005). EBP EST: Reply to Crits-Christoph et al. (2005). *Psychological Bulletin, 131*, 427–431.

- Wheeler, J., & Covell, C. (2013). Recidivism risk reduction therapy (3RT): Cognitive-behavioral approaches to treating sexual offense behavior. In R. C. Tafrate & D. Mitchell (Eds.), *Forensic CBT: A handbook for clinical practice* (pp. 302–326). Chichester, UK: Wiley Blackwell.
- White, P., Bradley, C., Ferriter, M., & Hatzipetrou, L. (1998). Managements for people with disorders of sexual preference and for convicted sexual offenders. *Cochrane Database of Systematic Reviews*, (2), CD000251.
- Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice*, 24, 190–195.
- Williams, F., Wakeling, H., & Webster, S. (2007). A psychometric study of six self-report measures for use with sexual offenders with cognitive and social functioning deficits. *Psychology, Crime & Law*, 13, 505–522.
- Yates, P. M. (2015). Models of sex offender treatment. In A. Phenix & H. M. Hoberman (Eds.), *Sexual offending: Predisposing conditions, assessments and management*. New York: Springer.
- Yates, P. M., Prescott, D. S., & Ward, T. (2010). *Applying the good lives and self-regulation models to sex offender treatment: A practical guide for clinicians*. Brandon, VT: Safer Society Press.
- Yates, P. M., & Ward, T. (2008). Good lives, self-regulation, and risk management: An integrated model of sexual offender assessment and treatment. *Sexual Abuse in Australia and New Zealand: An Interdisciplinary Journal*, 1, 3–20.