

Running Head: POST CONVICTION SEX OFFENDER POLYGRAPH TESTING

The Current Role of Post Conviction Sex Offender Polygraph Testing

In Sex Offender Treatment

Ron Kokish

Delson-Kokish Associates

Trinidad, California

Submitted: March 25, 2002

Revised: November 7, 2002

Revised: March 10, 2003

Accepted: April 22, 2003

Published - Journal of Child Sexual Abuse, Volume 12, numbers 3 & 4

Abstract

Polygraph testing is becoming increasingly important in sex offender treatment. Polygraph advocates cite dramatic increases in historical disclosures that presumably allow more precise targeting of treatment interventions, earlier detection of risky behaviors that often lead to new offenses, and improved treatment and supervision compliance. Based on this, they believe the procedure supports desirable behavior that continues to various degrees after treatment and supervision end. Opponents cite ethical problems related to inaccurate results, unproven accuracy rates, and the risk that examinees may be coerced into making false admissions. To counter these criticisms, proponents have developed standards, best practices, and examiner training and certification programs intended to reduce error rates and address ethical issues. Opponents argue that these measures have not been tested and that empirically established error rates and best practices may not be possible for a variety of reasons. This article reviews the current situation, leaving readers to decide the wisdom and ethics of using polygraph testing in their own practices.

Keywords: polygraph, sex offenders, post conviction sex offender testing, sex offender treatment

The Current Role of Post Conviction Sex Offender Polygraph Testing

In Sex Offender Treatment

How Polygraphy Works

Polygraph testing consists of taking physiological measurements, interpreting the results, and offering a professional opinion as to whether the examinee was attempting deception when she/he answered certain (“relevant”) questions. In practice, heart rate, blood pressure, breathing patterns, and galvanic skin response are the most typical measures used. This is because they are easily accessible and affordable, and accurate instrumentation is available. Changes in these physiological measures are believed to be associated with involuntary arousal in response to fear. This arousal is often called the “fight/flight response.” Theoretical underpinnings for the procedure include a hypothesis that almost all people experience fear of being discovered when they lie. Polygraph examiners attempt to isolate and measure this fear in relation to the issue in question. Since the measurements are physiological and their interpretation is based on psychological theory, polygraphy is considered a “psycho-physiological” procedure (Hassett, 1978).

Polygraph Testing and Treatment Goals

We treat sex offenders to achieve three major goals. These are, in order of importance:

1. To reduce the number of sexual reoffenses committed by treated offenders;
2. To reduce the number of general criminal offenses, but especially violent offenses committed by treated offenders; and
3. To change treated offenders in ways that lead to increased social fulfillment for them and those whose lives they touch.

Several recent meta-analyses indicate that we are doing a credible job reaching the first two goals. Based on these studies, we can reasonably claim to be achieving an overall 40% treatment effect for 5-10 years following discharge (Alexander, 1999; Grossman, Martis, & Fichtner, 1999; Hall, 1995; Hanson et al., 2002), though this effect certainly varies between offender categories (e.g., incest offenders, rapists) and from program to program. Our third goal is more difficult to measure. Only subjective reports from therapists, supervision officers, patients, and their families inform us about the degree to which we are attaining it, but this does not detract from its importance.

During treatment, we look at certain objectives that we believe are related to achieving our goals. Among these are:

1. Patient honesty with therapists, group members, family, and other important people;
2. Patient compliance with treatment procedures; and
3. Patient compliance with supervision conditions.

Proponents of Post Conviction Sex Offender (polygraph) Testing (PCSOT) assert that PCSOT significantly enhances their ability to attain these objectives, and that its use is thus in the community's and the offenders' best interests.

Polygraph examinations have been used to monitor probationers and particularly child abusers, since the mid 1960's (Abrams, 1989b). Sex offender treatment programs (SOTP's) began using polygraph examinations to obtain more complete sexual histories and to monitor program compliance as long ago as 1975 (Wolfe, personal communication, May 10, 1999). Over the years, polygraph examinations have become an increasingly important tool for sexual offender treatment. A 1996 survey found that over 30% of 960 responding programs utilized

PCSOT to varying degrees (Burton & Smith-Darden, 2000). A more recent survey found that at least some SOTP's in 33 of the 50 United States, Germany, and Canada use periodic polygraph examinations (Abrams & Simmons, 2000). The survey reported that 9.8% of probation and parole supervisors surveyed in 1994 said their agencies often or always used polygraph with adult sex offenders on their caseloads. This number rose to 16.3% in a similar 1998 survey. By 2000, approximately one-fourth of respondents reported often or always using polygraph with their agencies' adult sex offender probationers and parolees. At least one state (Colorado) now requires state recognized treatment programs to utilize polygraph examinations, and polygraphy use is encouraged by the Association for the Treatment of Sexual Abusers (ATSA) (ATSA, 2001).

Eliciting Increased Information

Polygraph utility flows from its ability to elicit information from offenders who might be less forthcoming with traditional interviewing techniques (ATSA, 1997). Sex offender treatment providers report significant, often dramatic increases in information disclosed by offenders who anticipate taking a polygraph examinations. For example, one study (English, Jones, Pasini-Hill, Patrick, & Cooley-Towell, 2000) comparing 180 offenders before and after testing found the following:

1. An 80 % increase in the number of offenders admitting male victims;
2. A 190% increase in the number of offenders admitting to both male and female victims;
3. A 230% increase in admission of both juvenile and adult victims;
4. A 60% increase in admission of high-risk behaviors;¹ and
5. A 196% increase in admission of more than one kind of high risk behavior.

Another study (Hindman & Peters, in press) compared 98 intra-familial sex offenders treated without PCSOT between 1978 and 1983 to 129 comparable offenders who were treated with polygraphy by the same program between 1983 and 1988. When they entered treatment, the non-polygraph group reported an average of 1.2 victims. This increased to 1.5 victims by the time they completed their sexual histories. The polygraph group reported an average of 1.3 victims at program entry with an increase to 9 victims by the time they completed their sexual histories. In addition to admitting fewer victims, 67% of the non-polygraph group reported being sexually abused as minors, and 21% reported sexually abusing others as minors. Those figures were nearly reversed for the polygraph group, with only 29% reporting they were sexually abused as minors and 71% admitting they sexually abused others as minors.

In a later replication study using data gathered between 1988 and 1994, the same researchers (Hindman & Peters, 2000) reported the non-polygraph group ($n = 76$) admitting an average of 2.5 victims, with 65% of the offenders claiming they were abused as minors and 22% admitting they had sexually abused others. The polygraph group ($n = 152$) admitted a mean average of 13.6 victims, only 32% claimed they were sexually abused as minors, and 68% admitted abusing others. Additionally, 17% of the non-polygraphed (primarily male) offenders admitted sexually abusing a male victim. Among the otherwise comparable polygraphed offenders, 30% admitted sexually abusing males.

Treatment providers who use polygraphy argue that complete historical information allows them to hold offenders more accountable and helps them to target treatment interventions more precisely, all of which ultimately helps lower reoffense rates for treated offenders. Many offenders apparently agree. Kokish, Blasingame, and Plaud (2002) conducted an anonymous survey of 95 sex offenders being treated in three separate programs that use PCSOT regularly

and adhere to current standards. Respondents stated by nearly 7-to-1 that periodic polygraph testing was more helpful than harmful. Many commented that being coerced into truthfulness with therapists and fellow group members was a new and positive experience that often carried over into everyday life, making them more honest with loved ones as well. Some said that if their programs stopped testing they would likely return to their former, more devious ways. However, they hoped that testing for the duration of treatment would eventually make honesty easier, perhaps even habitual. Even respondents who reported experiencing erroneous “deception indicated” (DI) ratings on one or more examinations considered polygraphy more helpful than harmful by a ratio of 4.3-to-1.

It appears then, that polygraph examinations integral to sex offender treatment help achieve the objective of increased honesty within and outside the treatment setting. One study addressed the possibility that learning openness and honesty helps offenders and their loved ones lead more fulfilling lives (Kokish et al., 2002). It found strong, albeit subjective support.

But does increased disclosure during treatment result in better post treatment outcomes? Many treatment providers assert that full disclosure of deviant sexual history is essential for effective treatment. How, they ask, can we focus on modifying behaviors we are not fully aware of? But, while this assertion has obviously attractive face validity, no controlled studies have been published comparing post-treatment recidivism between programs that use PCSOT and those that do not. Additionally, recidivism rates in Canada and Great Britain, where PCSOT is rarely used, do not appear substantially different from recidivism rates in the United States, where PCSOT is in wider use (Hanson & Bussière, 1996).

Achieving Increased Treatment and Supervision Compliance

Polygraph examinations have been used to monitor probationers and particularly child abusers since at least 1966 (Consigli, 2001). Abrams (1986) reported on 21 experimental subjects monitored by polygraph between 1983 and 1985, comparing them to 243 comparable control subjects. Sixty-nine percent of the experimental subjects successfully completed their probationary periods, compared to only 26% of controls. This result is all the more remarkable when one considers that polygraph monitoring made experimental subjects more vulnerable than control subjects to detection. The likelihood of false positives in this experiment is mitigated by the fact all Deception Indicated (DI) charts were verified by subsequent admissions. There were no known instances of unreliable "No Deception Indicated" (NDI) charts, but it is of course impossible to ascertain how many experimental subjects successfully lied to examiners.

Abrams' (1986) experimental group included 7 sex offenders. Five of them (71%) successfully completed probation, compared to only 3 of the 7 sex offender control subjects (43%). Although the small number of experimental subjects limits the value of this investigation, it does provide significant evidence for the effectiveness of polygraph monitoring as a deterrent to reoffending during supervision.

Abrams and Abrams (1993) also cite other studies reflecting impressive probation success rates when offenders (sexual and other) are monitored via polygraphy. For example, 173 polygraphed offenders supervised and treated by sex offender specialists in Jackson County, Oregon were (a) 95% free of new sex crime convictions, (b) 96% free of new felony convictions, (c) 89% free of any new criminal convictions, and (d) 65% free of parole/probation violations. Unfortunately, although local officials agreed that failure rates would have been considerably higher without polygraphy, the data is weakened by the absence of a control group.

According to Abrams (1993), the Oregon Polygraph Licensing Advisory Committee held two days of public hearings in July of 1990. The majority of those who spoke about sex offender testing were therapists, probation officers, and judges. They almost unanimously agreed about the value of PCSOT and were anxious to avoid proposed restrictions. Altogether, available evidence suggests that polygraph monitoring is an effective way to maximize supervision and treatment compliance while minimizing reoffense risk with sex offenders.

Accuracy Estimates

That polygraph examinations can generate increased disclosures is not surprising to social scientists. Roese and Jamieson (1993) summarized 20 years of research indicating that a mere belief that questioners have an independent “pipeline” to the truth is an effective way to obtain information that might not be offered up in the absence of that belief. However, belief that polygraph examiners can ascertain truth is obviously not synonymous with an ability to do so. The actual degree of that ability has been hotly debated for nearly a century. As recently as 2001, Iacono (2001) maintained that Comparison Questions Tests (CQT)² have no scientific validity. However, Raskin and Honts (2001) argue that CQT sensitivity and specificity likely exceed 90% when charts are scored by examiners who administer the examinations. They also assert that sensitivity is unchanged while specificity remains respectable at 75% when independent examiners score the same charts.

Examining literature relevant to polygraph use by National Security Agencies, the National Research Council of the National Academies (2002) recently concluded, “...in populations of examinees such as those represented in the polygraph literature, untrained in counter-measures, specific incidents polygraph tests can discriminate lying from truth telling at rates well above chance, though well below perfection” (p. 4).

For PCSOT specifically, the scientific debate is just beginning. Since PCSOT utilizes virtually no test format other than CQT (Abrams & Abrams, 1993; Blasingame, 1998; Consigli, 2001), the following discussion is limited to the CQT format.

Meaningful consideration of polygraph accuracy should address at least the following issues.

Laboratory Versus Field Data

Laboratory studies have the advantage of being able to control conditions so that ground truth is known with absolute certainty. However, since polygraph testing attempts to measure physiological reactions to a psychological phenomenon (fear of being caught in a lie), these studies are often criticized on the basis that laboratory conditions cannot simulate the psychological milieu of real-life testing.

Field studies have the advantage of looking at tests administered under real-life psychological conditions. However, since they look at tests given in cases where ground truth was unknown at time of testing, it is difficult to establish valid criteria for ground truth against which to compare examination results.

All Examiners Are Not Alike

Raskin, Barland, and Podlesny (in Blasingame, 1998) for example, reported a study wherein one examiner was accurate 53% of the time while fellow examiners with the same training were accurate 100% of the time. Many studies agree that examiners who score charts “blindly” (i.e., without interviewing the examinee and administering the actual examination) are usually less accurate than the examiner who administered the test (Raskin & Honts, 2001). This seems to underscore the fact that all examiners are not the same and that “art” may still play a significant role in polygraph accuracy.

All Examinees Are Not Alike

There is general agreement that actively psychotic people are inappropriate polygraph subjects and that test accuracy drops as examinee IQ goes into and below the borderline range (Blasingame, 1998). Guilt complex responding,³ although an apparently rare phenomenon, cannot always be discounted. During a 10-year period, the present author encountered two suspected guilt-complex responders and verified both cases with guilt complex tests (Abrams, 1989a).

All Tests Are Not Alike

Federal law no longer permits private sector pre-employment testing because of its alleged inaccuracy and the serious consequences (i.e., not getting the job) that usually follow from DI charts. Yet, even the most severe polygraph critics like David Lykken acknowledge high accuracy for Guilty Knowledge Tests⁴ (Lykken, 1981).

There are four kinds of sex offender tests (California Association of Polygraph Examiners [CAPE], 2001):

1. Specific issue examinations inquire into whether a particular act was committed at a particular time.
2. History examinations inquire into whether the examinee has fully disclosed his/her sexual and/or criminal history to the examiner and/or treatment program.
3. Maintenance examinations inquire into supervision, treatment, or safety plan violations.
4. Monitoring examinations inquire into the commission of unidentified sexual offenses while under supervision or since entering treatment.

Monitoring and maintenance issues have historically been addressed in a single examination (Blasingame, 1998; Consigli, 2001). However, some experts now think that

separating those issues will yield more accurate results (K. English, personal communication, November 29, 2001) and some professional organizations are incorporating this belief into their PCSOT standards (CAPE, 2001).

There is general agreement that narrow questions inquiring about specific acts at specific times produce the most accurate test charts. These narrow tests are suitable investigative tools. However, sex offender treatment programs (SOTPs) use polygraphy to obtain information for treatment leverage. Hence, there is a tendency to ask broad questions (e.g., “Since your 18th birthday, have you sexually touched anyone under the age of 18 other than...?”). While generating more information, broad questions can reduce chart accuracy.

Base Rates of Truthfulness and Attempted Deception

Professionals and examinees have distinct though related concerns about test accuracy. Professionals want to know the probability of a particular chart representing that examinee’s truthfulness (or lack of it) regarding the relevant questions (i.e., chart accuracy). Examinees already know how well their chart rating matches their truthfulness. What they want to know is the probability of their truthful or deceptive answers being accurately assessed by the examiner on each examination they are required to undergo (i.e., test “sensitivity” and “specificity”). These two sets of probabilities differ from one another. The former varies with base rates of truth telling and deception in the population being examined; the latter does not.

The Federal Government, which often relies on polygraphy for security and intelligence purposes, asked the Office of Technology Assessment (OTA) to conduct a major inquiry into polygraph accuracy and practices (Saxe, Dougherty, & Cross, 1983). OTA investigators found 10 CQT field studies with sufficient scientific merit to warrant inclusion in their analysis. Correct detection of guilt (sensitivity) ranged from 70.6% to 98.6% and averaged 86.3% when

inconclusive results were counted as errors. Correct detection of innocence (specificity) ranged from 12.5% to 94.1% and averaged 76% when inconclusive results were counted as errors. The OTA study elected to count inconclusive results as errors on grounds that a guilty or innocent person had not been correctly identified.

However, many experts do not consider inconclusive charts erroneous, arguing that an error is a mistaken belief while an inconclusive result is a correct belief about the limitation of one's knowledge. Raskin (1988) discarded inconclusive results from OTA's data. He calculated 90% accuracy for detection of guilt and 80% for correct identification of innocent subjects. There have been technological advances since the OTA report. Some more recent reports claim that correct identification of both guilty and innocent examinees exceeds 95% (Forensic Research, 1997; Raskin & Honts, 2001). Nonetheless, the OTA report is still cited as a conservative and reasonable estimate of CQT accuracy.

Laboratory studies included in the OTA report were less accurate. Counting inconclusive results as errors, sensitivity ranged from 35.4% to 100% (mean average 63.7%) and specificity from 32% to 91% (mean average 57.9%). However, when Raskin and Honts (2001) looked at nine laboratory studies conducted between 1978 and 1997, omitting inconclusive charts they arrived at average overall accuracy of 90%. None of these nine studies detected innocent subjects with less than 75% accuracy, although one study only detected guilt at the 53% level. Had Raskin and Honts omitted this "outlier," accuracy would have been even higher.

Altogether, it is fair to say there remains considerable disagreement in the scientific community about the accuracy of CQT polygraph examinations, OTA accuracy estimates (90% / 80%) remain approximately midway between the most pessimistic and most optimistic scientific opinions. If SOTP's operate from a working hypothesis that OTA rates are realistic, they appear

to have a very good clinical tool; indeed, many respected medical tests are not so accurate.

Unfortunately, the matter is not as simple as it first appears, because base rates of truthfulness and deception still have to be accounted for. Consider the following scenario:

An SOTP persuades its participants to answer examiners truthfully on four of every five examinations and tests 100 participants. According to Raskin's OTA projections, 80% (64) of the 80 truthful examinees will be correctly identified and 10% (4) deceptive examinees will be misidentified as truthful, yielding 68 NDI charts. From the program's point of view, over 94% of its truthful charts (64/68) will be accurate and 9 of every 10 lies will be detected. Unfortunately, 16 truthful examinees will be misidentified as deceptive, producing, in combination with 4 correct assessments of deception, 20 DI charts.

Thus, when the base rate of deception is relatively low (10%), many if not most examinees accused of deception will actually have been truthful. Paradoxically then, the more successful a program is at persuading its participants to be truthful, the higher its false positive rate becomes. To make matters worse, all studies included in the OTA report were based on specific issue examinations about real or mock crimes. Most PCSOT examinations focus on sexual histories or attempt to monitor a range of behaviors over time (e.g., "Since your last examination, have you...?"). Such examinations can be more akin to "screening tests," a format that is considered less accurate than the specific issue examinations typically administered in the criminal justice system (Krapohl, 2001; Saxe et al., 1983).

We have only one scientific study that examined PCSOT accuracy specifically (Kokish et al., 2002). The study asked participants in three SOTP's to complete an anonymous questionnaire that included questions about accuracy. The 95 study participants had taken 333

tests. They reported being truthful on 287 with examiners correctly identifying 265 and misclassifying 22 as DI. They reported attempting deception on 39 examinations, with examiners correctly detecting 28 and misclassifying 11 as NDI. This yields the following accuracy rates:

1. Correct identification of truthful examinees: 92.3% (265/287)
2. Correct identification of deceptive examinees: 71.8% (28/39)
3. NDI chart accuracy: 96.3% (287/298)
4. DI chart accuracy: 66.6% (11/33)

Interestingly, this study differs from most Criminal Justice and Intelligence studies in that reported specificity exceeded sensitivity. This may be a function of repeatedly testing the same group of examinees. The study is of course open to criticism in that offender self-report is, on its face, a poor way to establish ground truth. Nonetheless, the information was carefully collected to insure anonymity and the accuracy claims are not very far from OTA accuracy rates. Only 12 of these 95 offenders had been in treatment less than a year, and mean time in treatment was nearly 30 months ($SD = 18.12$ months). We would expect low rates of attempted deception from examinees immersed in a treatment culture that repeatedly reinforces the value of compliance and honesty while quickly sanctioning non-compliance and attempted deception. This expectation is supported by the fact that 85% of charts produced by examinees in these programs were rated NDI by examiners. This is very close to the 90% truthfulness claim from the anonymous respondents themselves. Finally, study credibility is further enhanced in that over 81% of respondents reporting experience with incorrect DI ratings nonetheless said periodic polygraph testing had a positive effect on their therapy and should be continued.

Altogether, available scientific data seem to suggest that SOTP's can have high but not absolute confidence in NDI charts. However, DI charts are more problematic and should be given less weight when case management decisions need to be made.

Critics caution that the coercive nature of polygraph testing may generate false confessions (Cross & Saxe, 1992, 2001). Kokish and colleagues (2002) appear to have conducted the only study that examined this question empirically. Five of 22 individuals (23%) who anonymously reported experiencing incorrect DI ratings said they reacted by admitting things they had not done. Four of these 5 individuals reported passing their next examination. It seems then that concern over coercing fictitious confessions bears further investigation.

Polygraph in Context

Although the term "therapy" is often used, mental health work with sex offenders is distinctly different from traditional psychotherapy. Traditionally, the patient-psychotherapist relationship is voluntary and patient oriented, in that patients take initiative to seek help regarding subjective distress. Sex Offender Treatment (SOT) is more community oriented. Community agents refer offenders because the community perceives the offenders' criminal behavior as a source of distress for others. Traditional patients are free to discontinue treatment at any time. They have discretion to decline certain procedures, and therapists are primarily accountable to their patients. Community agents, however, require sex offenders to remain in therapy regardless of their wishes and may require them to complete procedures they dislike, because the whole purpose of sex offender therapy is to modify "patient" behavior outside the therapy room. This occurs whether or not the patients are in agreement with the desired modifications.

Sex offender therapists are accountable to their patients for competence and ethical behavior, but are accountable to the community for treatment results. Thus, SOT is, in a word, coercive. Given all this, one could even argue that the term “therapy” is not appropriate for sex offender work. The term “treatment,” however, is certainly appropriate when defined simply as acting on something for the purpose of change (e.g., one can “treat” steel to harden or cloth to change its ability to resist stains, etc.).

There is wide consensus in the field that SOT success is improved when change efforts are made, not by a single treatment agent but by a team that has strong community support. This approach has come to be known as the “containment model” (California Coalition on Sexual Offending [CCOSO], 2001). English and Jones (1998) conceptualized Containment Teams as a metaphorical triangle consisting of supervision agents, therapists, and polygraph examiners. Team members engage in a variety of coordinated activities intended to achieve treatment goals and have distinct though overlapping roles. They share information as needed to support one another in fulfilling these roles. They consult regularly to assess progress, identify problems, plan treatment strategies, and make major case decisions.

Polygraphy is rooted in investigative work, but containment based SOTP’s use it as a clinical tool for obtaining diagnostic information and to monitor and shape offender behavior outside the treatment setting. Whereas opponents criticize polygraphy for its coercive nature (Cross & Saxe, 1992, 2001), proponents working within a containment framework embrace it for that very reason (English, Jones, Pasini-Hill et al., 2000; English, Jones, Patrick, Pasini-Hill, & Gonzalez, 2000). However, given its coercive nature and still uncertain accuracy, a majority of proponents recognize certain ethical concerns, especially protection from self-incrimination and appropriate responses to DI charts.

Blasingame (1998) addressed these concerns with a list of 16 recommendations for ethical and effective use. The American Polygraph Association (APA) sets national standards for sex offender testing (APA, 1995), some state organizations have followed (CAPE, 2001), and ATSA includes a section on polygraph in its practice standards and guidelines (ATSA, 1997, 2001). The various documents are essentially consistent with one another and generally address the following issues:

1. Maximizing accuracy and reliability by adhering to explicit standards for (a) examiner training and experience; (b) instrumentation and instrument calibration; (c) pre and post test interview procedures and formats; (d) chart scoring and interpretation; (e) examinee fitness; and (f) test format, environment, and frequency.
2. Appropriate and inappropriate application for making case decisions.
3. Protecting the treatment setting and examinee by (a) proper advisements and informed consents; (b) accurate record keeping and secure record storage; (c) limiting privilege and confidentiality in regard to new crimes and program violations; and (d) careful management of disclosures related to criminal acts committed before entering treatment.

The last item (3d) warrants further discussion. SOTP's use polygraph examinations for persuading examinees to disclose information about their behavior. Since allowing offenders to break rules and perhaps even reoffend with impunity would make little sense, there is general agreement that self-disclosure of illicit behavior during supervision and treatment will not remain confidential or privileged (e.g., non-approved contact with a minor). Programs advise participants accordingly, but ask them to disclose illicit behavior anyway. Participants are told that failed examinations will lead to increased surveillance and restrictions as well as thorough investigations, making discovery of illicit behavior quite likely. However, consequences may be

mitigated if offenders self-disclose rather than waiting to be independently discovered. This argument can be persuasive within a carefully managed treatment culture. The present author has seen offenders self-report many program violations and even some felonious acts. Additionally, the author has seen several offenders report violations committed by other participants because they were afraid keeping the knowledge to themselves would weigh on their minds, causing them to fail their next examination.

Reporting of pre-treatment crimes is another matter. As we saw earlier in this chapter, what sex offenders are prosecuted for rarely represents the true depth of their deviance or full extent of their criminal behavior. Polygraphy's usefulness as a clinical tool derives from its ability to elicit historical information, presumably allowing psycho-behavioral patterns to be more fully uncovered, better understood, and more effectively managed and changed. However, officially reporting incriminating historical information usually leads to additional prosecution, often ending the very treatment it was designed to enhance. This paradox assumes critical importance because all states have laws requiring therapists to report child abuse to law enforcement.

One way of resolving the dilemma is for local prosecutors to enter into various kinds of immunity agreements with the sex offenders being treated in their jurisdictions (English, Jones, Pasini-Hill et al., 2000). Typically, these immunity agreements stay in effect only so long as offenders comply with supervision and treatment conditions. If conditional release is revoked or, in the prison setting, if the offender is unsuccessfully discharged from treatment, the information can be used as a basis for investigation leading to prosecution.

When no immunity agreement is in place, programs can collect historical information, sans details that identify specific victims. Some programs prefer this strategy, because they

believe it is the most effective way of providing the confidentiality necessary for effective treatment. Others prefer the immunity route because it allows for treatment outreach to unreported victims and because they believe encouraging offenders to conceal anything at all supports victim objectification (English, Jones, Pasini-Hill et al., 2000). Some examiners are also concerned that directing offenders to withhold incriminating information could affect test accuracy (K. English, personal communication, November 29, 2001), although this is an empirical issue that should be amenable to scientific inquiry.

When it comes to PCSOT, issues of accuracy and ethics are inseparably intertwined. Many PCSOT proponents believe that adhering to standards will increase accuracy, thereby mitigating ethical dilemmas around managing examinees who produce DI charts. Staunch critics, however, believe that polygraphy is inherently flawed (Oksol & O'Donahue, 2001). They point out that in spite of efforts spanning more than a century, accuracy is still not firmly established. These critics fear that setting standards for such a flawed procedure may yield only a more powerful illusion of scientific credibility and argue that under such circumstances it may be better to avoid polygraphy altogether (Cross & Saxe, 1992, 2001).

Where Are We Now, Where Are We Going?

Available evidence suggests that PCSOT application is growing in the United States, with between 30-40% of existing SOTP's presently using it. User satisfaction appears high. ATSA sponsors a very active professional email discussion list for credentialed SOT professionals. The list has over 700 subscribers, has been in existence since 1998, and often averages between 10 and 20 posts per day. The present author sent an inquiry to that list asking whether any members who had, at some point, integrated polygraphy into their treatment work became dissatisfied with the procedure and abandoned it. No one responded affirmatively. The

one relevant study to date suggests that offenders are nearly as satisfied as professionals (Kokish et al., 2002). However, many issues and questions remain.

Polygraph utility for eliciting information has been studied and is well documented, though Kokish and colleagues (2002) suggest that eliciting false confessions may be a problem in some programs. Accuracy has also been widely studied, but both laboratory and field studies can be criticized, and firmly established accuracy rates remain elusive even for the criminal justice setting. Given psychological theory underlying polygraph practice, it seems likely that accuracy is somewhat context dependent and has simply not been scientifically studied in the PCSOT context. For that context, we have only one study submitted for peer review and nothing published to date. PCSOT-specific research is needed to maximize precision and help resolve ethical dilemmas around managing reportedly deceptive offenders in circumstances where base rates may combine with accuracy limitations to produce high false positive rates.

For the time being then, SOTP's are faced with a series of dilemmas that will have to be endured. Those that choose to incorporate polygraphy will gain much pertinent information about their offender-patients. Arguably, this information can make treatment more effective, helping lower post-treatment reoffense rates and perhaps even contributing to improved quality of life for the offenders and their families. Available information also suggests that polygraph monitoring increases program compliance and serves as a deterrent to reoffending during treatment. But there is no empirical evidence thus far that PCSOT reduces post-treatment recidivism, and programs that employ PCSOT risk generating false admissions and wrongfully sanctioning substantial numbers of program participants whose charts are incorrectly rated as being consistent with attempted deception.

Programs that decline polygraphy will avoid these ethical risks, but will almost certainly have to function with less information about their offender patients. Consequently, they will likely assume a higher risk of undetected program violations and reoffenses during treatment, and perhaps higher rates of post-treatment reoffending. In such circumstances, it falls to each SOTP to determine which advantages and liabilities to choose.

Kohlberg (in Weinrich-Haste, 1983) theorized that moral maturity is characterized by a person's ability to make ethical decisions relevant to context and to recognize and resolve moral dilemmas. It may be then, that those SOT providers who are most uncomfortable with their decision to use or not use polygraphy are modeling high levels of moral maturity for program participants. And it may be that their uncertainty is helping move the field to an ultimately satisfactory, though still elusive, resolution of the dilemmas PCSOT poses for sex offender therapists, supervision officers, examiners, and SOTP participants. Until these dilemmas are resolved, this author offers the following recommendations for ethical PCSOT use to those who choose to include PCSOT in their treatment repertoire:

1. Polygraph efficacy is maximized when examinees experience appropriate positive and negative consequences as a result of the charts they produce. It is advisable to share responsibility for such consequences across several disciplines. Therefore, treatment providers should work as members of Containment Teams (English, 1998).
2. Treatment providers should work with examiners who adhere to training and practice guidelines established by the American Polygraph Association and other professional polygraph organizations.

3. All crimes and rule violations committed during treatment, including but not limited to new crimes, should be immediately reported to appropriate officials. Patients should be informed in writing before beginning treatment that such reports will be made.
4. It is unreasonable to expect offenders to fully disclose as disclosure is likely to lead to additional prosecutions. Therefore, patients should be protected from prosecution for self-disclosed crimes committed prior to beginning treatment. Written use immunity agreements with prosecutors, not collecting victim identities, and legal consultation to patients regarding Fifth Amendment rights can contribute to appropriate protections for patients without unduly jeopardizing community safety.
5. Polygraphy should not be the only form of monitoring used by a treatment team. Other methods such as electronic surveillance, collateral contacts, chemical testing, and unannounced field visits should be employed on a regular basis.
6. NDI charts appear to be highly reliable. Together with other measures of patient progress, NDI charts can play an important role in supporting case management decisions.
7. DI charts do not appear to be as reliable at this time. Their primary value is to highlight a need for cautious case management and to trigger independent investigations of alleged facts and perceived risks. DI charts alone should never be the primary basis for case management decisions that could have significant adverse effects on a patient and/or her/his family.
8. Neither NDI nor DI charts should be used as the sole basis for decisions requiring court action (e.g., probation violations, early termination of probationary status, family reunification).

9. Some patients produce less reliable charts than others. Particular caution is warranted when testing patients who (a) are 13 years of age or younger; (b) manifest impaired reality testing; (c) take medications known to effect the physiological responses on which polygraphy relies; (d) appear unable to produce DI charts even when independent information makes it highly unlikely they are being deceptive; and/or (e) have an IQ less than 80.
10. Polygraph, corrections, and psychotherapy professionals should actively cooperate in joint research and other ventures to enhance PSCOT accuracy, utility and ethical practice.

References

- Abrams, S. (1986). Polygraph surveillance of probationers. *Polygraph*, 13(3), 174-182.
- Abrams, S. (1989a). *The complete polygraph handbook*. Lexington, KY: Lexington Books.
- Abrams, S. (1989b). Probation polygraph surveillance of child abusers. *The Prosecutor*, 22(3), 29-36.
- Abrams, S., & Abrams, J. B. (1993). *Polygraph testing of the pedophile* (1st ed.). Portland OR: Ryan Gwynner Press.
- Abrams, S., & Simmons, G. (2000). Post-conviction polygraph testing: Then and now. *Polygraph*, 29(1), 63-67.
- Alexander, M. A. (1999). Sexual offender treatment efficacy revisited. *Sexual Abuse: A Journal of Research & Treatment*, 11(2), 101-116.
- American Polygraph Association (1995). *Standards for post conviction sex offender testing: Practice guidelines*. Chattanooga, TN: Author.
- Association for the Treatment of Sexual Abusers. (1997). *Ethical standards and principles for the management of sexual abusers*. Beaverton, OR: Author.
- Association for the Treatment of Sexual Abusers. (2001). *Practice standards and guidelines for members of the association for the treatment of sexual abusers*. Beaverton OR: Author.
- Blasingame, G. (1998). Suggested clinical uses of polygraphy in community-based sexual offender treatment programs. *Sexual Abuse: A Journal of Research and Treatment*, 10(1), 37-45.
- Burton, D. L., & Smith-Darden, J. P. (2000). *1996 nationwide survey of treatment programs and models*. Brandon, VT: Safer Society Foundation.

California Association of Polygraph Examiners. (2001). *CAPE guidelines for clinical polygraph examination of sex offenders: Post conviction*: Redlands, California: Author.

California Coalition on Sexual Offending. (2001). *Effective management of sex offenders residing in open communities*. Sacramento, CA: Author.

Consigli, J. (2001). Post-conviction sex offender testing and the American polygraph association. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 237-250). New York: Academic Press.

Comment [JZ1]: Insert page numbers of chapter.

Cross, T. P., & Saxe, L. (1992). A critique of the validity of polygraph testing in child sexual abuse cases. *Journal of Child Sexual Abuse*, 1(4), 19-33.

Cross, T. P., & Saxe, L. (2001). Polygraph testing and sexual abuse: The lure of the magic lasso. *Child Maltreatment*, 6(3), 195-206.

English, K. (1998). The containment approach: An aggressive strategy for the community management of adult sex offenders. *Psychology, Public Policy and Law*, 4(1/2), 218-235.

English, K., Jones, L., Pasini-Hill, D., Patrick, D., & Cooley-Towell, S. (2000). *The value of polygraph testing in sex offender management* (Research Report Submitted to the National Institute of Justice No. D97LBVX0034). Denver, CO: Department of Public Safety, Division of Criminal Justice, Office of Research and Statistics.

English, K., Jones, L., Patrick, D., Pasini-Hill, D., & Gonzalez, S. (2000). We need you to become experts in post-conviction polygraph. *Polygraph*, 29(1), 44-62.

English, K.S.P., & Jones, L. (1998). The containment approach: An aggressive strategy for community management of adult sex offenders. *Psychology, Public Policy, and Law*, 4(1 / 2), 218-235.

Comment [JZ2]: Need page numbers of article.

Forensic Research. (1997). *The validity and reliability of polygraph testing* (industry Saturday).

Severna Park, MD: American Polygraph Association.

Grossman, L. S., Martis, B., & Fichtner, C. G. (1999). Are sex offenders treatable? A research overview *Psychiatr Services*, 50(3), 349-361.

Comment [JZ3]: Please list complete title of journal.

Hall, G. C. N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting & Clinical Psychology*, 63(5), 802-809.

Hanson, K. R., & Bussière, M. T. (1996). *Predictors of sexual offender recidivism: A meta-analysis* (No. Cat. No. JS4-1/1996-4E). Ottawa: Public Works and Government Services Canada, Department of the Solicitor General of Canada.

Hanson, K. R., Thornton, D., Gordon, A., Harris, A. J. R., Marques, J., 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(2), 169-194.

Hassett, J. (1978). *A primer of psychophysiology*. San Francisco: W. H. Freeman and Company.

Hindman, J., & Peters, J. M. (in press). How polygraphing adult sex offenders can lead us to better understanding the juvenile sex offender. *Federal Probation*.

Iacono, W. G. (2001). Forensic "lie detection": Procedures without scientific basis. *Journal of Forensic Psychology Practice*, 1(1), 75-86.

Kokish, R., Blasingame, G., & Plaud, J. (2002). An exploration of the accuracy, benefits and detriments of post-conviction sex-offender polygraph testing using offender self-report. Manuscript in progress.

Krapohl, D. J. (2001). The polygraph in personnel screening. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 217-236). New York: Academic Press.

Comment [JZ4]: Insert page numbers of chapter.

Lykken, D. (1981). *A tremor in the blood*. New York: McGraw-Hill.

Oksol, E. M., & O'Donahue, W. T. (2001). *A critical analysis of the polygraph*. Unpublished manuscript.

National Research Council of The National Academies; Division of Behavioral and Social Sciences and Education. (2002). *The polygraph and lie detection*. Washington, DC: Author.

Raskin, D. C. (1988). Does science support polygraph testing? In A. Gale (Ed.), *The polygraph test: Lies, truth and science* (pp. 96-110). Thousand Oaks, CA: Sage Publications.

Raskin, D. C., & Honts, C. R. (2001). The comparison question test. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 1-47). New York: Academic Press.

Comment [JZ5]: Insert page numbers of chapter.

Roese, N. J., & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. *Psychological Bulletin*, 114(2), 363-375.

Saxe, L., Dougherty, D., & Cross, T. (1983). *Scientific validity of polygraph testing: A research review and evaluation: A technical memorandum* (No. OTA-TM-H-15). Washington, DC: U.S. Congress Office of Technology Assessment.

Weinrich-Haste, H. (1983). Moral development. In R. Harre & R. Lamb (Eds.), *The encyclopedic dictionary of psychology* (pp. 398-400). Cambridge, MA: The MIT Press.

Comment [JZ6]: Insert page numbers of chapter.

Author Note

PO Box 476, Trinidad, CA 95570

ron@delko.net

Comment [JZ7]: Insert address to which correspondence may be sent. Please do not include a home/private address as this will be published.

Ron Kokish is a Licensed Family Therapist and Board Certified Clinical Social Worker with 17 years of experience in (government operated) Child Protective Services. He was in private practice from 1987 through July 2000, treating victims and perpetrators of sexual, physical and emotional abuse, preparing forensic evaluations, serving as an expert witness, and training social workers, psychologists, probation and parole officers, teachers and counselors through various accredited human services programs. He has been using polygraph in treatment since 1987. His present practice specializes in consultation, training and evaluations

Footnotes

¹ For example, bestiality, giving victim alcohol or drugs, drug abuse on offender's part, urination with sexual act, etc.

² There are various types of comparison question tests but all ask "relevant" and some form of "comparison" questions. Relevant questions are straightforward, focusing on what the examiner actually wants to find out. Comparison questions are about the same general subject as relevant questions (e.g. sex) but not about specific events relevant to the examination (e.g. whether a particular child's genitals were fondled or how many children the examinee victimized during the previous 5 years). Truthfulness or deception is inferred from numerical comparison of measured physiological responses to the two types of questions.

³ A "guilt complex responder" (misnamed, because the phenomenon relates less to guilt than anxiety) is an individual who cannot pass a test because of the extreme anxiety generated by the mere question as to whether she/he committed a particular act. A "guilt complex test" utilizes an issue that holds some threat to the examinee while the examiner knows ground truth in advance of the examination. If the truthful subject produces DI charts, he/she is presumed to be a guilt complex responder.

⁴ A test limited to situations in which only authorities and the guilty individual are likely to have knowledge of certain facts (e.g., the type of gun used in a crime, the number of times a victim was shot, etc).