Psychophysiological Assessment of Sexual Offenders: A Practitioner's Perspective

Wesley B. Maram

Psychophysiological assessment measures have been used in the assessment and management of sexual offenders for over half a century. Despite the large body of research as well as their extensive use in treatment programs (both residential and community-based) and the supervision for sexual offenders, critics continue to raise various issues about the nature of the findings and their application in clinical, forensic, and broader management settings for offenders. However, among those who use physiological assessment there is a recognized family of procedures with common aims that offer general guidance to effectively collect and interpret the results that provides a reasonable justification for the continued use of such measures (O'Donohue & Letourneau, 1992). Thus, the Association for the Treatment of Sexual Aggressive Abusers (ATSA, 2001, 2013) included several psychophysiological assessment measures in their Practice Standards. ATSA notes that it is recognized that psychophysiological assessment methods such as phallometry, viewing time, and polygraphy may have particular usefulness to (a) obtain objective behavioral data about an individual that may not be readily established through other assessment means, (b) explore the reliability of individual self-report, and (c) explore potential changes, progress, and/or compliance relative to treatment and other case management goals and objectives.

This chapter represents an effort to gather and organize the available literature on the procedural steps of administrating the Penile Plethysmograph along with other psychophysiological measures. However, this author departs from some practitioners regarding how to interpret findings in both clinical settings (in which the goal is to identify appropriate treatment targets for therapy) and forensic settings (when the purpose is to elicit information regarding the presences or absence of sexual deviance). It is posited that in clinical practice it is of greater importance to identify the potential presence of sexual deviance among known child molesters than it is to adhere to rigid interpretative guidelines. Having just one positive indicator or evidence of significant arousal to a sexually deviant stimulus provides valuable, clinically important information that should not be ignored even when the individual shows greater arousal to nondeviant stimuli. Important clinical information can be obtained about an individual who was aroused by a single image or deviant story about sex with a child, sexual violence, or nonsexual violence. This chapter includes discussions about and suggestions on how to apply viewing time measures and polygraph examinations as well as Penile Plethysmograph results in clinical and evaluating settings to increase and accelerate client disclosure of sexual deviance. Issues and concerns related to admissibility of this information in the courtroom are also discussed.

It is a formidable but critical process to attempt to determine what factors motivate a person to sexually molest a child, or rape a child, adolescent, or adult and then develop an effective intervention strategy to prevent future sexual abuse from occurring. The effort is complicated by the fact that there is a lack of uniform consensus on how to define the diagnostic criteria for sexual offending against children or nonconsenting individuals. Some even question the wisdom of classifying people with a Pedophilia, a Paraphilic Disorder-Not Otherwise Specified (NOS) (e.g., Coercive Paraphilic Disorder), or Sexual Sadism as mental conditions at all because historically sexual behavior with children and nonconsenting persons has been common (Green, 2002; Quinsey, 2010). Even if we settle on the diagnostic definition provided by the Diagnostic and Statistical Manual, Fourth Edition, Text Revision (DSM-IV-TR, 2000) which reads, "The paraphilic focus of Pedophilia involves sexual activity with a prepubescent child generally age 13 years or younger (p. 571)," different evaluators too often do not come to the same diagnostic conclusion (e.g., Levenson, 2004; O'Donohue, Regev, & Hagstrom, 2000; Wollert, 2007). The DSM-5 does not provide any additional clarity since the diagnostic criteria of Pedophilia remains the same. The only revision made was to change the name of the disorder from Pedophilia to Pedophilic Disorder. Also, efforts to establish

W.B. Maram (✉)
1234W. Chapman, Ste 203, Orange, CA 92868, USA
e-mail: drmaram@orangepsych.com

A. Phenix, H.M. Hoberman (eds.), *Sexual Offending*, DOI 10.1007/978-1-4939-2416-5_15

a new diagnosis of Paraphilia Coercive Sexual Disorder (to replace and improve upon the Paraphilia NOS nonconsenting partner diagnosis), referring to males who are sexually aroused by the coercive elements of rape, were debated and subsequently not included as specific disorder in the DSM-5 (DSM-5 Development, 2013; Knight, 2009). However, paraphilic disorders relating to sexual arousal, urges, and/or behavior to nonconsenting persons can still be categorized under DSM-5 as a Paraphilic Disorder-Not Otherwise Specified (NOS) (e.g., Coercive Paraphilic Disorder) or under the expanded definition of Sexual Sadism.

Problems with diagnostic reliability are likely to continue. Reliability is weakened when the evaluator who frequently only has at hand criminal records and the individual's self-report (where the former is too often a superficial behavioral description completely lacking details about internal motivation, and the latter is overly influenced by the desire for self-preservation). There is a common expectation that the accused individual will not be forthright about dimensions of their sexual deviance. Relying on the offender's self-report alone is problematic. This point was recognized by the DSM-V Paraphilia Workgroup (2013). They wrote, "…the fact that a substantial proportion—perhaps a majority—of patients referred for assessment for paraphilias is referred after committing a criminal sexual offense. Such patients are not reliable historians, and they are typically not candid about their sexual urges and fantasies." (pg. 1)

Official records list detected criminal events and may also include behavioral descriptions of criminal incidents. Information regarding the individual's fantasies and urges is generally discovered through clinical interview. Because of the negative societal repercussions associated with child molestation or sexual violence, there is a natural reluctance by the sexual offender to engage honestly in public self-examination of deviant fantasies and urges toward children to evaluators who are going to include this information in reports to the courts and other elements of a the legal system that will make decisions about their freedom.

A variety of strategies have been developed and employed to elicit the psychological factors of an individual related to his sexual offending, including using positive interview techniques that support the individual and avoid "shaming" the person for what they have done. In addition, standardized sex history questionnaires can extract information that might not otherwise be directly disclosed in a face-to-face interview. For example, questionnaires such as the Psychosexual History Questionnaire© (Nichols & Molinder, 1999) pose questions about the person's sexual experiences, fantasies, and urges in the context of gathering background information about the individual in a manner that has the potential to be objective, clinical, and less threatening than having to respond to a clinician who is asking specific questions to the person about their deviant sexual behaviors, fantasies, and urges. Even with the aid of structured questionnaires and skillful interviewing, an evaluator can rarely have confidence that the sex offender has made a full disclosure. Issues associated with subjective offender self-report and clinical judgment in making reliable and valid diagnosis suggest that objective measures might offer potential utility as another source of information (Bradford, Kingston, Ahmed, & Fedoroff, 2010). Sexual arousal testing can help by offering one means of identifying the possible presence of deviant sexual interests. Underscoring the importance of this information, meta-analysis studies (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2005; Mann, Hanson, & Thornton, 2010) demonstrated that deviant sexual interest in children is among the strongest predictors of sexual recidivism.

PPG, VRT, and Polygraph

Psychophysiological procedures have been developed that offer the advantage of providing a standardized and often more objective measure to help differentiate the situational offender from the persistent offender with preferential deviant sexual interests. These psychophysiological measures to be discussed in this chapter include the Polygraph in conjunction with Viewing Time, also referred to as Visual Reaction Time (VRT) assessments, and the Penile Plethysmograph (PPG), often referred to as phallometric assessment.

The Polygraph Sexual History Exam attempts to gain an individual's full disclosure of the extent and variety of their criminal, deviant, and nondeviant sexual behaviors using skillful interviewing in combination with measures of involuntary physiological reactions to stress, by monitoring patterns of blood pressure, heart rate, respiration, and galvanic skin response. VRT assessment [such as the *Abel Assessment for Sexual Interest*™ (AASI)] combines measures of self-reported sexual arousal with standard measures of how long the individual looks at slide images of persons of different age and gender categories (e.g., viewing time) along with a self-report sexual history questionnaire (Abel, Huffman, Warberg, & Holland, 1998). The PPG measures sexual arousal by changes in penile engorgement or tumescence (e.g., relative blood flow into the penis) while the individual is viewing clothed or nude images of children, adolescents, and adults; listening to audio recorded sexual stories involving children and sexual and nonsexual violence; or viewing and listening to these stimulus presentations. When these psychophysiological tools are used in tandem with other measures related to potential deviant sexual interests, the evaluator/clinician often obtains a more clear and comprehensive overview of the individual's sexual interests,

arousals, and behavioral history (Maram & Koetting, 2004). The availability of this information from these particular procedures can be used to create an opportunity for the individual to become more willing to disclose their inner motivation for deviant behavior.

Drawbacks and Advantages

As with any measurement approach, each of these psychophysiological measures has relative limitations as well as relative advantages. For example, those critical of the polygraph point to studies that suggest the polygraph lacks scientific validity (e.g., U.S. Congress Office of Technology Assessment, 1983). In contrast, polygraph supporters point to the polygraph's empirically demonstrated utility in post-conviction assessment of sex offenders (English, Jones, Pasini-Hill, Patrick, & Cooley-Towell, 2000; Grubin & Madson, 2006; Kokish, Levenson, & Blasingame, 2005; Raskin, 1988; Raskin, Barland, & Podlesny, 1976). The AASI has been criticized as being inaccurate in distinguishing child molesters from non-molesters (e.g., Fischer & Smith, 1999). Others have reported that the AASI results provide good discrimination between child molesters and non-child molesters (Card & Dibble, 1995; Letourneau, 2002).

Critics have also raised questions about the methodology and value of the PPG. These range from its lack of standardization and potential ethical concerns to questions as to whether penile engorgement is a reliable indicator of sexual arousal (Konopasky & Konopasky, 2000; Schouten & Simon, 1992). Alternately, a number of authorities advocate that the most well-established method for assessing positive evidence of sexual interest or arousal remains the Penile Plethysmograph (O'Donohue & Letourneau, 1992; Rosen & Keith, 1978; Zuckerman, 1971). Penile Plethysmograph responses to slide images and audio stimuli have been reported to provide relatively accurate information to classify child molesters and men who commit sexually coercive acts, such as rape, into more refined, delimited diagnostic groups and/or differentiate them from normals or other sex offender and non-sex offender groups (Barbaree & Marshall, 1984; Fedora et al., 1992; Lalumière, Quinsey, Harris, Rice, & Trautrimas, 2003; Quinsey, Steinman, Bergersen, & Holmes, 1975; Wormith, 1986).

Given the varied points of view and, to some extent, court rulings regarding the potential utility of the PPG, AASI, and polygraph, some forensic evaluators and clinical practitioners may be unclear about the value of such physiological assessment methods. The aim of this chapter is help lower the "noise level" and guide the evaluator in understanding, administering, interpreting, and then applying the PPG, AASI, and the polygraph in the forensic and clinical settings for optimal utility relative to exploring potentially useful additional sources of information about a particular individual related to sexual offending.

Before launching into a discussion on first how to use the PPG in applied settings, we need to be better grounded in its history, empirical support, and ethical concerns.

The Development of Penile Plethysmography

Sexual offenders are not the only category of individuals that might be motivated to deny their sexual interests or preference. In response to the Czechoslovakian government's concern about recruits attempting to evade military service by falsely claiming to be homosexual, Kurt Freund was commissioned to develop a procedure to differentiate sexual preference. In 1957, he employed a device, the volumetric method, to measure blood flow into the penis. He called this method the Penile Plethysmograph (commonly abbreviated as PPG). From this beginning, Freund's research evolved to focus on detection and diagnosis of sex offenders, particularly pedophiles (Wilson & Mathon, 2006).

The popularity of the PPG should be understandable, considering the unreliability of an offender's self-report and the fact that through such an evaluation, there is confidence that a meaningful erectile response to a sexual stimulus presentation of an adult or child is a psychogenic arousal and not simply a random erection (Bradford et al., 2010; DSM-5 Proposed Revision, 2010; Heilbrun, 2003; Janssen, Everaerd, van Lunsen, & Oerlemans, 1994). An offender's reluctance to be fully disclosing about both his current level of deviant arousal and his history of sexual offending is not difficult to understand, considering his fear of societal disapproval and the painful consequences that might follow an accurate accounting of past deeds and current deviant sexual interests. It is common for therapists working with recidivist, i.e., repeat sexual abusers of children and adults, even those with multiple detected victims spanning years, to hear from the sexual perpetrator denial or minimization of their offense history and/or denial of past or current deviant arousal. Abel, Mittelman, and Becker (1985) found that among 411 outpatient volunteers, the subjects initially provided very low reports of their past incidents of sexual crimes. However, when confronted with PPG results demonstrating erectile responses to sexually deviant stimuli, a large majority of those same sexual abusers in their study subsequently admitted that they had committed many more sexual offenses than they had previously disclosed. In a similar manner, Abel et al. (1988) discovered that among 561 nonincarcerated paraphiliacs, when provided assurance of confidentiality (e.g., a Federal Certificate of Confidentitality), most disclosed having engaged in as many as ten different types of sexually deviant behaviors that were previously unknown. This was

evidence of "crossover," with a significant number of sexual offenders reporting multiple types of atypical sexual behaviors as opposed to just one type, such as "rape" or "child molesting." Similar reports of previously non-disclosed victims and a history of varied sexual offending (e.g., crossover offending) have been reported by other researchers relying on polygraph examinations of sex offenders (English et al., 2000; Grubin, Madsen, Parsons, Susnowski, & Warberg, 2004).

Early Beginnings to Present

Contemporary Penile Plethysmography has generally changed from early Volumetric Plethysmograph measurement in which the blood flow into the penis is measured by the air displacement from a glass or rigid cylinder that is placed over the penis with an inflatable cuff. The cylinder is generally held in place with a leather harness that the technician places on the individual. This cumbersome process was first simplified by Bancroft, Jones, and Pullan in 1966 and later by Barlow, Becker, Leitenberg, and Argus in 1970 (Coric et al., 2005) to the method commonly used today, a circumference gauge. The gauge typically used today is a simple, thin mercury-filled elastic strain gauge that is placed on the midshaft or base of the penis. The gauge stretches as the penis circumstance expands with penile engorgement.

From its early beginning in the 1950s, the measurement of erectile responding became a central component and standard in the evaluation and treatment in the field (Marshall, 2006a, 2006b; McGrath, Cumming, & Burchard, 2003). The Penile Plethysmography equipment one is most likely to find in labs today consists of a computer with plethysmography software that includes sexual stimuli presentations and an attached wire leading from a circumferential mercury strain gauge (an elastic mercury band designed to measure electrical impedance to detect blood volume changes) that is calibrated on a calibrating rod prior to testing. The subject places the wire around the midshaft or base of the penis, and audio, video, or audio/video equipment is used for the examinee to listen to and view the stimuli. Some labs continue to use Volumetric Plethysmography, most notably the Centre for Addiction and Mental Health, which was described earlier. Volumetric Plethysmography requires placement of a cylinder over the penis and measures air displacement caused by erectile engorgement during stimulus presentations. Volumetric measurement is reported to be more accurate than mercury strain gauge circumference measurement for erectile response that is less than 10 % (2.5 mm) of full erection. However, both volumetric and strain gauge results were highly correlated for 10 % (2.5 mm) and greater of penile circumference increase (Barbaree, Blanchard, & Kuban, 1999).

Pros and Cons of Penile Plethysmography (PPG)

Proponents of Plethysmography hold that if it has been established that "Pedophilia" is a term that describes a sexual interest in largely prepubescent children, then the PPG is the most effective method for assessing such a sexual interest. Further, it is claimed that the PPG is useful in tracking erectile changes and it is the most well-established available method for assessing sexual interests (O'Donohue & Letourneau, 1992; Rosen & Keith, 1978; Zuckerman, 1971). PPG responses to slide images have been reported to be reasonably accurate in classifying child molesters into diagnostic groups and/or differentiating child molesters from normals or other sex offender and non-sex offender groups. Using the combined method of presenting erotic slides of nude children and adults and audio stimuli together with a self-report card sort (written scenarios of 13 categories of attractiveness to various description of sexual interest) (Laws, 1996) measures to differentiate boy-object and girl-object child molesters provides classification accuracy of 91.7 %, which is greater than any single measure (Barbaree & Marshall, 1984; Baxter, Marshall, Barbaree, Davidson, & Malcolm, 1984; Fedora et al., 1992; Freund, Watson, Dickey, & Rienzo, 1991; Laws, Gulayets, & Frenzel, 1995; Laws, Hanson, Osborn, & Greenbaum, 2000; Quinsey et al., 1975; Quinsey & Carrigan, 1978; Quinsey, Chaplin, & Carrigan, 1979; Wormith, 1986). This finding supports the notion that "more is better" in that using the PPG along with other measures of sexual interest is likely to give you the most comprehensive and accurate picture of the individual's sexual interests.

The primary focus of this chapter is on the assessment of child molesters and rapists in evaluation and treatment settings. Most sex offenders, including rapists are eventually released into the community. Some suggest that the results of PPG studies for rapists are a bit muddier than those for child molesters. A number of researchers have argued that PPG reliability with rapists is too low for its valid application in assessment (Eccles, Marshall, & Barbaree, 1994; Fernandez & Marshall, 2003). Barbaree, Baxter, and Marshall (1989) reported test–retest reliability of the rape index was extremely low ($r = 0.44$). However, Lalumière et al. (2003) revisited and updated quantitative reviews of studies that examined phallometric responses of rapists and other men. They discussed many laboratories assessing rapists have reported that approximately 60 % of rapists (perhaps a modest but still significant detection level) show rape indices that are larger than the rape indices of about 90 % of non-rapists. This 60/90 benchmark is a cut-point that can produce a score that determines interest for rape. In other words, good group discrimination between rapists and non-rapists is suggested

by these results. Lalumière et al. (2003) suggested that future research would be valuable in distinguishing among three potentially different sexual arousal patterns of profiles as these apply to rapists' phallometric responses: biastophilia (sexual arousal involving nonconsenting, struggling, resisting, but not necessarily to injury or cause physical suffering of the victim); sexual sadism (sexual arousal to pain, suffering, and injury); and the general antisocial or indifferent rapist (indifference to the interests, feelings and desires of others).

In a study of 586 male sex offenders convicted of contact sexual offenses assessed between 1982 and 1992 whose recidivism was studied over a 20-year follow-up, Kingston, Seto, Firestone, and Bradford (2010) investigated the predictive validity of sexual sadism, as indicated by psychiatric diagnosis, level of violence during the most recent sexual offense, the intrusiveness of the sexual activity, and phallometrically assessed sexual arousal to depictions of sexual or nonsexual violence. They found that the three behavioral operationalized indications (level of violence, sexual intrusiveness, and phallometrically assessed sexual arousal to sexual and nonsexual violence) were better predictors of sexual recidivism among sex offenders than the psychiatric diagnosis of Sexual Sadism. Of special interest here are the phallometric results of the study.

Kingston et al. calculated the *Pedophilia Assault Index* by dividing the highest response to an assault stimulus involving a child victim (nonphysical coercions of child, physical coercion of child, sadistic sex with child, or nonsexual assault of child) by the highest response to a child stimulus with no overt form of coercion. Similarly, they calculated the *Rape Index* by dividing the highest response to the rape stimulus by the highest response to the adult-consenting stimulus. The *Adult Assault Index* was calculated by dividing the highest response to a nonsexual assault stimulus against an adult by the highest response to a consenting adult. They then created a new index of sexual arousal of sexual and nonsexual violence, irrespective of victim age, that was simply the highest score from any of the three indices (Pedophilia Assault, Rape, and Adult Assault indexes). They found that phallometrically assessed sexual arousal to violence added to the prediction of violence (including sexual) recidivism after actuarially estimated risk to reoffense was controlled. This study's findings suggest that behaviorally operationalized measures, including the results of phallometric assessment, are preferred over psychiatric diagnosis because the phallometrically assessed deviant arousal to violence, including sexual violence, was associated with recidivism; whereas, in contrast, psychiatric diagnosis of sexual sadism was not associated with recidivism. This is yet another argument in support of the use of the PPG for assessment of sexual deviance for violent sexual offenders.

From the perspective of the evaluator and treatment provider, the available research provides sufficient support that the results of a PPG examination of an individual can produce findings that could be useful when discussing with the client that responds strongly to sexual and nonsexual violence. Some writers have suggested that since PPG testing with rapists may not be as discriminating as that for pedophilia, it may be less useful for clinical or forensic purposes. However, in addition to the available research that does support the utility of PPG evaluations with persons accused of or convicted of sexual assaults of adolescents and adults, it has also been long established that a significant proportion of sex offenders tend not to be specialists and some rapists also have sexual interest in children (Abel, Becker, Mittleman, et al., 1988; Marshall, 2006a, 2006b). Therefore, the PPG can still useful to rule out the possible presence of pedophilic interests in persons who assault adult females and is, therefore, also recommended to be used in settings assessing the sexually violent rapist as well as the child molester.

Several studies have assessed PPG sensitivity and specificity for diagnosing pedophilia (Camilleri & Quinsey, 2008). Sensitivity is defined as the probability that the test says a person has the "disease" or condition such as Pedophilia when in fact they do have the disease or condition. Specificity is defined as the probability that the test says a person does not have the "disease" (Pedophilia) when in fact they are "disease free" (Sensitivity and Specificity, 2013). "Sensitivity" is calculated by dividing the number of men identified as pedophiles by PPG assessment out of the total number of true pedophiles in the sample. For child molesters with multiple child victims, sensitivity was reported to range from 61 % to 88.6 %. Offenders with male victims had higher sensitivity scores. "Specificity" is calculated by dividing the number of men identified as gynephiles (men who prefer adult women) by PPG assessment out of the total number of true gynephiles in the sample. The specificity range was 80–96.9 %. For samples among adolescent sex offenders, sensitivity was lower but still acceptable at 42 % (Camilleri & Quinsey, 2008). In summary, the findings reflect moderate to robust sensitivity and robust or strong specificity, meaning we can have greater confidence in PPG findings that indicate the presence of pedophilia rather than in the absence.

Although the PPG has been around for over 40 years and there is a large body of research supporting its use as the best-validated tool for assessing pedophilia, a single standardized way of administering the test and published norms are lacking (Camilleri & Quinsey, 2008). Marshall (2006a, 2006b), an early proponent of the PPG, has more recently raised questions about the clinical usefulness of the PPG. He wrote:

> "… clinicians who rely on phallometrics must offer compelling arguments for doing so. The evidence of the reliability and validity of phallometrics presently available in the literature certainly offers little support for its use … some may find justification in the present review for abandoning the use of phallometric assessments altogether (p. 21)."

Similarly, at one point, Laws (2003) opined the PPG should be viewed as more an art than a science because of a perceived lack of universally agreed-upon standards and procedures. He reported on a national effort in the United Kingdom to develop standardization guidelines in 2007 (Thornton & Laws, 2009). There were also earlier attempts in North America to standardize the age and gender assessment of PPG administration for child molesters in 1987. However, of the five sites in the United States and three in Canada, only one Canadian site completed the study. Laws expressed disappointment about lack of plans to standardize the PPG and noted that the PPG is intrusive, invasive of privacy, and time-consuming. He expressed frustration that it had taken too long for the emergence of standardized procedures and explicit protocols (Laws, 2009). However, more recently, Laws (2009) has taken a less negative view on the utility of the PPG. He has retracted some of his earlier criticism because of improvements in the field. He has also acknowledged that PPG works well if implemented in a relatively consistent fashion. Laws acknowledged that many clinicians and researchers believe the PPG is a valid measure of deviant sexual interest, reporting that PPG measures correctly classified 82 % of the offenders by sex of victim and 74 % by both victim gender and use of force. Further, he reported encouragement based upon the implementation of a multisite study in the United Kingdom and the detailed procedure manuals that have been developed as a result. He concluded these results may, at least partially, solve many of the problems that have existed previously.

Similarly, O'Donohue and Letourneau (1992) have opined that although there does not appear to be a single standardized penile plethysmography assessment protocol, recognized procedures do exist and have shared aims. The British Psychological Society has, in fact, published Penile Plethysmography Guidance for Psychologists (British Psychological Society, 2008). The Association for the Treatment of Sexual Aggressors (ATSA, 2001, 2013) has long supported the use of PPG by experienced professionals using one of the more standardized procedures. Marshall and Fernandez (2003a, 2003b) have also supported the PPG's value. The authors stated that the psychometric data for assessments from tools such as card sorts (self-rating of 13 categories of attractiveness to various descriptive paraphilic sexual interests) (Laws, 2009), self-report measures, viewing time, and clinical interview results are less satisfactory than phallometry and that these alternative measures cannot yet be considered as a viable replacement for PPG testing. They opined phallometry would continue to have a role in effective clinical assessment of sexual offenders, but cautioned the role should be restricted to (1) determining which offenders need treatment, (2) targeting its application at modifying deviant interests, (3) estimating whether or not treatment intervention has reduced deviant tendencies, and (4) estimating the likelihood that an individual will reoffend. They indicated PPG evidence of deviant arousal for any sexual offender is an indication of problems that need to be addressed in all the decisions made about the offender, including treatment choices.

Justification for using a considered or qualified approach in interpreting PPG findings can be found in studies measuring PPG sensitivity (44–86 %) and specificity (approximately 95 %). The test sensitivity (accuracy of correctly categorizing individuals with sexual deviance) and specificity (accuracy of correctly categorizing nondeviant individuals as not being sexually deviant) findings tell us PPG test results are most informative when some signs of sexual deviance are revealed However, when no sexual deviance is revealed with PPG testing, it *cannot* be concluded that the person is not aroused by children because PPG false-negative rate can range from 14 % to 56 % (Freund & Blanchard, 1989; Freund & Watson, 1991; Hall, Hirschman, & Oliver, 1995). Consequently, PPG results indicating "nondeviance" do not confirm the absence of pedophilia or absence of arousal to coercive or violent sexual stimuli. In contrast, we can be fairly confident an individual was correctly classified as sexual deviant if their PPG results indicate the positive evidence of sexual deviance.

Further, relative to the value of the PPG assessment for sexual offenders, the Hanson and Bussière's (1998) meta-analysis of 61 scientific reports on the prediction of sexual reoffending involved approximately 40,000 sexual offenders. They found that PPG measure of deviant sexual arousal to male children was the single most distinguishing marker for sexual recidivism. In a second meta-analysis of 91 studies of 31,000 sexual offenders, Hanson and Morton-Bourgon (2005) found further justification for PPG results related to child stimuli in relation to sexual offense recidivism. They reported that phallometric measures of any deviant sexual interest and sexual interest in children were significantly related to sexual recidivism. Most recently, in an updated meta-analysis, Mann et al. (2010) found that measured sexual interest in violence was itself a significant risk factor for sexual recidivism.

PPG and the Polygraph

Laws (2009) described a personal communications with Thornton (5 April 2007), who has integrated the PPG and the Polygraph procedures for assessment and treatment. In Thornton's procedure, the client is instructed during PPG testing to allow himself to become sexually aroused with no attempt to control his response. The second phase is called the "enhanced non-suppression PPG." The procedural instructions are the same, except the client is asked a series of questions about the sexual stimulus 30 s after it has terminated.

The purpose is to encourage the client to process more deeply. Later, after the PPG, the client undergoes a polygraph examination focusing upon his compliance with PPG pretest instructions; he is asked more generally whether he deliberately tried to distort the results.

Clearly, while issues regarding PPG use have been raised and considered, it remains a primary method for the assessment of sexually deviant interests. The popularity of the PPG is irrefutable given its widespread use throughout North America. In a survey of North American treatment programs, out of 330 community-based programs for adult male sex offenders in the United States (U.S.) and 19 in Canada, 27.9 % of the U.S. programs and 36.8 % of the Canadian programs measure sexual interest reported using the PPG. Residential programs' use of the PPG is even higher. Of 85 U.S. residential programs participating in the 2009 North American Survey, 36.5 % were using PPG assessments, and of 8 Canadian programs, 87.5 % were using it (McGrath, Cumming, Burchard, Zeoli, & Ellerby, 2009). Hecker, King, and Scoular (2009), in their investigation of alternative approaches to the measurement of sexual interest, referred to the PPG as the "gold standard" for measuring sexual interest because of the extensive research literature of phallometric testing and the strengths and limitations of plethysmography are well know.

Some of the drawbacks to PPG that exist are simply logistical ones, For instance, equipment, laboratory space, and time required in setting up a lab and prepping for a test are expensive. Also, staff training can be both expensive and time-consuming. This likely means that the smaller program providing sex offender treatment is at a disadvantage without cooperative arrangements to refer clients to other facilities for PPG assessment.

Ethical Concerns with PPG Use

Few topics in North American culture draw as much attention and controversy as the subject of sex. Therefore, it should not be surprising PPG examination of sex offenders causes many to worry about ethics regarding its use with adolescent and adult offenders. The ethical challenges related to the fact that PPG stimuli are designed to evoke deviant sexual arousal and that the testing is intrusive. Examples of related concerns mentioned in the PPG literature should be considered (Association for the Treatment of Sexual Abusers, 2001; British Psychological Society, 2008; Marshall, 1996):

- Explicitly deviant stimuli can be seen as providing tacit approval for the material.
- Exposing impressionable juveniles and adults to explicit deviant stimuli material might shape future sexual interest patterns.
- Stimuli are inherently degrading to women and children.
- Exposure to explicit sexually deviant stimuli can produce anxiety, nervousness, depression, and other emotional upset.
- Lab procedures that require the subject to self-stimulate to achieve maximum arousal can increase subject embarrassment and humiliation, and may be contrary to religious beliefs.

These concerns have dampened research enthusiasm as well as "caused" the reluctance of human research ethics committees or Institutional Review Boards (IRB) to permit PPG studies has made it even more difficult to address the empirical limitations identified by some writers (Marshall, 1996; McAnulty & Adams, 1991; Murphy & Barbaree, 1994).

Here are some of the "Do not's" associated with PPG testing recommended by Marshall (1996) and British Psychological Society (2008):

1. Do not use PPG testing results as the sole criterion for determining deviant sexual interests.
2. Do not use PPG testing alone for estimating risk for engaging in future sexually abusive behavior.
3. Do not use PPG testing results exclusively regarding recommendations to release clients to the community.
4. Do not use PPG testing to determine that clients have completed a treatment program.
5. Do not use PPG test results to draw conclusions about whether an individual has or has not committed a specific sexual crime.
6. Do not test an individual with sexually transmittable diseases until their symptoms are in remission.
7. Do not interpret PPG results in absence of other relevant information to determine risk and treatment needs.

Yet some of these recommended prohibitions have been refuted. Regarding risk assessment, the available data (e.g., Hanson & Bussière, 1998, 2004; Marshall & Fernandez, 2000; Mann et al., 2010) from three meta-analysis studies clearly demonstrated that PPG results regarding deviant sexual arousal are significantly associated with future sexually abusive behavior.

Caution is recommended when using PPG results with clients who are developmentally disabled or have acute major mental illness based upon the paucity of normative data for these populations and the invasiveness of the techniques (National Clearinghouse on Family Violence, 1998). Also, the developmentally disabled population may, for example, have difficulty accurately perceiving the stimuli because of limited ability to discriminate the age and gender in each of the presentations in the assessment, and they might have problems associated with understanding the self-report

procedure (Haaven & Schlank, 2001). Also, it is clear that PPG results must be interpreted and applied in conjunction with other relevant information to determine the risk and treatment needs of a particular sexual offender.

How *Should* the PPG be Used

Lalumière and Harris (1998) offer a list of best practices for optimal discrimination using phallometric testing. They recommend the testing involves:

- Use of images that best discriminate age and gender preference.
- Use of graphic and violent audio narratives that best discriminate preference for coercive sex.
- Use of more than one stimulus per category (2–5 stimuli recommended by Lalumière & Quinsey, 1994).
- Collection of data tracing after the stimulus presentation has ended (recommended at least 30 s).
- Computation of a "deviance differential" index of relative preference between deviant (child or coercive sex) and nondeviant (adult or consenting sex).
- Using *Z*-score transformation to address individual differences in responding (high vs. low responders) to improve discriminate validity, or percent of full erection (PFE). Both methods are believed to provide good validity values.
- When auditory stimuli are used, the employment of anti-countermeasure procedures are recommended for use (such as semantic tracking tasks) to detect faking and encourage subject compliance.

The British Penile Plethysmography Guidance for Psychologists' (British Psychological Society, 2008) instructions on good practice indicate the clinical purpose of the PPG is to provide physiological evidence of patterns in sexual arousal, facilitate participant acknowledgement of their sexual arousal/interests and their engagement in treatment, develop formulation of problematic or offense-related sexual behavior, assist treatment and risk management planning, assist in measurement of changes in sexual arousal/interest, and again emphasize that PPG should not be used to establish guilt or innocence regarding offense behavior. It further indicates responsibility for correct administration, interpretation, and supervision of the PPG assessment should rest with a supervising psychologist who has substantial up-to-date knowledge of the relevant literature, practice, legal and ethical issues surrounding PPG assessment, and substantial experience working with men who have committed sexual offenses.

Establishing and Administering a PPG Lab

It is best to standardize PPG laboratory facility and operating procedures with other PPG labs to obtain greater confidence that assessments are done correctly. Jensen and Laws (1994) provided helpful videotape instruction on the "How-To" of phallometry where the viewer is walked through the physical layout of the lab facility. This author has also inspected a number of labs in California, Washington State, and Toronto. The common theme found in each of these labs is that the physical facility maximizes privacy and minimizes distraction. Typically, the subject is placed in a separate room from the lab technician with a window that permits the technician to both observe and communicate with the subject. Some labs are including audio/video equipment. It can be prudent to video record the administration of the PPG session to discourage false claims that the subject was embarrassed by being required to be exposed to the examiner. Also, the recorded session can be viewed afterward to look for so-called countermeasures (T. Buttle, personal communications, February 8, 2010). The need for countermeasure detection is underscored by the discovery that some males are able to bias or invalidate results by suppressing their arousal, and in some cases they are able to increase their arousal (Abel, Blanchard, & Barlow, 1981; Adams, Motsinger, McAnulty, & Moore, 1992; McAnulty & Adams, 1991). A well-trained technician pays close attention to potential cues suggesting deception such as furtive hand and arm movement, breathing rhythm changes, and unusual erectile tracing patterns displayed on the computer monitor.

The subject's room should have a comfortable, washable chair. Covering the chair with disposable paper medical drape sheets ensures good hygiene and conveys a message that the testing is conducted in a professional environment. The visual stimuli used today are usually displayed using a full-size television or computer monitor that is sufficiently large to dominate the subject's visual field. The room lighting should be dim and the subject should not be distracted by other objects or noises. When using audio presentations alone, you may use small computer speakers, or an audio headset that can be cleaned after each testing. Additional equipment connected to the computer includes the following:

- Mercury strain gauge, which the subject places on his penis in private outside of the view of the technician
- Visual and semantic tracking device such as a keypad, which is used to track the subject's attentiveness, provide self-report of erectile response, and encourage compliance with testing procedures

- Pressure sensor seat pad to detect artifact movement
- Respiration measuring device to monitor breathing pattern

The subject can cover his lap with a medical drape or put on a medical gown to avoid exposure when placing the gauge on the penis and during test administration.

The technician's room is designed to monitor and communicate with the subject during testing. Although a great deal of technological advancement has occurred since Kurt Freund developed the PPG device in 1957, fundamentally little has changed in the actual measurement of penile tumescence. Blood flow is still measured by change in the size of the penis. Most labs use an elastic circumference mercury strain gauge such as D. M. Davis, Inc., HgPC, which is durable and reusable when washed and soaked in a 10 % mixture of chlorine and water. The gauge is connected to a computer, which records electrical impedance that occurs as the penis expands, stretching the strain gauge placed around the circumference of the penis. Typically, software programs specifically designed for PPG testing transmit video, audio/video, or audio stimuli to the monitor and speakers.

The goal is to administer stimuli that will elicit erectile responses sufficient to discriminate between deviant and nondeviant interests. Problems can occur with stimuli such as the Auditory Stimuli for Penile Plethysmography (1993) that include vague audio stimulus descriptions such as, "You are with a young girl…the age you like the most." The subject may visualize that the "young girl" in his mind is a 6-year-old child and become sexually aroused, but then report he imaged the "young girl" to mean an 18- or 19-year-old female. To avoid this type of deception or confusion when using such audio stimuli that doesn't clearly state the age of the sexual partner, the technician or the stimuli materials should give clear and specific instructions stating what the age category will be on each stimulus prior to beginning the assessment, and reinforcing the instructions periodically throughout the testing session.

A variety of types of stimuli have been produced over the years (film, videotapes, slides, audio recordings), but not all stimulus sets generate discriminating responses from the subject. For example, Abel and Blanchard found that videos generated the greatest levels of arousal, but that the strong arousal obscured differential responding, that is, the video overstimulated the subjects causing undifferentiated arousal to deviant and nondeviant stimuli (Marshall, 2006a, 2006b). Conversely, other stimuli sets may not be sufficiently arousing to generate meaningful erectile responses. Further, the duration of stimulus time matters. In a study of 31 child sex offenders aged 21–44, Avery-Clark and Laws (1984) reported that there was a significant difference in the arousal levels achieved between 2 and 4 min of stimulus presentations, suggesting the need for stimuli presentation longer than 2 min; a recommended stimulus time is 3 min. Another important consideration is at what point should the sexually significant event identifying the deviant or nondeviant theme occurs during the sexual vignette in the audio stimulus (Marshall, 2006a, 2006b). If the audiotaped stimulus provides over a 2-min description of sexually arousing behavior and then only toward the end reveals that the sexual partner is a prepubescent child, it is difficult to discern to what the subject is responding. Marshall's finding suggests the introduction of the sexually deviant stimuli aspect of the vignette should occur early on in the stimulus presentation to remove doubt about what is arousing the subject.

Laws and Court Decisions Impacting PPG Use

The stimuli used in PPG testing of sex offenders have not been without controversy. The government of Canada allows the use of what might be termed pornography for scientific or clinical purposes (Howe, 1995), whereas this is not the case in the United States. Concern over distribution of child pornography and legal sanctions against transporting child pornography across state lines, even for evaluation or research purposes, have made it difficult to standardize PPG procedures across evaluation sites (Howe, 1995). In the United States, federal statutes and state laws exist prohibiting the use of nude images of children for the purpose of sexual arousal (e.g., Federal Law, 18 U.S.C. § 1466A (2008) § 1466A. Obscene Visual Representation of the Sexual Abuse of Children; California Penal Code 311.3 & 311.11, Obscene Matter of a Minor). Legal prohibition of images of children construed to be sexually abusive or obscene have encouraged the development of clothed slide images of children and audio stimuli depiction of deviant sexual behavior described above for use in PPG testing. However, research shows that audio stimuli is reasonably accurate when compared to the nude stimuli (Barbaree & Marshall, 1984; Fedora et al., 1992; Lalumière et al., 2003; Quinsey et al., 1975; Wormith, 1986), suggesting valid PPG testing has not been significantly diminished by governmental restrictions.

PPG Stimuli

There are a number of sources for standardized stimuli that can accurately classify deviant and nondeviant subjects and which have demonstrated reliability and validity. Some stimuli are encrypted and are only commercially available through the software manufacturers, so it may be necessary to purchase the hardware equipment and software products from commercial distributors such as Limestone Technology, Inc. and Behavioral Technology, Inc. in order to obtain the stimulus set. The apparent reason for this is to encourage

potential users to purchase their product. This competitive spirit may make good business sense, but it is unfortunate for professionals in the field who seek improved standardization because it limits ready accessibility, thereby discouraging published reports on the same standardized stimuli.

One recently standardized audio/video stimulus, Real Children Voices (RCV) (T. Buttle, personal communications, August 8, 2010), is more creative than many of the older stimuli versions found and may prove to better capture an individual's unique arousal profile. The RCV aural portion of the stimuli includes the voices of the sexual partners. In other words, the subject hears both the adult male talking about the sexual behavior and the voice of the child sexual partner responding as the sexual scene is enacted.

American Psychological Association Ethical Principles 9.03 (2010) requires informed consent be obtained for PPG testing. The subject must agree to volunteer for the testing. Otherwise his resistance is likely to sabotage the testing. It is important to recognize that few individuals find much enjoyment sitting in a chair in a monitored room for over an hour with a gauge on their penis and being directed to pay attention to a variety of deviant and nondeviant sexual stimuli. Preparation of the individual for PPG testing starts with helping the individual client reduce his anxiety by explaining the testing protocol and familiarizing him with the lab before he begins testing. An examinee needs to be trained on how to correctly put the gauge on and how to use the keypad or other devices that encourage his cooperation. He must sign his informed consent to the test and be reassured that the facility and gauge are clean and his privacy is respected.

The test instructions should be standardized. The easiest way to do this is by reading a prepared script describing the protocol, or as with the RCV, the protocol may be described on the subject's monitor to get him ready for the stimulus presentation (T. Buttle, personal communications February 8, 2010). This automatic protocol assures greater standardization. The instructions walk the subject through the testing, starting with how to put the gauge on. The RCV stimulus set begins with the subject in the chair reading instructions on the monitor that inform him about what to expect in each stimulus presentation, then inform him how to respond to the visual and audio attention cues, and how to rate his level of sexual arousal on the keypad. Then the examination begins.

The subject should be instructed before the examination date not to masturbate to ejaculation 48 h before the examination because the sexual refractory period can last from a few minutes to days, depending on age, frequency of sexual activity, and other factors (Crooks & Baur, 2008). To ensure subjects' compliance subsequent to PPG testing, he can be administered a polygraph test and asked if he engaged in dissimulation behaviors to bias the PPG results. Just prior to testing, it should be suggested that he take a restroom break. Also, he should be discouraged from consuming liquids such as coffee or soda beforehand because he will be sitting in a chair for over an hour with no restroom opportunity.

Administration of the PPG

Once the subject is in his seat and the testing has begun, the technician observes both the subject and the computer screen tracings displayed on the monitor. The technician monitors the subject's compliance with the testing protocol and may need to remind the subject how to correctly follow the anti-countermeasure cues. For example, with the RCV stimuli the subject is required to press the OK button on the response pad when the picture shows a different person from the one shown in the previous photo.

The technician watches for a subject attempting to "beat the test" by moving, tensing his muscles, not paying attention, holding his breath, or breathing rapidly. Keen alertness to countermeasures is necessary because research has shown that the phallometric test is easy to fake (Laws, 2003; Wilson, 1998). Indications of faking may also include a wavy arousal pattern viewed on the technician's computer monitor. This pattern suggests the subject is attempting to control his arousal. Similarly, a flat tracing pattern during the presentation followed by arousal after the presentation and continuing beyond 30 s also suggests the subject may be suppressing his arousal. It has also been reported that high responses to neutral stimuli may be another sign of faking (Freund, Watson, & Rienzo, 1988).

Interpreting the Data

Many labs use cutoff scores for low response levels. This author is aware of labs that use cutoff scores ranging between 10 % and 20 % of full erectile response. A full erectile response has been determined by various means, such as requiring the subject to masturbate to full erection and then stop before ejaculation, displaying highly erotic stimuli during a pretest examination designed to measure full erectile responses, and estimating average erectile circumference.

The masturbatory procedure is viewed as problematic. Many subjects are uncomfortable being monitored while masturbating and when full arousal is achieved it might influence tumescent response to subsequent stimuli. Others object to the masturbation procedure because of religious reasons. Viewing highly erotic stimuli before testing might also exaggerate sexual responsiveness during the test administration.

An estimated range of full erection, a preferred procedure in this author's opinion, was reported to vary from between 25 mm to 30 mm, until the results of Howe's (2003) study of

circumference scores were obtained from 724 respondents at nine North American correctional facilities. He reported that flaccidity to full erection for male sexual offenders has a mean of 32.6 mm and a standard deviation of 8.8 mm, and the scores are normally distributed. Ninety-five percent of circumferential change scores can be expected to fall at or below 47 mm. He contended that using anything less than 47 mm as an estimate of full erection is unacceptable by conventional scientific standards. Following this recommendation, 10 % erectile response converts to about 4.5 mm, 15 % 7 mm, and 20 % would be about 9 mm. For clinical purposes with circumference measurement, we use 10 % percentile response as the minimum response level. If the subject does not achieve 10 % erectile response to any of the stimuli, then he is classified as a nonresponder, meaning his responses were too low to interpret the PPG results. Fifteen percent and 20 % erectile responding is often used as the cut score indicating some minimal degree of sexual arousal.

However, Lalumière and Harris (1998) indicated that cutoff scores might not be needed. They reported that they were unable to find any minimum response level that increased validity. They noted that there is no discriminant validity data of which they were aware of that supported declaring low responder data as useless, except when responses to the neutral stimuli are higher. They also argued that there is a mathematical advantage to using standard scores because one can obtain information from low responders that is just as good as from high responders. One might question the advisability of interpreting data on low responders knowing the circumferential strain gauge is inaccurate below 2.5 mm, and low levels of responding might not be due to sexual arousal (Barbaree et al., 1999).

This author recommends use of both standard score and percentage of full erection (PFE) when interpreting data in the clinical setting. The clinician will oftentimes discover that the PPG results reflect the same rank order of age and gender preference using either the *Z* score (standard deviation from the mean) or PFE. This means that the individual's sexual arousal profile looks about the same regardless of whether the raw score is converted into *Z* scores or percentage of full erection. If the individual has higher sexual arousal to children than to adults, both measures are likely to reflect this same arousal pattern.

After each presentation has ended, the circumference tracing should continue to be collected for at least 30 s (Barbaree et al., 1999). I have observed with older subjects their sexual responsiveness is generally more gradual and takes longer than 30 s before reaching peak arousal. Therefore, you may want to continue to observe the tracing another 30 s. The next presentation should not begin again until the subject has returned to baseline. Unfortunately, some subjects may never return to their original flaccid baseline during the testing session. For example, in some cases a subject may have a stable baseline reading of 95 mm after placing the gauge on his penis, and then after 30 min his baseline drops 10–85 mm. Others might become aroused during a presentation and not return to the original baseline during the rest of the session. Variation of baseline of 3 mm appears to be acceptable (W. Burke, personal communication, July 10, 2009) and will not affect either the *Z* score or PFE because data collection begins at the start of each presentation. In cases where there are dramatic baseline shifts, one has less confidence in the accuracy of the data.

In a clinical practice, the PPG results may be of limited value unless the findings are shared with the subject; sharing results may provide the impetus for a particular subject to move past denial and minimization and help focus and motivate the individual to address his deviant sexual interests as treatment needs. During the testing the subject has been asked to estimate the percentage of his highest erectile response after each presentation. Therefore, it makes sense to provide him with the findings of the PPG in PFE as well. It is not uncommon for the subject to report during the testing that he was not aroused by any of the stimuli. Therefore, it may also be helpful to show him his arousal-tracing pattern of PFE on the computer monitor to assist in effectively communicating the PPG findings. This is an important juncture of communication between clinician and client. The client may have started the PPG session admitting to his offense behavior, but denying any ongoing deviant fantasies or urges. Presented with the results of a PPG, this may change; that is, an examinee presented with evidence of demonstrated sexual arousal to deviant stimuli might subsequently acknowledge recent or current experience of deviant sexual fantasies or urges (Schwartz & Cellini, 1999). This writer has heard candid admissions of sexual deviance on countless occasions after PPG testing results are shared with the subject.

Lalumière and Harris (1998) recommended using deviant index scores to determine the presence of deviant sexual interest (e.g., highest average deviant score to children minus highest average nondeviant score to adults). Their reasoning is that computing a numerical deviant index score enhances validity. They also indicated that if one obtains a single clinically significant score suggesting arousal to children, for example, but the deviant index does not support the conclusion of sexually deviant interest, one does not give weight to the deviant score because the subject's average sexual preference is not deviant. This author argues that in testing individuals with a history of inappropriate sexual contact with children, even a single clinically significant "hit" (e.g., response) to a child or sexual violence is clinical information that should not be ignored. Such information provides an important starting point for a dialogue with a client who may be denying or minimizing his interest in children. The clinician needs to bring this finding to an examinee's attention and ask him if he was aware of his arousal; under what other

circumstances this arousal has occurred; how often it has occurred; frequency of recent deviant sexual thoughts; and frequency of fantasies about children during masturbation. Faced with evidence of his sexual arousal to one deviant stimulus even when he had greater responsiveness to adults, the client may become more inclined to identify transient or persistent sexual interests in children. This is especially true when the findings from a PPG examination are combined with those of VRT (Maram & Koetting, 2004), and then further clarified when followed by a polygraph examination. The value of the polygraph in revealing previously unknown sexually deviant behaviors was demonstrated in a study of 109 individuals under the jurisdiction of the Colorado Department Corrections (Ahlmeyer, Heil, McKee, & English, 2000). The researchers reported the mean number of victims revealed by Presentence Investigative Reports and Sexual History Disclosure form nearly doubled from a mean of 61 to 109 victims after the first polygraph.

Court Challenges and Research Regarding the Penile Plethysmography

In U. S. v. Powers, 59 F.3d (1995), the Federal court determined the penile plethysmograph test did not meet the scientific validity prong of Daubert. Laws (2003) summarized the legal literature surrounding the admissibility of PPG evidence in court as follows: (1) the technique has been tested; (2) it has been peer reviewed and published; (3) the procedure has a known or potential rate of error; (4) there is standardization for operation and the PPG is generally accepted in the scientific community. Laws then opined PPG should not be accepted in Daubert. He indicated that to some extent items 1 and 2 have been met. The PPG has been tested thousands of times, and with highly unvariable results. Also, it has been peered reviewed and published hundreds of times. However, the absolute error rate is unknown (item 3), and adequate standards do not exist for administration of the procedure, failing to meet the criteria of item 4. In Powers it was argued that the district court erred in excluding the testimony that the penile plethysmograph test did not indicate pedophilic characteristics. The district court excluded this evidence because, in its opinion, the test did not satisfy the "scientific validity" prong of Daubert. The U.S. Court of Appeals 4th Circuit affirmed the lower Court's ruling, excluding the PPG results. However, the court also noted the plethysmograph test is "useful for treatment for sex offenders" and permitted the district court to impose this condition on an individual as part of their supervised release from custody. This particular decision prohibited the use of PPG results in the guilt phase of Federal criminal proceedings, but allowed it to be used in treatment.

In Berthiaume v. Caron, 142 F.3d 12, 1st Cir (1998), the PPG was described as "an accepted tool" and "a standard practice" in the field of sex offender treatment. Barker and Howell (1992), at the time, reported that there was much research to support the claim that the penile plethysmograph is a reliable and valid method of assessing erectile response in male sex offenders. They concluded that while the PPG is the best objective measure of male sexual arousal and could be useful in assessing and treating sex offenders, caution still must be exercised because of its limitations. These limitations included a lack of standardization, a high incidence of both false negatives and false positives, and the use of the PPG unsupported by other data as a predictive test. Barker and Howell (1992) suggested that the PPG is most effective in predictive situations when it is used in conjunction with multiple data sources.

In State of North Carolina v. Spencer, 459 S.E. 2d 812, 815, N.C. Ct. App (1995), the court reviewed the literature and case law and concluded that penile plethysmography was scientifically unreliable. They concluded that despite the sophistication of the current equipment technology, question remains whether the information emitted is a valid and reliable mean of assessing sexual preference.

In a more recent review, it was reported that a substantial amount of research data has been gathered and reviewed, and significant steps have been taken toward standardization. According to Launay (1999), "[T] he the validity of the technique for research and clinical assessment has been is now established."

Other than in the guilt-determination phase of court proceedings, phallometry is now widely considered appropriate for treatment and supervision of convicted sex offenders. The courts are now permitting plethysmographic testing for monitoring compliance of sex offenders with the conditions of their community placement as part of crime-related treatment for sexual deviancy (Sachsenmaier & Peters, 2002).

The scientific validity and reliability of the procedure has also earned acceptance in many jurisdictions during the presentencing stage of criminal proceedings, as well as for the parolee or probationer who is under community supervision. The standard of evidence required need only be sufficient indicia of reliability to support "probable accuracy," a standard analogous to preponderance of the evidence; this is a standard less stringent than the Daubert standard (U.S. v. Silverman, 976 F.2d, 1992; U.S. v. Herrera, 928 F.2d 769, 772, 6th Cir, 1991; U.S. v. Lee, 1998). As a result, the Courts are now more routinely upholding the use of PPG testing in administrative law cases and with probationers for evaluation and treatment (Berthiaume v. Caron, 142 F.3d 12, 1st Cir, 1998).

The debate over PPG use is not limited to the United States. In Canada, the Canadian Supreme Court (R. v. J.-L. J., 2000) ruled against admitting penile plethysmography into

evidence in a case in which a psychiatrist (who was a Canadian pioneer in the field) attempted to testify about the results of the penile plethysmography (previously recognized by the Courts as a therapeutic tool), as a forensic tool in criminal procedures. Similar to the United States, the court opined that although PPG test level of reliability in a court of law was not necessarily sufficiently reliable to identify or exclude an accused individual as a potential perpetrator of an offense (a criminal application), they identified it as quite useful in therapy because it yields information about a recommended course of treatment.

In summary, court rulings thus far have not provided a bright line regarding the admissibility of the PPG in court proceedings. Generally speaking, the PPG is not allowed in the guilt phase of criminal proceeding (an exception to this will be discussed later from People v. Stoll, 1989). The PPG, however, may be allowed in the sentencing phase of court proceeding, is often introduced during Sexually Violent Predator cases and in other civil proceedings, and is permissible for use in sex offender treatment.

Summary on Penile Plethysmograph

The PPG has been in use for sex offender assessment and treatment for over 50 years. Although there are a large number of published articles on the subject, there continues to be controversies regarding its validity and reliability. In particular, concerns have also been expressed as to the lack of common standards and procedures and the variability of PPG administration, results, and interpretation. Consequently, some have suggested and continue to believe that using the PPG is more of an art than science (Laws, 2003, 2009). Others strongly support its role in clinical assessment (ATSA, 2001, 2013; Marshall & Fernandez, 2003) and point out that although the procedure may show varied standardization, there are recognized functions for its use that share common aims and features and utility (O'Donohue & Letourneau, 1992), such as treatment planning and hypothesized motivation of the underlying offense.

The recommended procedures described here include using a lab that ensures privacy and discretion; using discriminating stimuli that have been standardized and validated; carefully screening the PPG session for faking; using video recording as well as other devices that require the subjects' consistent visual and auditory attention; and using measures to check breathing and muscle movement to reduce threats to reliability and validity. In a clinical practice, it is important that data be shared with the client in a manner that is readily understood. Therefore, during the posttest interview, it is recommended that data interpretation be shared in "percentage of full erection" to encourage a client to report his inner deviant experiences and actual behaviors with greater openness and accuracy. The presence of any significant deviant arousal, even when there is greater average arousal to adults, is important clinical information and should be included in the dialogue with the client.

At present, the PPG continues to be the most sensitive and reliable available physiological measure of sexual arousal (Howe, 2003). However, until there is a convergence of standardization, the PPG will continue to be colored by controversy, and its admissibility in court for certain functions will remain uncertain. Effective sex offender assessment requires leadership and communication to establish professional consensus, to yield agreed upon standards of practice in which improved validity and reliability studies can follow. In the interim, although the clinician must rely upon less than ideal guidance when administering and interpreting PPG data, in this author's opinion, there are sufficient arguments supporting its value to justify its continued use in clinical practice.

Viewing Time Measures

Several other measures have been used in an attempt to reliably measure sexual arousal and interest. These include facial electromyography, measures of penile temperature, volume, circumference, and motion (Krueger, Bardford, & Glancy, 1998). It is postulated that sexual arousal is not a unitary construct and identifies three stages of sexual attraction in males: (1) aesthetic response, a hedonic feeling response to the sexual stimuli in which the individual may keep the object of interest in view; (2) an approach response where the individual moves toward the sexual object of attraction with a desire for body contact; and (3) the genital response characterized by greater penile engorgement. The third stage is the purview of the PPG, which was the subject of the previous discussion in this chapter. An increasingly popular method for assessing sexual interest involves Visual Reaction Time (VRT) measures, which rely on increased visual response to potential objects of attraction (or the first component of Sing's model, the aesthetic or hedonic response).

Abel, Jordan, Hand, Holland, and Phipps (2001) reported visual reaction time measurement was originally based on the work of Rosenzweig (1942). Rosenzweig, in a study with 20 schizophrenics using the photoscope, was the first to report that VRT was a good objective device for identifying sexual interest. Subsequently, Wright and Adams (1994) (in a study of 80 subjects using a VRT test) found that sexual arousal interfered with cognitive processing. Specifically, they found that individuals showed a longer reaction time to slides depicting preferred sexual partners than to nonpreferred sexual partners or neutral scenes.

Abel et al. (2001) addressed the difficulty and the importance of determining what motivates a person to sexually molest a child when developing an effective intervention

strategy to prevent future molestations from occurring. Support for the use of VRT as a measure to differentiate child molesters from non-child molesters and to identify individuals who are concealing their interest in children has been repeatedly reported in the literature (Abel et al., 1998, 2004; Harris, Rice, Quinsey, & Chaplin, 1996). Although Harris et al. (1996) reported that VRT was less intrusive than the PPG and significantly discriminated between child molesters and normals, they also noted the PPG was still better at discriminating these categories than VRT measures. Maram (2005; Maram & Koetting, 2004) found that use of VRT and the PPG incrementally increased discriminant validity, especially with child molesters of male youth. This finding supports the combined use of these instruments in the clinical setting.

Perhaps the best-known and most frequently used VRT measure is the *Abel Assessment for sexual interest*™ (AASI). The AASI has hundreds of licensed sites throughout North America authorized to administer the AASI testing instrument. The AASI was developed to function as a viable alternative to the PPG (Abel Screening, Inc. 2004). The licensed site setup cost to run AASI testing is significantly less than what is required for a PPG lab. The test comes with a training manual, Abel and his associates conduct training workshops several times a year and there is an online examination for qualified users. In addition, Abel and others at Abel Screening are available for consultation. In comparison to the PPG setup cost for the hardware, software, office space, laboratory equipment, and training, the AASI is a particularly economic alternative. In addition, it is not nearly as intrusive as the PPG (no one has to take their clothes off or attach devices to their genitals) and it does not cause the anxiety and distress often experienced by individuals taking the PPG. Using AASI as an inexpensive replacement to the PPG can be appealing to many sex offender evaluators and treatment programs and may explain the popularity and wide use of this instrument. However, as previously indicated, using the AASI and the PPG together has the advantage of increasing the individual's candor and willingness to cooperate with treatment.

Both the AASI and PPG have advantages and disadvantages based on the specific methods used. Although PPG and VRT measure different phenomenon, they both should be at least conceptually related to sexual interest and sexual attraction.

Research and development of the AASI demonstrated the test has criterion validity based on its ability to discriminate between non-child molesters and admitting child molesters. Abel has also demonstrated that the AASI was resistant to falsification based upon a statistical regression model designed to discriminate between "lier-deniers" child molesters and non-child molesters (Abel et al., 2001).

About the AASI

The AASI is a two-part examination. The first part is the VRT procedure relative to images of individuals of varied ages and races as well as a detailed questionnaire examining sexual interest, arousal, and behaviors. Both VRT data and self-report data are used together to assess respondents' sexual interest(s) and to calculate probability values that reflect the likelihood that a respondent has pedophilic sexual interests (Abel et al., 2001) The VRT is an ipsative measure, meaning the individual's standard scores are not normed to others, but compared only to the individual's personal scores to the visual stimuli administered as part of the procedure. In other words, the person's sexual interest to images of children, adolescents, and adults is normed only for the individual and is not compared to others.

Questions About AASI Reliability and Admissibility in Court

Like other assessment instruments, the AASI-2 has its detractors and its admissibility in court has been challenged. Smith and Fischer (1999) reported in their study that the Abel Assessment for Interest in Paraphilias used with juvenile sexual offenders in residential and day treatment failed to demonstrate adequate reliability and validity. They concluded there was no evidence that the test produced reliable scores for adolescents that could screen deviants from normal individuals or could diagnose specific pathology in deviant subjects. Abel et al. (2004) responded to Smith and Fischer's article with a counterargument to their criticism, citing numerous flaws in their study, the most central of which was the authors' failure to determine whether members of their control group were really "non-child molesters" or lacked sexual interest in children. The importance of this determination was underscored in a 2001 study (Zolondek, Abel, Northey, & Jordan, 2001) that reported information gathered from 485 males younger than 18 who were being evaluated as possible juvenile sex offenders. More than 60 % reported involvement in child molestation. Of the boys who reported never being accused of child molestation, 41.5 % reported they had molested a younger child.

Abel (personal communication, February 5, 2008) indicated that the AASI should not be used in making a diagnosis, nor does he claim it to serve that purpose. He pointed out that not all sexual abusers of children have a sustained sexual interest in children. He describes unpublished data showing VRT sensitivity as .44 and specificity of .81 when using a very high cutoff score to limit false positives. His study was based on 7,773 admitted sexual abusers of minors (children

or adolescents) and 365 non-sex offender community volunteers. He notes, "the sensitivity would be even higher if one only considered individuals who sexually abused children instead of combining the sexual abusers of children and sexual abusers of adolescents."

Letourneau's (2002) study demonstrated the utility of the AASI with adult male offenders. She investigated the reliability and validity of VRT and PPG in a sample of 57 sex offenders incarcerated at a high-security military prison. She reported the results indicated adequate internal consistency for both measures. The convergent validity and assessment of clinical usefulness indicated that both measures accurately identified sexual offenders against boys. The VRT, but not the PPG, also significantly identified offenders against adolescent girls. However, neither measure reached statistical significance in identifying offenders against adult women or against young girls.

The AASI has been used in various criminal court proceedings and on numerous occasions has been ruled as admissible evidence. For example, in U.S. v. Stoterau, 07-50124 524 F.3d 988, 9th Cir (2008), the district court ruled that Mr. Stoterau could be subjected to the AASI as a condition of his supervised release. Similarly, the Ninth Circuit, Central District of California, ruled that the district court may require AASI as a condition of supervised release (2006). The U.S. District Court of Louisiana ruled the AASI met the Federal Daubert standard (G. Abel, personal communications, February 5, 2008).

However, the AASI has not been uniformly accepted by all the courts. In Ready v. Commonwealth of Massachusetts, 2002, the AASI was found not to meet the Daubert standard for scientific validity because the original research study that developed the "rule of thirds" used to score the VRT (The "rule of thirds" refers to the cutoff score used to determine if sexual interest to an age and gender category was detected by the test) was never published (G. Abel, personal communications, February 5, 2008).

The AASI, as well as the PPG and polygraph examination, results are admitted into evidence generally without challenge when both sides stipulate to its use in an evaluation of sexual interest and arousal. Additionally, the California Superior Court has ruled that an expert may rely on standardized psychological tests in formulating an expert opinion (People v. Stoll, 1989). This allows the expert to offer an opinion about an individual's character relying upon their assessment of the individual using standardized testing such as the AASI, PPG, and polygraph without the Frye standard applying, which prohibits admissibility of evidence in the court of new or novel scientific techniques that are not generally recognized as sufficiently established by the scientific community (Frye v. United States, D.C. Cir, 1923, 293 Fed. 1013). In other words, in Stoll, the scientific technique(s) which formed the basis for the expert's testimony are not required to be tested as generally accepted by the scientific community. This opens the door for ethical issues since Stoll allows the defendant to present expert opinions of good character to show non-commission of a crime. However, an expert may not ethically report the findings of the AASI, PPG, and polygraph as evidence that the individual did not commit a sexual offense.

VRT in the Clinical Setting

Using the AASI in the clinical setting is fairly straightforward. Testing should be conducted in a location that minimizes distraction. Testing begins after the individual receives standardized instructions and successfully completes a practice test of 15 slide images. Following the practice session, the client is administered 160 slide images of the AASI on a computer screen. The slides consist mainly of clothed male and female models (there are no nude images) of different ages, ranging from age two to adulthood. The client is instructed to imagine being sexual with each image and then after becoming familiar with each slide, they are to indicate their perception of how aroused or disgusted he or she would become by the idea of being sexually involved with the slide image depicted. The individual indicates their degree of "disgust" or "arousal" by pressing numbers ranging from 1 to 7, which ranks the image from a "1" of very low interest and highly disgusted to a "7" which indicates the image is highly sexually arousing. Approximately, a client requires30 min to complete the VRT.

The second part of the AASI test is a comprehensive questionnaire concerning sexual interest, arousal, and behaviors, as well as questions about history of sexual abuse, cognitive distortion questions about statements which individuals who molest children often endorse, and social desirability questions to assess the individual's willingness to be truthful. There are different versions of the questionnaire for men, women, adolescent boys and girls, and special needs populations. Typically, the questionnaire requires about 45–60 min to complete. However, some individuals agonize over their responses to questions and take much longer. For clients with reading, cognitive problems with special needs, evaluators may need to anticipate at least 2 h and/or multiple testing sessions to complete the questionnaire.

Disingenuous responding or faking is an obvious concern of any type of psychological testing. As discussed previously, PPG is vulnerable to faking. Lanyon and Thomas (2008) reported that no research on AASI could be found that utilized non-admitters or deliberate faking. They concluded that the ability of VRT procedures to detect deceptive responders is unknown. Gray and Plaud (2005), in an investigation of test sensitivity of PPG and AASI with 63 participants (17 subjects were excluded because of low

responding on PPG testing) in an outpatient treatment program. Reported that both the AASI and PPG measures were able classify pedophiles to a high degree. The plethysmograph was able to classify 65 % of the participants correctly, while the AASI was able to correctly classify 79 % of those participating in this study. Gray and Plaud (2005) also observed what they referred to as a reflective responder, i.e., individuals who attempt to employ dissimulation on the AASI. They devised a formula to be used with the Abel Assessment graphs to detect for reflexive responders and report significant improvement in sensitivity.

THE AASI provides interpretation and training to its authorized users on the administration and interpretation of an individual's test results. The data from a particular test administration is electronically transmitted back and forth between Abel Screening, Inc. and the evaluation site. The raw data is computer-scored and returned as a sexual interest graph, which displays eight bars of the appropriate ethnicity, gender, and age category for the Caucasian and African-American client. Each bar has a *Z* score associated with each of the two age and gender categories of children, plus two gender categories of adolescents and adults, totaling eight individual age and gender categories. There are also other potential paraphilic interests measured, the most significant of which are the measures of sexual interest in male and female adult object sadomasochism. Bars at or below the vertical cutoff score (using the rule of thirds as the cutoff determinate) showed on the graph reflect suspected high sexual interest in that category. For instance, a client may have bars at or above the cutoff of adult females, adolescent females, and male children 2–4 years old. This profile suggests the individual appears to have sexual interest in adult and adolescent females, with suspected interest in prepubescent children. The mathematical formula used to calculate the *Z* score is proprietary information and not available for public dissemination. There has been criticism of AASI about secrecy surrounding the specific mathematical formula embedded in the test interpretation. Abel (Personal communications, February 5, 2008) defended this position saying the release of such information would compromise the test's utility for future test-takers. He maintains the accuracy and validity of the AASI, like the majority of psychological tests, is partially dependent on the test-taker not knowing how the test works. Also, to disseminate such information compromises the aspect of a naïve normative group. Further, as Abel noted, the Standards for Educational and Psychological Testing (specifically Standards 11.7 and 11.8) address such protection of copyrighted material. He pointed out with the advent of the Internet it is even more critical to safeguard the information that would compromise the usefulness of the test. In addition, Dr. Abel fully acknowledges his commercial interest in the proprietary nature of the software, which is within his rights protected under law.

Discussing the Result with the Client

After the individual completes the AASI testing, the clinician may discuss the test results with the client. More often than not, when describing the findings to a client with sexual interest detected to prepubescent children, the client reacts defensively and denies this attraction, sometimes claiming that they may have accidentally pressed the wrong number, or somehow used the computer incorrectly. It is helpful at this point to then discuss the results of the PPG, which generally reveals an arousal pattern to children similar to the AASI. Frequently, the resistive client begins to disclose more about their sexual appetite, but usually not everything. It is after the completion of the AASI-2, PPG, and then the polygraph examination that the client is likely to be most revealing about past sexual behaviors and current interests.

The Polygraph: Its History

Knowing what is truth and what is a lie has likely been a subject of conversation among people since language first evolved. Daniel Defoe in 1730 was not the first to suggest that "taking the pulse" was a practical and more humane method of identifying a criminal in his essay entitled "an Effectual Scheme for the Immediate Preventing of Street Robberies and Suppressing all Other Disorders of the Night." Defoe wrote:

> Guilt carries fear always about with it; there is a tremor in the blood of a thief, that, if attended to, would effectually discover him; and if charged as a suspicious fellow, on that suspicion only I would always feel his pulse, and I would recommend it to practice. The innocent man which knows himself clear and has no surprise upon him; when they cry "stop thief" he does not start; or strive to get out of the way, much less does he tremble and shake, change countenance or look pale, and less still does he run for it and endeavor to escape. (Matte, 1996)

In the 1900s, C. Lombroso, M. D. (an Italian criminologist) applied blood pressure-pulse test to actual criminal suspects on several occasions while assisting the police in identification of criminals. By the turn of the twentieth century, Verdin of Paris, a manufacturer of physiological apparatus, was producing ink-recording polygraphs with pneumatic tambours. Later, S. Veraguth (1907) began using word-association tests with the galvanometer. His observations of the galvanic phenomena and emotions noted that emotional complexes, unveiled in word-association experiments, made an ascending galvanometer curve, in contrast with the rest curve of non-crucial stimuli (as reported in Matte, 1996).

Larson, a University of California medical student, employed by the Berkeley California Police Department, invented the modern portable lie detector in 1921 (Matte, 1996).

Since then the polygraph has been used in many thousands of police interrogations and investigations and now is an essential part of many sex offender treatment programs. However, as with other psychophysiological measures, it is controversial among researchers and is not always judicially acceptable (Bellis, 2013).

Polygraph Research

Modern-day research conducted on psychophysiological veracity (PV), which is the polygraph examination, involves three types of validation studies: the analogue study, the field study, and the hybrid study. The analogue study employs a mock crime paradigm, whereas a field study involves testing of real-life suspects of criminal offenses. The hybrid study attempts to avoid weaknesses of both analogue and field studies by combing the best features of each. Most of the research conducted consists of analogue studies, which is problematic because the psychodynamics are quite different in mock paradigms (analogue) studies than in real-life cases. The analogue studies are appealing to researchers because absolute truth is known and it is easier to study because the investigator has complete control over the experiment. However, the analogue studies fail to duplicate three major emotions normally responsible for autonomic arousal in real-life suspects: fear of detection by the guilty examinee, fear of error by the innocent examinee, and anger by the innocent examinee. In spite of the shortcomings of analogue studies, many studies have shown remarkable results attesting to the validity of the PV examinations (Matte, 1996). In a field test of real-life criminal guilty knowledge tests (an investigation of 40 innocent and 40 guilty subjects), Elaad, Ginton, and Jungman (1992) reported that over 97 % of the innocent and nearly 76 % of the guilty subjects were correctly classified. Incredibly, some investigators have reported correctly identification of 100 % of innocent as truthful with no inconclusive findings and no errors (Mangan, Armitage, & Adams, 2008).

However, as a scientific tool, some researchers continue to find polygraphy of questionable validity (Iacono, 2008). Iacono considered the Mangan et al.'s methods flawed because the confessions in that study were obtained by the polygraph examiner who interrogated the examinee after deciding the test was failed. Iacono wrote, "Although largely ignored by the polygraph profession, this flaw inherent on confession-based field studies of polygraph validity has been known to confound these studies for over two decades. Hence, contrary to Mangan et al. (2008), their study design does not provide for an adequate estimate of polygraph test accuracy." (p. 25)

The American Polygraph Association website (http://www.polygraph.org), not surprisingly, reports studies more supportive of polygraph testing. They report researchers having conducted 12 studies of the validity of field examinations, following 2,174 field examinations, which reported an average accuracy of 98 %. Further, researchers conducted 11 studies involving the reliability of independent analyses of 1,609 sets of charts from field examinations confirmed by independent evidence, providing an average accuracy of 92 %. Researchers also conducted 41 studies involving the accuracy of 1,787 laboratory simulations of polygraph examinations, producing an average accuracy of 80 %. It was also reported that in 16 studies involving the reliability of independent analyses of 810 sets of charts from laboratory simulations, there was an average accuracy of 81 %. In summary, these studies indicate between 80 % and 98 % accuracy, giving strong support for the continued use of polygraph testing.

Polygraph: How It Works

The polygraph examination is really just a measurement tool of a person's autonomic nervous system. In the psychophysiological veracity (PV) examinations, the ear of the subject is the receptor that receives the potentially threatening question or stimulus from the polygraph examiner. The stimulus is transmitted from the ear to the reticular activating system (RAS). The RAS, part of the brain that regulates sleep–wake transitions, can influence the state of arousal depending on the nature of the stimulus. When a question is perceived as threatening, impulses trigger a sympathetic system response, which when activated prepares the body for "fight or flight" with secretion of hormones (epinephrine and norepinephrine). This causes constriction of the arterioles leading to the stomach, significantly reducing the amount of blood normally routed to the stomach, producing the nauseated feeling sometimes referred to as "butterflies." Norepinephrine affects the skin capillaries in the same manner, producing pallor in the face seen when one experiences severe fright, as well as coldness or clamminess of the hands and fingers due to the reduction in the volume of the blood in those extremities. The heart begins to pump blood harder and faster, increasing blood pressure, and pulse rate. Salivary glands in the mouth secretion change causing "dry mouth." There is a tensing of the involuntary muscles, in addition to constriction of the cardiovascular system, causing a tightening of the involuntary muscles in the stomach inhibiting diaphragm-intercostal muscular complex, causing less than average air intake at a time when the brain needs more than an average amount of oxygen because of increased mental activity. Consequently, stimulation of the respiratory muscles by the brain will also cause some breathing changes. Sweat glands are stimulated releasing perspiration, the iris of the eyes dilate, and contraction of the anal and urinary sphincters occurs, along with relaxation of the bladder (Matte, 1996).

When the test subject is presented with a threatening question, the previously described physiological reactions can occur. The polygraph instrumentation attached to their body records a variety of these changes.

There are numerous commercial companies supplying polygraph equipment (Lafayette Instrument, Axciton System, Stoelting, and Limestone Technologies). For about five or six thousand dollars, one can purchase a polygraph computerized system. The polygraph system uses four basic components to record the examinee's physiology. There are the thoracic and abdominal breathing devices consisting of two hollow corrugated tubes attached to a unit by a rubber hose and fastened around the subject's upper body with a beaded chain or Velcro® trap. This breathing or pneumo unit is low pressured and measures the inhalation/exhalation causing the tubes to expand and contract reflecting changes in the subject's breathing pattern (Matte, 1996).

The galvanic skin response component senses small changes in the skin resistance to electricity caused by the sweat glands activity in the bodies' fight or flight protective response to threat or danger. Galvanic skin conductance is measured by electrodes placed on the fingertips of the examinee's nondominant hand (Matte, 1996).

The fourth component is a cardio-sphygmograph, which measures blood pressure, rate, and strength of the pulse beat. The cardio-sphygmograph is a medical blood pressure cuff containing a rubber bladder that is wrapped around the upper arm against the brachial artery (Matte, 1996).

Before one can become a polygraph examiner, basic polygraph training is required. Accredited APA training generally consists of 12 weeks full-time residential training including theory and hands-on lessons with simulated cases, followed by several weeks of practical training with actual examinees. The topics covered include scientific foundations of polygraph, physiology, psychology, testing protocols, instrumentation, and interviewing and interrogation techniques. Regarding polygraph examination of sex offenders, APA By-Laws require a polygraph examination to be administered by a well-trained and competent polygraph examiner who has completed an additional 40 h of specialized instruction and certification training approved by APA on Post Conviction Sex Offender Testing (PCSOT). In addition, to maintain certification, 30 h of continuing education training is required every 2 years. Examiners conducting PCSOT are required to spend at least 15 h specific to the issues dealing with testing, treatment, or supervision of sex offenders.

The PCSOT examiner uses three screening tests: Maintenance Exam, Sex Offense Monitoring Exam, Sexual History Exam I—Victims, Sexual History Exam II—Compulsivity. The PCSOT diagnostic exams included: Instant Offense Exam—event-specific; Instant Offense Investigative Exam—multi-facet; Prior Allegation Exam—event-specific; and Prior Allegation Investigative Exam.

In practice, the types of polygraph examinations conducted in the PCSOT field are as follows (T. Tipton, personal communications July 23, 2007):

- Sex History examination: Covers several different activities and sexual behaviors, excluding the offense for which the examinee is on probation. Areas covered include past sexual habits, other victims, sexual deviance, sex abuse, physical abuse, alcohol/drug use, etc., during the examinee's lifetime.
- Disclosure examination: Specifically refers to the offense(s) for which the examinee is on probation. The test should be conducted if there is a significant discrepancy between the offender's version of the offense and the reported version of the offense. Used to assist in evaluating denial of offense.
- Maintenance/Monitoring examination: Refers to the period of time since examinee last took a polygraph examination, generally a 3–6 month period. Issues covered include compliance to probation/therapy rules, alcohol/drug use, contact with victim or minors, exposing or peeping, use of pornography, etc.
- Monitoring examination: Involves the commission of sexual offenses or other probation/therapy restrictions to a narrower line of questions. Focus on whether or not the offender had committed a sexual reoffense during the period of supervision.
- Specific issue/Incident examination: An exam concerning one issue or incident, identical to the disclosure in that it concerns one event, possibly travel out of state.

Although PCSOT coursework includes familiarity with the psychological issues relevant to sex offenders and some interviewing techniques (American International Institute of Polygraph, 2009), the polygraph examiner must acquire the knowledge and understanding of how a sex offender might think and feel about their sexual behavior and interests. The pretest and posttest interviewing skill required of a good examiner is an art that can be acquired with experience.

Working with the Polygraph Examiner

The polygraph examiner is an important member of the collaborative effort of a sex offender management strategy and must work closely with the community supervision officer and the sex offender treatment provider. Generally, the PCSOT are conducted in the polygraph examiners' office. However, this can create a time and information gap. Too often after information is revealed for the first time in a PCSOT and by the time the report reaches the clinician, the client has created a story to minimize the importance of the new information. Practically, what works best is when the supervision officer and clinician are both nearby during

the examination. This allows the examiner direct access to the officer and clinician and client to clarify issues and concerns that may arise during the pretest interview. Then, subsequent to the posttest interview, the officer and clinician appear in the examining room and hear directly from the client what he has disclosed to the polygraph examiner. The client still may attempt later to distort or somehow minimize any new revelations in his therapy group about his sexual behaviors and interests. However, the clinician is now better able to assist the client to remain on track.

Appling the PPG, AASI, and Polygraph in the Clinical Setting

Thus far, the history, science, and the application of the PPG, AASI, and Polygraph examination have been considered for the sex offender client in evaluative and clinical settings. Integrating these physiological tools is not a new suggestion, nor is it complicated. It was reported earlier that combining measures of sexual interest and arousal incrementally increases the validity of classifying individuals with sexual deviance. Thornton has also suggested combining the PPG and the polygraph to increase PPG sensitivity (Laws, 2009). Information obtained from polygraph examination can have obvious value to the individual's treatment as well as contributing to community safety. This information is enhanced by the obtaining data from the AASI and PPG. The combined results of these three procedures can be used to help the client better identify (and become more motivated to face) their problems associated with sexual deviance. It is suggested that the evaluator and clinician work closely with the client with deviant sexual interests, sharing findings reflective of sexual deviance on the AASI, which is likely to be denied, and then following that with information about deviant sexual arousal from the PPG. The next step requires the evaluator and clinician to communicate with the polygraph examiner about the individuals' testing results and have the polygraph investigation probe further about the individual's behavior. This layering and integration of information from AASI to PPG and polygraph can be expected to increase the client's treatment motivation. When an individual is candid about their internal experiences and past and recent behaviors, he can be encouraged and guided in treatment toward focusing on and overcoming psychological or criminogenic needs that contribute to the potential for sexual recidivism. Thus, the use and integration of psychophysiological assessment practices, over time, can provide an individual information and assistance in the development of a more balanced, nondeviant lifestyle, with the development of motivation, understanding, and skills to minimize the chances of sexual reoffending.

References

Abel, G. G., Becker, J. V., Mittleman, M., Cuningham-Rather, J., Rouleau, J. L., & Murphy, D. W. (1988). Self-reported sex crimes of nonincarcerated paraphiliacs. *Journal of Interpersonal Violence, 2*(1), 3–25.

Abel, G. G., Blanchard, E. D., & Barlow, D. H. (1981). Measurement of sexual arousal in several paraphilias: The effects of stimulus modality, instructional set and stimulus content on the objective. *Behavior Research and Therapy, 19*, 25–33.

Abel, G. G., Huffman, J., Warberg, B., & Holland, C. L. (1998). Visual reaction time and plethysmography as a measures of sexual interest in child molesters. *Sexual Abuse: A Journal of Researcher and Treatment, 10*(2), 81–95.

Abel, G. G., Jordan, A., Hand, C. G., Holland, L. A., & Phipps, A. (2001). Classification models of child molesters utilizing the Abel assessment for sexual interest. *Child Abuse & Neglect, 25*, 703–718.

Abel, G. G., Jordan, A., Rouleau, J. L., Emerick, R., Barboza-Whitehead, S., & Osborn, C. (2004). Use of visual reaction time to assess male adolescents who molest children. *Sexual Abuse: A Journal of Research and Treatment, 16*(3), 255–265.

Abel, G., Mittelman, M. S., & Becker, J. (1985). Sexual offenders: Results of assessment and recommendations for treatment. In H. H. Ben-Aron, S. I. Hucker, & C. D. Wester (Eds.), *Clinical criminology* (pp. 191–205). Toronto, ON: MM Graphics.

Adams, H. E., Motsinger, P., McAnulty, R. D., & Moore, A. L. (1992). Voluntary control of penile tumescence among homosexual and heterosexual subjects. *Archives of Sexual Behavior, 21*, 17–31.

Ahlmeyer, S., Heil, P., McKee, B., & English, K. (2000). The impact of polygraphy on admissions of victims and offenses in adult sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 12*(2), 123–138.

American International Institute of Polygraph. (2009, July). *Polygraph testing for the identification, treatment & monitoring of sex offenders. American Polygraph Association recognized course. Post conviction sex offender testing (PCSOT)*. Morrow, Georgia: American International Institute of Polygraphy.

American Polygraph Association: Polygraph Validity Research. Retrieved February 18, 2013, from http://www.polygraph.org/section/resources/polygraph-validity-research

American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct.* Retrieved February 16, 2013, from http://www.apa.org/ethics/code/index.aspx

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Text Revision (DSM-IV-TR). Washington, DC: American Psychiatric Association.

Association for the Treatment of Sexual Abusers. (2001). *Association for the treatment of sexual abusers: Practice standards and guidelines for evaluation, treatment, and management of adult male sexual abusers*. Beaverton, OR: Author.

Association for the Treatment of Sexual Abusers. (1993). *Auditory stimuli for penile plethysmography* (Cassette Recording DC No. 136C). Beaverton, OR: Author.

Association for the Treatment of Sexual Abusers. (2013). *Association for the treatment of sexual abusers: Practice standards and guidelines for evaluation, treatment, and management of adult male sexual abusers*. Beaverton, OR: Author.

Avery-Clark, C., & Laws, D. R. (1984). Differential erection response patterns of sexual child abusers to stimuli describing activities with children. *Behavior Therapy, 15*, 71–83.

Barbaree, H. E., Baxter, D. J., & Marshall, W. L. (1989). Brief research report: The reliability of the rape index in a sample of rapists and nonrapists. *Violence and Victims, 4*, 399–306.

Barbaree, H. E., Blanchard, R., & Kuban, M. (1999). A comparison of volume and circumference phallometry: Response magnitude and method agreement. *Archives of Sexual Behaviour, 28*(4), 345–360.

Barbaree, H. E., & Marshall, W. L. (1984). Deviant sexual arousal, offense history, and demographic variables as predictors of reoffense among child molesters. *Behavioral Sciences and the Law, 6*, 267–280.

Barker, J. G., & Howell, R. J. (1992). The plethysmograph: A review of recent literature. *Bulletin American Academy of Psychiatry and Law, 20*(1), 13–25. Electronic revision.

Baxter, D. J., Marshall, W. L., Barbaree, H. E., Davidson, P. R., & Malcolm, P. B. (1984). Deviant sexual behavior: Differentiating sex offenders by criminal and personal history, psychometric measures, and sexual response. *Criminal Justice and Behavior, 11*, 477–501.

Bellis, M. (2013). Police technology and forensic science history of the lie detector or polygraph machine. About.com Guide. Retrieved February 17, 2013, from http://inventors.about.com/od/fstartinventions/a/forensic_2.htm

Berthiaume v. Caron, 142 F.3d 12 (1st Cir. 1998).

Bradford, J. M. W., Kingston, D., Ahmed, G., & Fedoroff, J. P. (2010). Commentary: Sildenafil in phallometric testing—An evidence-based assessment of sexual offenders. *Journal of American Academic of Psychiatry and Law, 38*, 512–515.

Card, R. D., & Dibble, A. (1995). Predictive value of the Card/Farrall stimuli in discriminating between gynephilic and pedophilic sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 7*(2), 129–141.

Coric, V., Feuerstein, S., Fortunati, F., Southwick, S., Hemporini, H., & Morgan, C. A. (2005). Assessing sex offenders. *Psychiatry, 2*(11), 26–29.

Crooks, R. & Baur, K. (2008). *Our sexuality.* Retrieved February 10, 2013, from http://www.soc.ucsb.edu/sexinfo/article/the-sexual-response-cycle

DSM-5 Proposed Revisions. (2010). Retrieved February 10, 2010, from http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision

DSM-5 Workgroup Proposals. (2013). Retrieved January 18, 2013, from http://www.psych.org/File%20Library/Advocacy%20and%20Newsroom/Press%20Releases/2012%20Releases/12-43-DSM-5-BOT-Vote-News-Release--FINAL--3-.pdf

DSM-V Workshop Proposed Development. (2013). U 03 Sexual Masochism. Retrieved February 18, 2013, from http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=187#

Eccles, A., Marshall, W. L., & Barbaree, H. E. (1994). Differentiating rapists and non-offenders using the rape index. *Behaviour Research Therapy, 32*(5), 539–546.

Elaad, E., Ginton, A., & Jungman, N. (1992). Detection measures in real-life criminal guilty knowledge tests. *Journal of Applied Psychology, 77*(5), 7757–7767.

English, K., Jones, L., Pasini-Hill, D., Patrick, D., & Cooley-Towell, S. (2000). *The value of polygraph testing in sex offender management*. Lakewood, CO: National Institute of Justice, Colorado Department of Public Safety Division of Criminal Justice, Office of Research & Statistics.

Fedora, O., Reddon, J. R., Morrison, J. W., Fedora, S. K., Pascoe, H., & Yeudall, L. T. (1992). Sadism and other paraphilias in normal control and aggressive and nonaggressive sex offenders. *Archives of Sexual Behaviors, 21*, 1–31.

Fernandez, Y., & Marshall, W. L. (2003). Victim empathy, social self-esteem and psychopathy in rapists. *Sexual Abuse: Journal of Research and Treatment, 15*(1), 11–26.

Fischer, L., & Smith, G. (1999). Statistical adequacy of the Abel assessment for interest in paraphilias. *Sexual Abuse: A Journal of Research and Treatment, 11*(3), 195–205.

Freund, K., & Blanchard, R. (1989). Phallometric diagnosis of pedophilia. *Journal of Consulting and Clinical Psychology, 57*, 100–105. Electric version.

Freund, K., & Watson, R. J. (1991). Assessment of the sensitivity and specificity of a phallometric test: An update of phallometric diagnosis of pedophilia. Psychological assessment. *A Journal of Consulting and Clinical Psychology, 3*, 254–260.

Freund, K., Watson, R., Dickey, R., & Rienzo, D. (1991). Erotic gender differentiation in pedophilia. *Archives of Sexual Behavior, 20*, 555–566.

Freund, K., Watson, R., & Rienzo, D. (1988). Signs of feigning in phallometric test. *Behavior Research and Therapy, 26*, 105–112.

Frye v. United States (D.C. Cir. 1923) 293 Fed. 1013

Gray, S. R., & Plaud, J. J. (2005). A comparison of the Abel assessment for sexual interest and penile plethysmography in an outpatient sample of sexual offenders. *Journal of Sexual Offender Civil Commitment: Science and the Law, 1*, 1–10.

Green, R. (2002). Is pedophilia a mental disorder? *Archives of Sexual Behavior, 31*(6), 467–471.

Grubin, D., Madsen, L., Parsons, S., Susnowski, D., & Warberg, B. (2004). A prospective study on the impact of polygraph on high risk behavior on adult sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 16*(3), 209–222.

Grubin, D., & Madson, L. (2006). Accuracy and utility of post-conviction polygraph testing of sex offenders. *British Journal of Psychiatry, 188*, 479–483.

Haaven, J., & Schlank, A. (2001). The challenge of treating the sex offender with developmental disabilities. In A. Schlank (Ed.), *The sexual predator: Legal issues, clinical issues, special populations* (Vol. II, pp. 13-1–13-19). Kingston, NJ: Civic Research Institute.

Hall, G. C. N., Hirschman, R., & Oliver, L. L. (1995). Sexual arousal and arousability to pedophilic stimuli in a community sample of normal men. *Behavior Therapy, 269*(4), 681–694. Electric version.

Hanson, R. K., & Bussière, 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., & 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.

Harris, G. T., Rice, M. E., Quinsey, V. L., & Chaplin, T. C. (1996). Viewing time as a measure of sexual interest among child molesters and normal heterosexual men. *Behaviour Research and Therapy, 34*(4), 389–394.

Hecker, J. E., King, M. W., & Scoular, R. J. (2009). The startle probe reflex: An alternative approach to the measurement of sexual interest. In T. Ward & D. R. Laws (Eds.), *Sexual deviance: Issues and controversies* (pp. 82–102). Thousand Oaks, CA: Sage.

Heilbrun, K. (2003). Principles of forensic mental health assessment. Implications for the forensic assessment of sexual offenders. In R. A. Prentky, E. S. Janus, & M. C. Seto (Eds.), *Sexually coercive behavior, understanding and management* (pp. 167–184). New York: Annals of the New York Academy of Sciences.

Howe, R. (1995). A survey of plethysmographic assessment in North America. *Sexual Abuse: A Journal of Research and Treatment, 7*(1), 9–24.

Howe, R. J. (2003). Circumferential change scores in phallometric assessment: Normative data. *Sexual Abuse: A Journal of Research and Treatment, 15*(4), 365–375.

Iacono, W. G. (2008). Accuracy of polygraph techniques: Problems using confessions to determine ground truth. *Physiology & Behavior, 95*(1–2), 24–26. Electronic version.

Janssen, E., Everaerd, W., van Lunsen, R. H. W., & Oerlemans, S. (1994). Visual stimulation facilitates penile responses to vibration in men with and without erectile disorder. *Journal of Consulting and Clinical Psychology, 62*(6), 1222–1228.

Jensen, S. H., & Laws, D. R. (1994). *A clinician's guide to phallometry*. Brandon, VT: Safer Society Press. Videotape.

Kingston, D. A., Seto, M. C., Firestone, P., & Bradford, J. M. (2010). Comparing indicators of sexual sadism as predictors of recidivism

among adult male sexual offenders. *Journal of Consulting and Clinical Psychology, 78*(4), 574–584.
Knight, R. (2009). Is a diagnostic category for paraphilic coercive disorder defensible? *Archives Sex Behavior, 39*(2), 419–426. doi:10.1007/s10508-009-9571-x.
Kokish, R., Levenson, J. S., & Blasingame, G. D. (2005). Post-conviction sex offender polygraph examination: Client-reported perceptions of utility and accuracy. *Sexual Abuse: A Journal of Research and Treatment, 17*(2), 211–221.
Konopasky, R. J., & Konopasky, A. W. B. (2000). Remaking penile plethysmography. In D. R. Laws, S. M. Hudson, & T. Ward (Eds.), *Remaking relapse prevention with sex offenders: A sourcebook* (pp. 257–284). Thousand Oaks, CA: Sage.
Krueger, R. B., Bardford, J. M. W., & Glancy, G. D. (1998). Report from the committee on sex offenders: The Abel assessment for sexual interest—A brief description. *Journal of the American Academy of Psychiatry and the Law, 26*(2), 277–278.
Lalumière, M. L., & Harris, G. T. (1998). Common questions regarding the use of phallometric testing with sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 10*(3), 227–237.
Lalumière, M. L., & Quinsey, V. L. (1994). The discriminability of rapists from non-sex offenders using phallometric measures: A meta-analysis. *Criminal Justice and Behavior, 1*(150), 150–175.
Lalumière, M. L., Quinsey, V. L., Harris, G. T., Rice, M. E., & Trautrimas, C. (2003). Are rapists differentially aroused by coercive sex in phallometric assessments? In R. A. Prentky, E. S. Janus, & M. C. Seto (Eds.), *Sexually coercive behavior, understanding and management* (pp. 211–224). New York: Annals of the New York Academy of Sciences, 989.
Lanyon, R. I., & Thomas, M. L. (2008). Detecting deception in sex offender assessment. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 285–296). New York: Guilford Press. Electronic version.
Launay, G. (1999). The phallometric assessment of sex offenders: An update. *Criminal Behaviour and Mental Health, 9*(3), 254–274. Electronic version.
Laws, D. R. (1996). Marching into the past: A critique of Card and Olsen. *Sexual Abuse: Journal of Research and Treatment, 8*(4), 273–277.
Laws, D. R. (2003). Penile plethysmography: Will we ever get it right? In T. Ward & D. R. Laws (Eds.), *Sexual deviance: Issues and controversies* (pp. 82–102). Thousand Oaks, CA: Sage.
Laws, D. R. (2009). Penile plethysmography: strengths, limitation, innovations, cognitive approaches to the assessment of sexual interest in sexual offenders. In D. Thornton & D. R. Laws (Eds.), *Cognitive approaches to the assessment of sexual interest in sexual offenders* (pp. 7–29). West Sussex, UK: Wiley.
Laws, D. R., Gulayets, M. J., & Frenzel, R. R. (1995). Assessment of sex offenders using standardized stimulus procedures: A multi site study. *Sexual Abuse: A Journal of Research and Treatment, 7*, 45–66.
Laws, D. R., Hanson, L. K., Osborn, C. A., & Greenbaum, P. E. (2000). Classification of child molesters plethysmographic assessment of sexual arousal and a self-report measure of sexual preference. *Journal of Interpersonal Violence, 15*, 1297–1312.
Letourneau, E. J. (2002). A comparison of objective measures of sexual arousal and interest: Visual reaction time and penile plethysmography. *Sexual Abuse: A Journal of Research and Treatment, 14*, 207–224.
Levenson, J. S. (2004). Reliability of sexually violent predator civil commitment criteria in Florida. *Law and Human Behavior, 28*(4), 357–368.
Mangan, D. J., Armitage, T. E., & Adams, G. C. (2008). A field study on the validity of the Quadri-track zone comparison technique. *Physiology & Behavior, 95*, 17–23.
Mann, R. E., Hanson, R. K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the psychological meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment, 22*(2), 191–217.
Maram, W. B. (2005). Comparison of PPG and AASI in assessing offenders. *Perspective, California Coalition on Sexual Offending, 3*(1), 1.
Maram, W. B., & Koetting, M. G. (2004). *Phallometry and visual reaction time: Criterion-related validity using phallometric change scores and visual reaction time*. Unpublished raw data.
Marshall, W. L. (1996). Assessment, treatment, and theorizing about sex offenders. In D. R. Laws & W. O'Donohue (Eds.), *Sexual deviance: Theory, assessment, and treatment* (pp. 162–199). New York: Guilford.
Marshall, W. L. (2006a). *Clinical and research limitations in the use of phallometric testing with sexual offenders*. http://www.sexual-offender-treatment.org
Marshall, W. L. (2006b). Diagnostic issues, multiple paraphilias, and Comorbid disorder in sexual offenders: Their incident and treatment. *Aggression and Violent Behavior, 12*, 16–35.
Marshall, W. L., & Fernandez, Y. M. (2000). Phallometric testing with sex offenders: Limits to its value. *Clinical Psychology Review, 20*, 807–822.
Marshall, W. L., & Fernandez, Y. M. (2003). *Phallometric testing with sexual offenders: Theory, research, and practice*. Brandon, VT: Safer Society Foundation, Inc.
Matte, J. A. (1996). *Forensic psychophysiological: Using the polygraph: scientific truth verification—Lie detection*. Williamsville, NY: J. A. M. Publications.
McAnulty, R. D., & Adams, H. E. (1991). Voluntary control of penile tumescence: Effects of an incentive and a signal detection task. *Journal of Sex Research, 28*, 557–577.
McGrath, R. J., Cumming, G. F., Burchard, B. L., Zeoli, S., & Ellerby, L. (2009). *Current practice and emerging trends in sexual abuser management: The Safer Society 2009 North American survey*. Brandon, VT: The Safer Society Foundation, Inc.
McGrath, R. J., Cumming, G. F., & Burchard, B. L. (2003). *Current practices and trends in sexual abuser management: The Safer Society 2002 nationwide survey*. Brandon, VT: The Safer Society Foundation, Inc.
Murphy, W. D., & Barbaree, H. E. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology, 62*, 783–792.
National Clearinghouse on Family Violence. (1998). *Addressing the needs of developmentally delayed sex offenders, a guide.* Minister of Public Works and Government Services Canada.
Nichols, H. R., & Molinder, I. (1999). *Psychosexual life history (adult male form)*. Fircrest, WA: Nichols & Molinder Assessments, Inc.
North Carolina v. Spencer, 459 S.E. 2d 812, 815 (N.C. Ct. App. 1995).
O'Donohue, W., & Letourneau, E. (1992). The psychometric properties of penile tumescence assessment of child molesters. *Journal of Psychopathology and Behavior Assessment, 14*, 123–174.
O'Donohue, W., Regev, L. G., & Hagstrom, A. (2000). *Sexual Abuse: A Journal of Research and Treatment, 12*(2), 95–105.
Paraphilia Work Group. (2013). Parawww.dsm5.org/ProposedRevisions/Pages/proposedrevision DSM-V Proposed Develop. U 03 Pedophilic Disorder. Retrieved February 18, 2013, from http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=186#
People v. Stoll. (1989). 49 C3d 1136, 265 Cal Rptr 111.
Quinsey, V. L. (2010). Coercive paraphilic disorder. *Archives of Sexual Behavior, 39*, 405–410.
Quinsey, V. L., & Carrigan, W. F. (1978). Penile response to visual stimuli: Instructional control with and without auditory sexual fantasy correlates. *Criminal Justice and Behavior, 5*, 333–341.
Quinsey, V. L., Chaplin, T. C., & Carrigan, W. F. (1979). Sexual preferences among incest and nonincestous child molesters. *Behavior Therapy, 10*, 562–565.
Quinsey, V. L., Steinman, C. M., Bergersen, S. G., & Holmes, T. F. (1975). Penile circumference, skin conductance, and ranking

responses of child molesters and "normals" to sexual and nonsexual visual stimuli. *Behavior Therapy, 6*, 213–219.

R. v. J.-L.J., 2 S. C. R. 600 Canada Sup. Ct. (2000).

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.

Raskin, D., Barland, G. H., & Podlesny, J. A. (1976). *Validity and reliability of detection of deception: Final report (Government Study Contract # 75-NI-00-0001)*. Salt Lake City, Utah: National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice.

Rosen, R. C., & Keith, F. J. (1978). The measure of human penile tumescence. *Psychophysiology, 15*, 366–376.

Rosenzweig, S. (1942). The photoscope as an objective device for evaluating sexual interest. *Psychosomatic Medicine, 4*, 150–158. Electronic version.

Sachsenmaier, S. J., & Peters, J. M. (2002). Sexual offender risk assessment methods and admissibility as expert witness evidence. In J. M. Peters (Ed.), *Assessment and management of sex offenders: What prosecutors need to know*. Washington, DC: United States Department of Justice, Child Exploitation and Obscenity Section. Electric version.

Schouten, P. G. W., & Simon, W. T. (1992). Validity of phallometric measures with sex offenders: Comments on Quinsey, Laws, and Hall debate. *Journal of Consulting and Clinical Psychology, 60*(5), 812–814.

Schwartz, B.K., & Cellini, H.R. (1999). Sex offender recidivism and risk factors in the involuntary commitment process. In B. K. Schwartz (Ed.), The sex offender: Theoretical advances, treating special populations and legal developments (pp. 8-1-8-22). Kingston, NJ: Civic Research Institute.

Scientific Validity of Polygraph Testing: A Research Review and Evaluation. (1983). Washington, DC: U.S. Congress Office of Technology Assessment, 1983. Author. Retrieved February 2, 2012.

Sensitivity and Specificity. (2013). Retrieved January 17, 2013, from http://courses.ncssm.edu/math/Stat_Inst/Stats2007/Stat%20and%20Calc/Sensitivity%20and%20Specificity

Smith, G., & Fischer, L. (1999). Assessment of juvenile sexual offenders: Reliability and validity of the Abel assessment for interest in paraphilias. *Sexual Abuse: A Journal of Research and Treatment, 11*(3), 207–216.

The British Psychological Society, Professional Practice Board. (2008, September). *Penile plethysmography: Guidance for psychologists.* British Psychological Society.

U. S. v. Powers, 59 F.3d 1460 (4th Cir. 1995).

U.S. Congress Office of Technology Assessment. (1983). *Scientific validity of polygraph testing: A research review and evaluation: A technical memorandum.* November 1983 (OTA-TM-H-15), Washington, DC: Author.

U.S. v. Dotson, 324 F. 3rd 256, 02-4208 (4th Cir. 1984).

U.S. v. Herrera, 928 F.2d 769, 772 (6th Cir. 1991).

U.S. v. Silverman, 976 F.2d 1502, 1511 (6th Cir. 1992).

U.S. v. Stoterau, 07-50124 524 F.3d 988 (9th Cir. 2008).

U.S. v. Weber, 05-50191 (9th Cir. 2006).

Wilson, R. J. (1998). Psychophysiological signs of faking in phallometric test. *Sexual Abuse: Journal of Research and Treatment, 10*, 113–126.

Wilson, R. J., & Mathon, H. F. (2006). Remembering Kurt Freund (1914–1996). *ATSA Forum.* Beaverton, OR: Association for the Treatment of Sexual Abusers.

Wollert, R. (2007). Poor diagnostic reliability, the Null-Bayes logic model, and their implication for sexually violent predator evaluations. *Psychology Public Policy and Law, 13*(3), 167–203. Electronic version.

Wormith, J. S. (1986). Assessing deviant sexual arousal: Physiological and cognitive aspects. *Advances in Behavior Research and Therapy, 6*, 101–137.

Wright, L. W., & Adams, H. E. (1994). Assessment and sexual preference using a choice reaction time task. *Journal of Psychopathology and Behavioral Assessment, 16*, 211–231. Electronic version.

Zolondek, S. C., Abel, G., Northey, W. F., Jr., & Jordan, A. D. (2001). The self-reported behaviors of juvenile sexual offenders. *Journal of Interpersonal Violence, 16*(1), 73–85.

Zuckerman, M. (1971). Physiological measures of sexual arousal in humans. *Psychological Bulletin, 75*, 297–329.