

Phallometric Assessments of Sexual Interests: An Update

W. L. Marshall

Published online: 7 December 2013
© Springer Science+Business Media New York 2013

Abstract This paper reviews recent evidence on the value of phallometric (i.e., erectile) measures of sex offenders plus earlier crucial papers. Distinctions are made between types of measuring devices with the volumetric instrument appearing to be the most sensitive. Considering the meaning of responses it is concluded that arousal below 10 % of full erection is not reliably interpretable and that only deviant profiles (those displaying equal or greater arousal to deviant sex) can be confidently interpreted. The specificity and sensitivity of phallometry is satisfactory with child molesters but not with other types of sex offenders, although there remains a need to satisfactorily distinguish among subtypes of child molesters. There are enduring disagreements concerning the meaning of rapists' responses and no test has yet reliably identified deviance among exhibitionists. It is concluded that despite its long history, problems remain with phallometric assessments and clinicians should be cautious in interpreting the meaning of these responses.

Keywords Phallometry · Sex offenders · Standardization · Specificity · Sensitivity · Response magnitude

Introduction

The assessment and treatment of inappropriate sexual interests has been a fixed component of sex offender programs since McGuire, Carlisle and Young [1] first outlined their conditioning theory of the etiology and maintenance of sex offending. Accurate assessment of sexual interests is meant to serve a number of purposes including contributing to the determination

of current and likely future risk, and to identify treatment needs and treatment gains, to name just a few of its uses.

The early attempts at assessing sexual interests involved a range of strategies (see Zuckerman [2] for descriptions) but since the late 1950s measurement devices for male offenders have all been directed at monitoring changes in the penis in response to the presentation of various sexual images or scripts. Female sexual responses have also been measured typically by either vaginal probes employing light-emitting diodes or by thermistors tracking temperature changes in the labia (see Rosen & Beck [3]). This paper will address only the measurement of responses in males since that has been the group which the majority of studies have addressed. Furthermore the focus will only be on issues concerning devices that have measured erectile responses. Recently there has been considerable interest in alternative more cognitively-oriented approaches which show some promise (see Thornton & Laws [4]).

Devices

There are currently three devices used to determine the magnitude of erectile responses to appropriate and deviant stimuli: an instrument that measures changes in the volume of the penis; and two that measure only circumferential changes in the penis. The volumetric instrument was first introduced by Freund [5] but has had limited application outside Freund's laboratory. The most common circumferential device is a rubber strain gauge that contains within the hollow rubber tube one of various substances that readily change their electrical conductivity with changes in the penis. In early applications this substance was mercury which for health-risk reasons has been replaced by non-threatening materials such as indium-gallium or graphite. This rubber strain gauge was first introduced by Bancroft, Gwynne Jones and Pullan [6], and is currently the most widely used device. The alternative to the

This article is part of the Topical Collection on *Sexual Disorders*

W. L. Marshall (✉)
Rockwood Psychological Services, Kingston, Ontario, Canada
e-mail: bill@rockwoodpsyc.com

rubber strain gauge for assessing circumferential changes in the penis is a mechanical device initially designed by Barlow, Becker, Leitenberg and Agras [7]. However problems with this mechanical gauge have limited its popularity. In this paper all devices recording erectile change will be described as “phallometric” measures although other terms such as penile plethysmography (or PPG) are commonly used.

The evidence on the relative merits of these devices is limited, but the volumetric instrument seems to have advantages. For example, Earls and Marshall [8] demonstrated that the initial response of the penis to sexual arousal is a significant lengthening matched by a reduction in circumference as blood races into the penis. These initial changes in length were shown to account for 30 % of the length changes at full erection but these early changes are typically ignored by researchers and clinical evaluators. Unfortunately, a set of strain gauges developed by Earls and Marshall to simultaneously assess length and circumference changes, proved too difficult to manage for clinical purposes. They noted that volumetric measures tracked all changes in the penis including length and circumference, and they accordingly recommended such devices as the assessment strategy of choice due to their greater sensitivity. Although these recommendations may have provided comfort to those who use volumetric measures, they do not appear to have influenced anyone to drop their use of circumferential assessments. It must be said, however, that generally both these devices produce similar outcomes although the volumetric instrument is clearly more sensitive. Earls and Marshall [8] observed that circumferential gauges often miss changes at the high and low end of normal responding and that the client’s subjective experience of arousal typically fails to match changes in the penis.

Continuing Problematic Issues

Standardization

An enduring problem concerns the lack of standardization of the various aspects of phallometric assessments. In addition to different devices, there is no universally accepted set of stimuli to elicit arousal. Some centers use visual images (either still photos or moving images), while others use audio recordings of sexual encounters. Overall the evidence supports the use of visual images (still slides of naked people of various ages) for identifying pedophilic interests while audio descriptions appear most suited to assessing rapists. However, Fernandez [9] found that audio descriptions of sexual interests between adults and children, produced evidence of deviance across a broader range of child molesters than did still slides. Even when several centers use the same mode of presentation, the content of the stimuli is often idiosyncratic. There are numerous other differences across centers (see Howes [10] and

O’Donohue & Letourneau [11] for concerns about these differences) which has led to recommendations for multi-site standardization studies but these have been difficult to bring to fruition (see Laws, Gulayets, & Frenzel [12]).

The Meaning of Responses

One feature of the literature on phallometry is a clear lack of agreement on what constitutes a meaningful response. Some laboratories treat all responses as indicating interpretable arousal while others use an arousal level cutoff below which responses are treated as meaningless. Those who use a cutoff set different standards with some declaring that 1 mm of change in circumference is meaningful, others say 3 mm is the minimum arousal required for interpretation, while others claim that 6 mm is essential (see Howes [13], for a discussion). Recently in an evaluation of this issue, Lykins et al. [14] showed that establishing a cutoff of approximately 10 % of a full erection (i.e., approximately 3 mm circumferential change from baseline) results in a marked improvement in test-retest reliability which previous research had suggested was poor (see reviews by Marshall & Fernandez [15, 16]). This level of erectile response also represents the point at which “circumference becomes monotonically related to penile blood volume” (Lykins et al. [14], p. 54); that is, this is the point at which circumferential and volumetric measures produce a near perfect match (Kuban, Barbaree, Blanchard [17]). In addition to these observations, Blanchard et al. [18] demonstrated that a relative index of arousal (i.e., responses to deviant stimuli proportional to responses to normative stimuli) was a far more satisfactory way to identify deviance than simply examining the maximum magnitude of response to deviant stimuli.

In addition to these issues there is the problem of the proper interpretation of the resultant data. There are three potential profiles resulting from these assessments: 1) equal or greater arousal to deviant stimuli than to non-deviant images; 2) greater arousal to non-deviant stimuli; or 3) no arousal to any stimuli. The latter occurs in approximately 20 % to 25 % of all subjects (although rarely reported in studies) and should be considered an uninterpretable profile. The second type of response profile cannot be considered as indicative of normal sexual interests since all men, including chronic offenders, can control their responses so as to appear normal (Freund, Watson, & Rienzo [19]; Laws & Holmen [20]; Wydra, Marshall, Earls, & Barbaree [21]). In fact, it is only the first profile (i.e., greater arousal to deviant images/scripts) that can be confidently interpreted.

One of the continuing problems that besets this field has been the inability to provide a standard against which to compare the results of phallometric testing. This would seem to be essential to establish the validity of the results for each of the different groups of offenders. For example, the results

might be compared to a carefully established diagnosis. According to this strategy, all child molesters diagnosed as pedophiles, and only those so diagnosed, should display arousal to children that is equal to, or greater than, their arousal to adults. Unfortunately the inter-clinician reliability of the Diagnostic and Statistical Manual of Mental Disorders diagnoses has been called into question (see reviews by Marshall [22, 23]; Marshall & Marshall [24]; O'Donohue, Regev & Hagstrom [25]). In fact Freund & Blanchard [26] viewed the phallometric test itself as a diagnostic tool which, while having some merit, leaves researchers seeking to establish the psychometric properties of phallometry with no obvious recourse. Despite these reservations, phallometry remains the best approach to establishing a sex offender's sexual interests and, so long as the resultant data are interpreted with caution, they can usefully contribute to risk assessment and to defining targets for treatment.

Sensitivity and Specificity

A final general issue concerns the specificity and sensitivity of phallometry. The accurate establishment of these critical features is essential to any type of test if it is to meet psychometric standards. In the case of phallometry, sensitivity refers to the accuracy with which the test identifies the problematic population (i.e., sex offenders of a particular type) while specificity indicates the accuracy with which phallometry identifies healthy nonproblematic men. Unfortunately these issues are not as straightforward as they may seem. For example, no one expects all sex offenders to have enduring fantasies or arousal to deviant images and, indeed, that is what the research shows. Similarly a comparison group of non-offenders may include some men who have unreported victims or who have deviant fantasies that they have not yet acted on. As a result we would not expect all members of a comparison group to display strictly normative responses.

Wormith [27] found phallometry to correctly classify 64 % of his subjects and Day, Miner, Sturgeon and Murphy [28] found 74 % classification accuracy. In a study using extensive data, Harris, Rice, Quinsey, Chaplin and Earls [29] reported unusually high classification accuracy. They were able to correctly classify 83 % of child molesters and 87 % of rapists using audio stimuli as well as correctly classifying 79 % of child molesters using slides of naked children and adults. While these data are quite impressive, Fernandez [9] found lower accuracy. She reported 69 % correct classification of non-familial child molesters and 51 % of incest offenders using visual stimuli and accuracy rates of 67 % with nonfamilial child molesters and 75 % with incest offenders using audio descriptions. For rapists, Fernandez reported that phallometric depictions of sexual violence against an adult woman, was only able to correctly classify 31.7 % of the offenders but more damaging was the incorrect classification of 45 % of child molesters as rapists using these stimuli. Barbaree and Peacock [30] also reported

poor sensitivity for an Age-Gender assessment as did Malcolm, Andrews and Quinsey [31].

Thus with the exception of the Harris et al. [29] study, the evidence on the specificity and sensitivity of phallometric testing, does not encourage confidence but only if the expected standard is that all sex offenders should be shown to have deviant interests and all non-offending males should display normative interests. There is clear evidence that neither of these expectations is true. Thus the problems associated with the accurate classification of sex offenders and nonoffenders will remain unresolved until a gold standard of what should be expected is established. This seems unlikely to ever be realized.

Recent Research Findings

Despite the above noted concerns, phallometry still enjoys some degree of popularity for the assessment of various types of sexual offenders. Research on the phallometric assessments of sex offenders continues although the number of publications employing this measurement approach appears to be diminishing as the interest in deviant arousal patterns has generally ebbed (see Laws [32]). For those researchers who still evaluate deviant tendencies, alternative ways of evaluating these interests have been utilized (Gress & Laws [33]). The evidence on the value of these more cognitively-based approaches, however, is not yet sufficiently strong enough to justify abandoning phallometry.

The following section reviews the findings with the major types of sex offenders. In all cases there continues to be issues that need empirical resolutions but, when carefully done, phallometric appraisals can contribute to a clearer understanding of individual sex offenders.

Child Molesters

It is typical among researchers and clinicians to distinguish incest offenders from men who molest nonfamilial children. This distinction, however, has not proved to be as useful as was once thought. First the definitions attached to both groups are inconsistent across studies with some researchers including in their definition of incest, not only fathers, but also grandfathers, stepfathers, adoptive fathers, cousins, uncles, and the mother's live-in male partner. Other studies identify only men acting as fathers while allocating the rest to the broad category of nonfamilial child molesters. This latter strategy means that all offenders who are distantly related to the child, as well as those who are involved in caregiving activities with their victim (e.g., teachers, sports coaches, clergy) are all placed in the same category as "stranger-victim" offenders. Marshall and Marshall [34] have suggested that a better distinction would be between men who molest children whom they know and for whom the offender is in a care-taker role, and those who seek out children unknown to the abuser.

Marshall and Marshall describe the first group as “affiliative” child molesters and the stranger-victim offenders as “non-affiliative”. Since true stranger-victim offenders typically constitute less than 10 % of any group of child molesters and appear to abuse significantly more victims than affiliative offenders, we can expect them to be the most deviant. Freund ([35]; Freund, Chan & Coulthard [36]; Freund & Watson [37]) early on showed that among child molesters, it was only those who admitted to having offended and who had three or more victims, who displayed deviant response profiles (i.e., equal or greater arousal to children than to adults).

In a promising strategy, Barbaree and Marshall [38] examined the individual profiles of a mixed sample of female-victim child molesters and non-offending males. They reported five distinctive profiles: 1) greatest arousal to adults; 2) equal arousal to adults and teenagers; 3) a nondiscriminating profile; 4) strong arousal to children with little arousal to adults; and 5) an odd profile of strong arousal to children, little or no arousal to teenage girls, and strong arousal to adults. Twenty-seven percent of non-familial child molesters displayed profile four; 10 % showed either profile 1, 3 or 5; and 20 % displayed the teen-adult profile (i.e., profile two). This profile analysis approach seems more clinically relevant than the usual method of simply calculating a deviance index (i.e., arousal to children as a proportion of arousal to adults). Recently, Michaud and Proulx [39] conducted a similar profile analysis of phallometric data from child molesters and found profiles that were quite similar to those described by Barbaree and Marshall [38].

Rapists

Disagreements across research findings continue in the phallometric evaluation of men who rape women. Lalumière and Quinsey’s [40] review reported highly significant and consistent differences between rapists and non-rapists with the rapists showing greater arousal to scenes of a violent rape on a nonconsenting woman than they did to a consenting scene. The majority of studies in this review, however, evaluated rapists housed in psychiatric facilities for offenders found to be not criminally responsible. Barbaree, Marshall and Lanthier [41] and Baxter, Marshall, Barbaree, Davidson and Malcolm [42], evaluating a federal prison population, found that while rapists showed greater arousal to rape than did other men, the peak responses of the offenders to rape were somewhat lower than their responses to consenting sex. From these data, Barbaree [43] concluded that the rapists in these latter studies simply failed to be inhibited, as were the nonoffenders, by the presence of force and nonconsent in the rape scenes. Lalumière and Quinsey [40] on the other hand, interpreted their data as indicating that the sexual arousal of rapists is incited by these coercive and violent elements.

Since the studies by both these groups were scientifically sound, there appears to be just two possible sources of the different findings. It may be that the violence and brutality of the rape depictions in the Lalumière and Quinsey stimulus set is critical, as Lalumière and Quinsey maintain, or the differences might simply reflect the different populations in the two sets of studies. Looman ([44]; Looman & Marshall [45]) tested these possibilities by examining the responses of prison-based rapists and other offenders to both stimulus sets. Looman found that the two sets elicited approximately equivalent responses among the rapists. In response to both stimulus sets these prison-based offenders produced arousal patterns that were essentially a match for the Barbaree data; that is, the rapists’ responses to both sets failed to reveal a preference for coercive, nonconsenting sex.

The explanation for these findings compared to Lalumière and Quinsey’s (1994) observations, that was offered by Looman ([44]; Looman & Marshall [45]), points to sample differences. Looman suggested that the offenders in the psychiatric facilities were likely to be both more disturbed and more vicious in their assaults, and Marshall and Fernandez [16] added that the Lalumière and Quinsey groups may have included a disproportionate number of sexual sadists. If there is any substance to these claims, then we can expect rapists assessed in prisons or at community-based clinics, to be less deviant than these in long-term psychiatric facilities. In Barbaree’s ([43]; Barbaree et al. [38]) early reports, only 32 % of the rapists could be classified as deviant and this dropped to 25 % in Looman and Marshall’s [45] report. Assessors should be careful interpreting the meaning of rapists’ responses given that Marshall and Fernandez [16] report data indicating that the responses of 26 % of carefully selected nonoffenders produced deviance indexes that matched those of rapists. It is appropriate, therefore, to be cautious in interpreting the meaning of a rapist’s response profile, particularly when arousal to rape and consenting sex are similar.

Sexual Sadists

In considering the results of phallometric assessments of sexual sadists it is first necessary to note the serious reliability problems associated with the diagnosis of these problematic offenders (see [46] for a summary of the evidence). Most studies have simply employed the same stimulus sets used to appraise rapists, although some of these sets, like those in Lalumière and Quinsey’s [40] review, involve depictions of very violent attacks which might be expected to elicit arousal from sadists.

Proulx, Blais, and Beauregard [47] added to their sexual assault scenes not only physical violence but also, quite crucially for sadists, elements depicting the assailant saying and doing things to the victim that reflect clear attempts to humiliate the woman. This fits with the *Diagnostic and Statistical Manual’s* (DSM-5) (American Psychiatric Association [48])

definition of sadism. Proulx et al. found that this combination of violent and humiliating elements elicited strong arousal among sexual sadists but not among non-sadistic rapists. Marshall, Hucker, Nitschke and Mokros [49] recommended the use of Proulx et al.'s stimuli for the assessment of sadists but they noted that access to crime scene data would strengthen confidence in the diagnosis. Nitschke, Osterheider and Mokros [50] have generated a scale for sexual sadism that attends to both crime scene data and phallometric results. This scale has been shown to meet all psychometric requirements and is to be preferred over phallometry alone in establishing a diagnosis of sexual sadism.

Exhibitionists

A phallometric test for exhibitionists was first described by Kolarsky and Madlafousek [51] but these researchers did not employ ecologically valid stimuli. Some subsequent studies used more appropriate depictions but found no differences between exhibitionists and other men (Freund, Scher, & Hucker [52, 53]; Langevin et al. [54]; Murphy, Abel, & Becker [55]).

Marshall, Payne, Barbaree and Eccles [56] designed audiotapes that portrayed either exposing or consenting sex with each occurring in three different locations: the man's apartment, his car, or a secluded park. These settings matched those where the men in the study reported having exposed themselves. Although these chronic exhibitionists displayed somewhat greater arousal than did normal men to exposing scenes and less to consenting scenes, they nevertheless showed far greater arousal to normative sex (42.5 % full erection) than they did to exposing (19.2 % full erection). There have been no studies since that time that have employed similarly ecologically sound stimuli. When Marshall et al. asked the exhibitionists in their study what it was they fantasized during masturbation, these offenders indicated that they imagined their victim being sufficiently impressed and, as a result, inviting them to repair to the woman's house for consensual sex. These offenders indicated that they realized this was an unlikely outcome but fantasy knows no such prosaic limits. Whether stimuli designed according to these fantasies will prove consistently effective with exhibitionists has not yet been tested.

Conclusions

There remain ongoing concerns regarding the psychometric status of phallometric assessments. In response to these observations it has been argued (O'Donohue & Letourneau [11]) that phallometry is better viewed as a behavioral evaluation rather than as a psychometric test and in that view, the standards applicable to such tests are irrelevant. However many centers continue to utilize phallometry to determine the presence or absence of deviant sexual interests in sex offender

clients and research reports continue to be integrated as if all studies are utilizing a standardized and psychometrically sound procedure.

Despite these reservations it is possible to attempt an integration of data from across sites. What this integration reveals is that the results with child molesters appear reasonably consistent in showing that some, but far from all, of these offenders display deviant arousal that requires treatment intervention. Deviant interest displayed by child molesters is also a moderate predictor of future risk to reoffend, so the value of phallometric assessments with these offenders appears to be clear. This, however, does not appear to be true with rapists where controversy over the results has divided the field. With sexual sadists there have been recent promising advances that await further development and research. The research on exhibitionists does not appear to have produced an acceptable phallometric evaluation procedure.

Overall the data suggest some value to phallometric assessments but caution in the interpretation of results is recommended. In particular an apparently "normative" profile should not be viewed as indicating the client does not have persistent deviant interests since offenders and nonoffenders alike can readily control their erectile responses and these manipulations cannot be easily discerned. While Freund, Watson and Rienzo [19] were able to identify attempts by some sex offenders to alter their responses, the only attempts they could reasonably consistently discern were those where the client employed physical manipulations. These manipulations involved voluntary contractions of the muscles of the perineum. Unfortunately these tactics appear to be relatively uncommon with the most common by far involving cognitive strategies such as distraction [57] which, of course, cannot be readily observed.

Hopefully future research will lead to a refinement of phallometric evaluations although they may yet be replaced by the recent emergence of more cognitively-based evaluations.

Compliance with Ethics Guidelines

Conflict of Interest W.L. Marshall declares that he has no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by the author.

References

- McGuire RJ, Carlisle JM, Young BG. Sexual deviations as conditioned behaviour: A hypothesis. *Behav Res Ther.* 1965;2:185–90.
- Zuckerman M. Physiological measures of sexual arousal in the human. *Psychol Bull.* 1971;75:297–329.

3. Rosen RC, Beck JG. Patterns of sexual arousal: Psychophysiological processes and clinical applications. New York: Guilford Press; 1988.
4. Thornton D, Laws DR. Cognitive approaches to the assessment of sexual interest in sexual offenders. Chichester: John Wiley & Sons; 2009.
5. Freund K. Diagnostika homosexuality u muszu. *Ceskoslovak Med*. 1957;53:382–93.
6. Bancroft JH, Gwynne Jones HC, Pullan BP. A simple transducer for measuring penile erections with comments on its use in the treatment of sexual disorders. *Behav Res Ther*. 1966;4:239–41.
7. Barlow DH, Becker R, Leitenberg H, Agras WS. A mechanical strain gauge for recording penile circumference change. *J Appl Behav Anal*. 1970;3:73–6.
8. Earls CM, Marshall WL. The simultaneous and independent measurement of penile circumference and length. *Beh Res Methods and Instrum*. 1982;14:447–50.
9. Fernandez YM. Phallometric testing with sexual offenders: An examination of reliability and validity issues. Doctoral thesis, Queen's University, Canada; (2001).
10. Howes RJ. A survey of plethysmographic assessment in North America. *Sex Abuse A J Res Treat*. 1995;7:9–24.
11. O'Donohue WT, Letourneau W. The psychometric properties of the penile tumescence assessments of child molesters. *J Psychopathol Behav Assess*. 1992;14:123–74.
12. Laws DR, Gulayets JJ, Frenzel RR. Assessment of sex offenders using standardized slide stimuli and procedures: A multisite study. *Sex Abuse A J Res Treat*. 1995;7:45–66.
13. Howes RJ. Circumferential change scores in phallometric assessment: Normative data. *Sex Abuse A J Res Treat*. 2003;15:365–75.
14. Lykins AD, Cantor JM, Kuban ME, Blak T, Dickey R, Klassen PE, et al. Sexual arousal to female children in gynephilic men. *Sex Abuse A J Res Treat*. 2010;23:279–89.
15. Marshall WL, Fernandez YM. Phallometric testing with sexual offenders: Limits to its value. *Clin Psychol Rev*. 2000;20:807–22.
16. Marshall WL, Fernandez YM. Phallometric testing with sexual offenders: Theory, research, and practice. Brandon: Safer Society Press; 2003.
17. Kuban M, Barbaree HE, Blanchard R. A comparison of volume and circumference phallometry: Response magnitude and method agreement. *Arch Sex Behav*. 1999;28:345–59.
18. Blanchard R, Kuban ME, Blak T, Cantor JM, Klassen PE, Dickey R. Absolute versus relative ascertainment of pedophilia in men. *Sex Abuse A J Res Treat*. 2009;21:431–41.
19. Freund K, Watson RJ, Rienzo D. Signs of feigning in the phallometric test. *Behav Res Ther*. 1988;26:105–12.
20. Laws DR, Holmen ML. Sexual response faking by pedophiles. *Crim Justice Behav*. 1978;5:343–56.
21. Wydra A, Marshall WL, Earls CM, Barbaree HE. Identification of cues and control of sexual arousal by rapists. *Behav Res Ther*. 1983;21:469–76.
22. Marshall WL. Pedophilia: Psychopathology and theory. In: Laws DR, O'Donohue W, editors. *Sexual deviance: Theory, assessment, and treatment*. New York: Guilford Press; 1997. p. 152–74.
23. Marshall WL. Diagnostic issues, multiple paraphilias, and comorbid disorders in sexual offenders: Their incidence and treatment. *Aggress Violent Behav*. 2007;12:16–35.
24. Marshall WL, Marshall LE. The paraphilias. In: Hofmann SG, editor. *The Wiley handbook of cognitive behavioral therapy*. Chichester: John Wiley & Sons; 2014. p. 763–91.
25. O'Donohue WT, Regev LG, Hagstrom A. Problems with the DSM-IV diagnosis of pedophilia. *Sex Abuse A J Res Treat*. 2000;12:95–105.
26. Freund K, Blanchard R. Phallometric diagnosis of pedophilia. *J Consult Clin Psychol*. 1989;57:1–6.
27. Wormith JS. Assessing deviant sexual arousal: Psychological and cognitive aspects. *Adv Behav Res Ther*. 1986;8:101–37.
28. Day D, Miner M, Sturgeon V, Murphy J. Assessment of sexual arousal by means of physiological and self-report measures. In: Laws DR, editor. *Relapse prevention with sex offenders*. New York: Guilford Publications; 1989. p. 115–23.
29. Harris GT, Rice ME, Quinsey VL, Chaplin TC, Earls CM. Maximizing the discriminant validity of phallometric assessment data. *Psychol Assess A J Consult Clin Psychol*. 1992;4:502–11.
30. Barbaree HE, Peacock EJ. Phallometric assessment of sexual preferences as an investigative tool in cases of alleged child abuse. In: Ney T, editor. *Allegations of child sexual abuse: Assessment and case management*. New York: Brunner/Mazel; 1995. p. 242–59.
31. Malcolm PB, Andrews DA, Quinsey VL. Discriminant and predictive validity of phallometrically measured sexual age and gender preference. *J Interpersonal Violence*. 1993;8:486–501.
32. Laws DR. Penile plethysmography: Will we ever get it right? In: Ward T, Laws DR, Hudson SM, editors. *Sexual Deviance: Issues and controversies*. Thousand Oaks: Sage Publications; 2001. p. 82–102.
33. Gress CLZ, Laws DR. Measuring sexual deviance: Attention-based measures. In: Beech AR, Craig LA, Browne KD, editors. *Assessment and treatment of sex offender*. Chichester: John Wiley & Sons; 2009. p. 109–28.
34. Marshall WL, Marshall LE. Affiliative and nonaffiliative child molesters: A suggestion for new subcategories. *J Sex Aggress*. (in press)
35. Freund K. Reflections on the development of the phallometric method of assessing sexual preferences. *Ann Sex Res*. 1991;4:221–8.
36. Freund K, Chan S, Coulthard R. Phallometric diagnoses with “nonadmitters”. *Behav Res Ther*. 1979;17:451–7.
37. Freund K, Watson RJ. Assessment of the sensitivity and specificity of a phallometric test: An update of phallometric diagnosis of pedophilia. *Psychol Assess A J Consul Clin Psychol*. 1991;3:254–60.
38. Barbaree HE, Marshall WL. Erectile responses amongst heterosexual child molesters, father-daughter incest offenders, and matched nonoffenders: Five distinct age preference profiles. *Can J Behav Sci*. 1989;21:70–82.
39. Michaud P, Proulx J. Penile-response profiles of sexual aggressors during phallometric testing. *Sex Abuse A J Res Treat*. 2009;21:308–34.
40. Lalumière ML, Quinsey VL. The discriminability of rapists from non-sex offenders using phallometric measures: A meta-analysis. *Crim Justice Behav*. 1994;21:150–75.
41. Barbaree HE, Marshall WL, Lanthier RD. Deviant sexual arousal in rapists. *Behav Res Ther*. 1979;8:229–39.
42. Baxter DJ, Marshall WL, Barbaree HE, Davidson PR, Malcolm PB. Deviant sexual behavior: Differentiating sex offenders by criminal and personal history, psychometric measures, and sexual responses. *Crim Justice Behav*. 1984;11:477–501.
43. Barbaree HE. Stimulus control of sexual arousal. In: Marshall WL, Laws DR, Barbaree HE, editors. *Handbook of sexual assault: Issues, theories, and treatment of the offender*. New York: Plenum Press; 1990. p. 115–42.
44. Looman J. Sexual arousal in rapists as measured by two stimulus sets. *Sex Abuse A J Res Treat*. 2000;12:235–48.
45. Looman J, Marshall WL. Sexual arousal in rapists. *Crim Justice Behav*. 2005;32:367–89.
46. Marshall WL, Yates PM. Diagnostic issues in sexual sadism among sexual offenders. *Journal of Sex Aggress*. 2004; 10:21–27. Marshall, W. L., & Marshall, L. E. (in press).
47. Proulx J, Blais E, Beauregard E. Sadistic sexual aggressors. In: Marshall WL, Fernandez YM, Marshall LE, Serran GA, editors. *Sexual offender treatment: Controversial issues*. Chichester: John Wiley & Sons; 2006. p. 61–77.
48. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: Author; 2013.
49. Marshall WL, Hucker SJ, Nitschke J, Mokros A. Assessment of sexual sadism. In L. A. Craig & M. Rettenberger (Eds., *The Wiley-Blackwell Handbook on the assessment, treatment and theories of*

- sexual offending (Volume: Assessment). Chichester, UK: John Wiley & Sons. (in press).
50. Nitschke J, Osterheider M, Mokros A. A cumulative scale of severe sexual sadism. *Sex Abuse A J Res Treat.* 2009;21:262–78.
 51. Kolarsky A, Madlafousek J. The inverse role of preparatory erotic stimulation in exhibitionists: Phallometric studies. *Arch Sex Behav.* 1983;12:123–48.
 52. Freund K, Scher H, Hucker S. The courtship disorders. *Arch Sex Behav.* 1983;12:369–79.
 53. Freund K, Scher H, Hucker S. The courtship disorders: A further investigation. *Arch Sex Behav.* 1984;13:133–9.
 54. Langevin R, Paitich D, Ramsey G, Anderson C, Kamrad J, Pope S, et al. Experimental studies in the etiology of genital exhibitionism. *Arch Sex Behav.* 1979;8:307–31.
 55. Murphy WD, Abel GG, Becker JB. Future research issues. In: Cox DJ, Daitzman RJ, editors. *Exhibitionism: Description, assessment, and treatment.* New York: Garland STPM Press; 1980. p. 339–92.
 56. Marshall WL, Payne K, Barbaree HE, Eccles A. Exhibitionists: Sexual preferences for exposing. *Behav Res Ther.* 1991;29:37–40.
 57. Golde JA, Strassberg DB, Turner CM. Psychophysiological assessment of erectile responses and its suppression as a function of stimulus media and previous experience with plethysmography. *J Sex Res.* 2000;37:53–59.