



# Diagnostic Issues in the Paraphilias

William L. Marshall<sup>1</sup> · Drew A. Kingston<sup>2</sup>

Published online: 21 July 2018  
© Crown 2018

## Abstract

**Purpose of Review** To examine the empirical bases underlying the diagnoses of the paraphilias. We address issues concerning the reliability of these diagnoses and their implications for etiology, treatment, and prognosis.

**Recent Findings** Research on these issues with the paraphilias is quite limited except for those paraphilics whose interests lead them to sexually offend. Even among these clients, research has, for the most part, failed to distinguish those who meet criteria for a paraphilia from those who do not, thereby limiting the possibility of drawing firm conclusions regarding the value of a paraphilic diagnosis.

**Summary** Speculations regarding the etiology of the paraphilias are for the most part limited to those who sexually offend and these theories do not distinguish those who do, or do not, meet paraphilic criteria. Treatment of sex offenders, when effective, appears to have the same impact regardless of whether or not clients meet criteria for a paraphilia. In terms of prognosis, it was only among untreated child molesters that a paraphilic diagnosis (in this case “pedophilia”) predicted long-term outcome. In the face of these problems, we suggest a dimensional approach to diagnoses may represent an improvement over the current categorical model.

**Keywords** Diagnosis · Paraphilias · Pedophilia · Sexual sadism · Etiology · Treatment · Prognosis

## Introduction

The paraphilic diagnoses have appeared in all editions of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) of the American Psychiatric Association (APA). However, in the first two editions [1, 2], the paraphilias were placed in the section on “Personality Disorders.” In DSM-III [3], they were moved to their own section where they have been ever since. From DSM-III onward through the course of subsequent editions, the criteria have been changed progressively with the defining features becoming more behav-

ioral. This has gone some way in reducing the vagueness of the criteria but not completely eliminating the need for diagnosticians to interpret the meaning of several aspects of the putative characteristics. This paper will attempt to address these and other concerns regarding the paraphilic diagnoses.

First, however, a brief description will be provided of the main aspects of the diagnostic criteria as specified in the current edition of the DSM [4]. These notes will be restricted to those paraphilias (i.e., pedophilia and sexual sadism) that involve criminal behaviors as these are the disorders for which there is the greatest amount of relevant evidence. The comments on these specific diagnoses, however, should be seen as pertinent to the other paraphilias (for more detailed discussions of the other paraphilias see [5, 6•]).

It should be noted that in much of the literature on child molesters, the term “pedophilia” is used as a general descriptor without due regard to diagnostic criteria [7]. This is also how the popular media tend to describe child molesters. Such casual applications of the term, particularly by professionals, are not helpful if we accept that a distinction needs to be made. As will be suggested in this

---

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

✉ William L. Marshall  
bill@rockwoodpsyc.com

Drew A. Kingston  
drew.kingston@theroyal.ca

<sup>1</sup> Queen’s University and Rockwood Psychological Services, Kingston, Canada

<sup>2</sup> Integrated Forensic Program, Royal Ottawa Health Care Group and Institute of Mental Health Research, Ottawa, Canada

paper, however, the distinction may not carry much empirical weight.

## DSM-5 Paraphilic Diagnoses

According to DSM-5 [4], a paraphilia “denotes any intense and persistent sexual interest other than sexual interests in genital stimulation or preparatory fondling with phenotypically normal, physically mature, consenting human partners [4 (p685)]. A paraphilia is distinguished from a paraphilic disorder where the latter is described as “concurrently causing distress or impairment to the individual (and) whose satisfaction has entailed personal harm, or risk to harm to others” [4 (pp685–686)]. Both criteria must be met in order for a diagnosis of paraphilic disorder to be applied, and once applied, the disorder demands intervention. Paraphilias, on the other hand, only require treatment if the clients request it.

Not surprisingly, many sex offenders, including those who admit to having committed an offense, deny having any enduring interest in the activities involved in their crimes, which is to say they deny having a paraphilic disorder. DSM-5 offers a solution in such cases when it declares that if a person denies these interests, the diagnosis may be applied when there is “substantial objective evidence to the contrary” [4 (p696)]. Unfortunately, exactly what constitutes “substantial evidence” is not made clear.

## Issues in the Paraphilic Diagnoses

There are many issues that could be, and have been, raised concerning the value and precision of DSM diagnoses. Marshall [8–10] noted several issues of concern with DSM criteria, particularly the various purported defining features. Other authors have made similar challenges by detailing the lack of precision in the criteria [11, 12]. In addition to the vagueness of the criteria, the diagnoses are not accompanied by specific implications for treatment [10]. Also, the paraphilic diagnosis of pedophilia appears not to predict future risk to reoffend [13–15]. In addition, as will be illustrated later in this paper, paraphilic diagnoses have been shown to lack the necessary levels of reliability to function as useful guides even for diagnostic purposes. Moreover, the structural nature of the DSM can be challenged with the question being “Is the DSM simply a nomenclature (i.e., a system of naming) or does it meet the more rigorous standards of a nosology?” In fact in the introductory remarks to DSM-5, the authors suggest it is both. This leaves open the possibility of examining the nosological status of the manual.

## Nosological Issues

The value of a nosological system, over that of a nomenclature, is that it should have clear implications for the etiology, treatment, and future course of each of the specified disorders [16–18]. It might, therefore, be fruitful to explore how relevant the paraphilic diagnoses are for etiology, treatment, and long-term outcome (i.e., prognosis). Among those working in the field of sexual offending, theorists have attempted to explain the origins of paraphilic behaviors, treatment providers have outlined approaches addressing these problems, and researchers have identified characteristics that predict future risk to reoffend. Clearly, practitioners want answers to these nosological questions.

There are numerous theories purporting to explain the etiology of various types of sexual offending [19••]. These theories, for the most part, focus on the post-birth social experiences of those children who become sex offenders, particularly problematic care-giver attachments and experiences of being sexually abused when they were themselves children. However, none of these theories are restricted to those offenders who later meet criteria for a paraphilia.

Two factors these theories neglect concern pre-birth issues (i.e., genetic and inter-uterine factors) and post-birth head trauma, each of which has been implicated in the genesis of pedophilia. Cantor and his colleagues [20, 21] have shown that pedophiles have lower intelligence and relative memory impairments and are shorter in stature than are matched non-pedophilic men. These researchers also found that pedophiles are more likely to be left-handed than other men. Studies by Blanchard [22, 23] indicate that a disproportionate number of pedophiles suffered traumatic head injuries during their developmental years. More recent studies using magnetic resonance imaging techniques [24] have also shown differences in brain structure, particularly lower white matter volume among pedophilic men. This has led to the suggestion that pedophiles may more readily interpret environmental stimuli as sexually relevant. Both Cantor’s and Blanchard’s studies suggest that at least among some pedophiles, biological substrates may play an important role in the origin of these deviant sexual interests.

Although the findings of neurobiological underpinnings of pedophilia appear to be relatively robust, future research should further explore the possible indirect effects of such factors on the early environmental experiences that may contribute to the development of deviant sexual interests in children. It is possible that some of the neurobiological correlates of pedophilia could contribute to considerably stressful developmental years of those who later become pedophiles. Marshall [25, 26] has suggested that such unfortunate experiences may predispose developing youngsters to form age-inappropriate attachments which may, in the long run, contribute to a sexual interest in children. Child molesters certainly

do experience greater emotional congruence with children than do other men and they have corresponding problems in emotional attachments with adults [27]. In any case, the issue concerning what factors play a direct versus an indirect role in the genesis of pedophilia has not yet been fully settled.

Focusing on both treatment effects and long-term outcome (a proxy for prognosis), Marshall [28] reported a study that has nosological relevance for pedophilia. He found that among the 68 treated child molesters, there was no significant difference in long-term outcome between those diagnosed as pedophiles and those who were determined to not meet criteria for this diagnosis; the re-offense rates were 5.9% for treated pedophiles and 7.8% for the non-pedophiles who received treatment. What these data appear to suggest is that the same form of treatment is effective regardless of diagnostic status.

Contrary to the findings with the treated clients, Marshall found a significant difference in outcome among the untreated clients ( $N = 58$ ). Untreated pedophiles reoffended at a higher rate (23.8%) than did the non-pedophilic child molesters (12%), suggesting that the long-term prognosis for untreated pedophiles is rather dismal compared to that with other child molesters. Thus, at least among untreated pedophiles, the diagnosis appears to have prognostic relevance. However, it is important to note that researchers disagree about the meaning of the absence of the behavioral features of pedophilia after treatment.

A 2002 special edition of the *Archives of Sexual Behavior* was devoted to a consideration by numerous authors of several aspects of pedophilia. Some of these authors [29–31] claimed that the disorder is a lifelong persistent sexual orientation akin to homosexuality and that it is, accordingly, unchangeable. Seto [32] has also suggested that pedophilia is best thought of as a sexual orientation. Indeed, DSM-5 declares pedophilia “to be a lifelong condition... (although) the propensity to act out sexually” may change [4 (p699)]. Other authors writing in the special edition of the *Archives* [33, 34] vigorously deny that pedophilia represents an unchangeable sexual orientation.

Seemingly consistent with this latter view are findings reported by Muller et al. [35]. They examined the stability of pedophilic interests over time as assessed by phallometry. Muller et al. found significant reductions in their derived pedophilic index over an extended follow-up period (6 to 259 months). Muller et al. took these changes to mean that pedophilic interests can be reduced and are, thus, not unchangeable. Several authors, however, took issue with this claim. As Bailey [36] pointed out, arousal patterns assessed by phallometry have been shown to include measurement error, and that this, he noted, was evident in Muller et al.’s data. Thus, the apparent changes might simply reflect this error in measurement. Both Cantor [37] and Lalumière [38] pointed to other serious methodological problems with Muller et al.’s paper and Mokros and Habermeyer [39] noted that both

low reliability and regression to the mean could explain Muller et al.’s data.

Finally, it should be noted that at least some sexual offenders and paraphilics as well as non-offending men can control their sexual responses during phallometric testing if they are instructed to do so [40, 41]. Specifically, Quinsey and Carrigan [42] showed that child molesters could fake “normal.” Contrary to these observations, Babchishin et al. [43] reported that after controlling for measurement error, 83% of pedophiles were unable to demonstrate control over their arousal to child stimuli during phallometric testing. However, 15% were able to change their pedophilic profile to a normal pattern of responding.

Thus, it seems that the two competing hypotheses (pedophilia can or cannot be changed) are not open to empirical validation or rejection. Even if it is accepted that apparent reductions in pedophilic responses or behaviors occur as a result of treatment or simply as a function of time, these results could equally be interpreted as indicating that pedophilia has been eliminated or that the pedophiles have simply learned to inhibit the overt manifestation of their still persisting sexual interests.

## Reliability

The vagueness of the criteria noted earlier suggests that it is unlikely that reliability across diagnosticians will meet acceptable standards. However, as Nelson-Gray [44] has pointed out, establishing reliability is a particularly important aspect of the application of any diagnosis. An estimate of reliability in the case of the paraphilias could be established by having several clinicians independently assess the same set of cases with a subsequent analysis of the degree of agreement across these evaluators. Studies of inter-rater agreement are in fact what the authors of DSM have initiated for the majority of disorders since at least DSM-III, although such studies do not appear to have been repeated with the paraphilias. The degree of inter-rater reliability is typically determined by calculating the kappa coefficient which corrects for chance agreement between diagnosticians. In interpreting the meaning of coefficients of reliability in such cases, it should be noted that Cicchetti and Sparrow’s [45] psychometric standards indicate that for very important decisions, a kappa ( $k$ ) of 0.9 is essential. Given the impact diagnoses of pedophilia or sexual sadism have on the offender’s prospect of early release, and the potential impact on the public if a dangerous offender is released on parole, then clearly, these diagnoses meet Cicchetti and Sparrow’s standard of a very important decision. Accordingly, we should expect inter-diagnostician agreement to be at or above  $k = 0.9$ .

Appropriate data on reliability can be derived from examining diagnoses applied during independent assessments conducted by trained evaluators for the courts in hearings of

applications for Sexually Violent Predator status. With regard to these civil commitment laws, the US Supreme Court has stipulated that in order for sex offenders to meet commitment requirements, they must exhibit a mental condition that predisposes them to commit a sexually violent offense [46]. Several paraphilic conditions, notably pedophilia, sexual sadism, and what is now termed “Other specified paraphilic disorder” (previously known as “Paraphilia not otherwise specified”), have served as the requisite mental conditions [47, 48]. Levenson [49] has described her examination of extensive files on these assessments conducted by independent evaluators in the state of Florida. She found unacceptably low inter-assessor reliability for four sets of diagnoses: pedophilia ( $k = 0.65$ ); sexual sadism ( $k = 0.30$ ); exhibitionism ( $k = 0.47$ ); and paraphilia NOS ( $k = 0.36$ ); as well as the presence of any paraphilia ( $k = 0.47$ ). As can be seen, none of these coefficients even approach acceptable levels for an important decision such as indefinite confinement.

In a series of studies, Marshall and his colleagues examined the status of the diagnosis of sexual sadism. They began with a detailed review of the literature [50] which unfortunately revealed that no center adhered to DSM criteria but rather created their own based on their clinical experience. This meant that comparisons across settings were compromised.

Next, Marshall et al. [51] reviewed the clinical application of the diagnosis of sadism completed in a secure psychiatric center located in a maximum security federal prison in Canada. The inmates of this institution were often referred by the National Parole Board to independent psychiatrists for the assessment of the presence of sexual sadism. These independent psychiatrists were all experienced in dealing with dangerous offenders. The psychiatrists’ decisions had very important implications because if they decided that an offender was a sadist, then he was very unlikely to get parole; if he was seen to not meet criteria for the disorder, he had a strong chance of obtaining early release. Errors of diagnosis could, therefore, put the public at risk or unnecessarily extend an offender’s sentence.

Marshall et al. examined 51 of these evaluations and checked the relationship between each decision and an array of information contained in the files upon which the examining psychiatrists based their opinions. Included in this information was past psychiatric evaluations, phallometric test results, various psychological test results, detailed case notes, and the results of their own interviews with the offenders. When Marshall et al. examined the diagnoses made by these evaluators in relation to the same information they used, they were surprised by the results. Those offenders deemed not to be sadists and showed greater arousal to depictions of violent rapes than did those who were said to be sadists. Furthermore, the so-called non-sadists had been more likely than the psychiatrist-diagnosed sadists to have tortured and humiliated their victims.

Given these disappointing results, Marshall et al. [52] extracted the complete set of information to which the psychiatrists in the previous study had access. This set of information on 12 of the offenders (6 who were judged to be sadists and 6 who were not) was sent to 15 internationally acknowledged experts in sexual sadism. The primary question put to these experts simply asked them to indicate whether or not each client met criteria for sexual sadism. The analysis of the resultant data revealed very low inter-diagnostician agreement ( $k = 0.14$ ). However, these experts provided more uniform responses to the secondary question put to them. In this instance, they were asked to identify the criteria they believed most accurately identified sadists and to rate the relative importance of each feature they said was critical. This information later served as the basis for the development of a scale meant to assist clinicians in making a diagnosis of sexual sadism [53].

Overall, the available data do not encourage confidence in the reliability of paraphilic diagnoses, even when they are applied under conditions which we might expect to generate high reliability. There are, however, alternative ways to derive paraphilic diagnoses.

## Alternative Diagnostic Strategies

### Phallometrics

Phallometry involves the measurement of erectile responses to sexually explicit stimuli [54]. Freund [55, 56] suggested that the most accurate way to identify pedophilia is to employ phallometric testing where the relevant stimuli include sexualized depictions of children and adults. Responses to children that were greater than responses to adults would, so Freund declared, serve as diagnostic indicators of pedophilia. Other researchers have attempted to generate relevant phallometric stimuli that might serve as bases to diagnose sadists [57].

The use of phallometry in the assessment of paraphilics has won widespread acceptance as a method to identify deviance in need of treatment. This is not to say that such assessments are free of problems. As noted earlier, Cantor [37], Lalumière [38], and Mokros and Habermeyer [39•] have pointed to several potential measurement and methodological issues with the way in which phallometric assessments are often done. As a consequence, serious errors may result in the interpretation of the findings. It has also been noted [54] that the test-retest reliability of phallometry is unsatisfactory which also suggests a limit to this assessment procedure. Despite these caveats, phallometry remains popular as can be seen in the chapters on the various paraphilias in Laws and O’Donohue’s [5, 6••] two volumes. It is clear, however, that more research needs to be done before we can rely on this

methodology, particularly in its capacity to track changes over time. While, as we will show, there are alternatives, all of them have limitations.

### Indirect Measures

These measures cover an array of approaches [58], all of which attempt to circumvent the offenders' understandable disposition to hide their true sexual interests. These procedures include discrepancies in viewing time between identifying child and adult images [59], choice reaction time measures [60], implicit association tests [61], and the emotional Stroop test [62].

Unfortunately, the empirical support for these various measures has not always clearly discriminated paraphilics from other males [63], although the emotional Stroop test has generated interesting results with sex offenders [64–66]. While these alternative assessment strategies have the advantage of obscuring from the subjects the real intent of the evaluations, there is still considerable work to be done to make most of them acceptable alternatives for diagnostic and treatment evaluation purposes.

### Diagnostic Rating Scales

Seto and Lalumière [67] developed a simple scale meant to be completed based on a review of available file information. They showed that this scale accurately predicted a diagnosis of pedophilia. A critical aspect of this scale concerns the fact that the presence of two or more prior victims is a crucial factor in identifying pedophilia, just as Freund and Watson [68] had earlier shown to be the case. Subsequent validation of this scale applied in a separate setting revealed its robustness in accurately diagnosing pedophilia [69].

While empirically derived rating scales have been developed for assessing sexual sadism [53, 70], Seto's scale remains the only other scale for the paraphilias. Nitschke et al.'s [70] *Severe Sexual Sadism Scale* (SeSaS) has been thoroughly explored for its psychometric properties and been found to meet high standards of scale-ability, reliability, and reproducibility, and to generate high inter-rater agreement [71]. It has subsequently been shown to be applicable to dangerous female sexual offenders [72] and to the assessment of men detained in an American civil commitment center [73]. The SeSaS has the advantage that it can be scored categorically, to produce a diagnosis, or dimensionally, to provide an estimate of the degree of sadistic interest. Thus, it appears to have significant potential and can provide a model for the development of similar scales for the various other paraphilias.

### Summary of Alternative Diagnostic Strategies

It appears that phallometric testing can be a valuable objective tool to assist diagnostic approaches with the various paraphilias. There seems to be some promise for one or more of the alternative indirect methods and particularly for the recent developments of diagnostic scales. Unfortunately, these conclusions seem, at present, to be valid only for pedophilia and sexual sadism.

### A Dimensional Perspective

An important question for the committee formulating the next edition of the DSM is whether or not the paraphilias can best be construed as dichotomous categories (i.e., as indicating the presence or absence of a disorder), or as degrees of disorder located on dimensions anchored at one end by normal functioning and at the other by severe disorder. One problem with categorical diagnoses is that they view people with "disorders" as qualitatively different from the so-called normal people. As some authors have observed, dichotomous diagnoses do not represent natural classes [74, 75] as the majority of the aspects of human behavior range along continua.

The DSM's categorical view of the paraphilias does not appear to match research findings. As has been shown, those child molesters who have two or more victims are the ones most likely to display arousal patterns indicative of pedophilia. What these findings suggest is that child molesters range along a continuum in terms of their numbers of victims which seems to imply that a dimensional approach to diagnosing pedophilia might be more useful.

Matching this view are the results of a study by Barbaree et al. [76]. They compared the phallometric responses of rapists and non-offending males to depictions of normative sex and to depictions of sexual violence. They found that the differences between these two groups were accounted for by an inhibitory response to the elements depicting forcefulness that was most evident among the normal subjects but that varied in degree in the rapists. Lalumière and Quinsey [77] also reported that a number of rapists displayed normal responses at phallometric testing while several non-offenders responded to sexual assault stimuli. As Nitschke and Marshall [78] have suggested, these data imply a continuum of responses to sexual violence with extreme deviants lying at one end of the spectrum while the other end captures men without any apparent sexually abusive tendencies.

Consistent with this idea that paraphilias may be best thought of as lying along a dimension are the observations of several authors working with sexual aggressors. Knight [79] and Knight et al. [80] concluded, after examining their own evidence and the observations of Barbaree et al. [76], that the responses of rapists can be ordered along a continuum

ranging from no arousal to sexual violence to very strong arousal to such cues. Similarly, Mokros et al. [81•] and O'Meara et al. [82] consider the evidence to indicate that sexual sadism is best viewed as dimensional rather than simply categorical.

Research aimed at identifying features indicative of future risk among sex offenders [83, 84] also suggests a dimensional quality to these risk factors and so could be seen as providing the basis for the development of dimensional scales. As we have seen, deviance revealed by phallometric assessments has a natural dimensional quality, as do other risk factors such as deficits in relationship skills, sexual preoccupation, features of dysregulated behavior, and emotional congruence with children. Features associated with other paraphilias could also be rendered into dimensional scales as has been suggested for voyeurism by Mann et al. [85].

Hopefully, the development of appropriate scales will soon be realized so that all paraphilias can be identified as continua. If so, this will then provide the bases for determining the degree of deviant sexual interests while at the same time retaining the possibility of discerning a categorical diagnosis. Such scales would also eliminate the need for diagnosticians to interpret the meaning of the vaguely stated DSM criteria.

## Conclusions

This paper has considered various facets of the DSM's attempt to provide categorical diagnoses for the paraphilias. These considerations were limited to the diagnoses of pedophilia and sexual sadism since they are the only diagnoses for which reasonable evidence is available.

The criteria for the paraphilias, as specified in the various editions of the DSM, were found to be vague and, as a consequence, unreliable. This lack of reliability was particularly evident when examined under conditions where the highest reliability might be expected. In evaluating diagnostician agreement between independent assessors preparing reports for Sexually Violent Predator hearings, the resultant levels of reliability were unacceptably low. Similar disappointing results were found with diagnoses of sexual sadism made by internationally renowned experts. The capacity of the DSM to meet the standards of a nosological system (i.e., the identification of etiology, treatment, and prognoses) was also examined and found wanting.

As a result of these various observations, alternative methods of diagnosing the paraphilias were considered. The use of phallometric testing and the deployment of various indirect measures of sexual interests were considered and seemed to offer promise. The recent emergence of rating scales was also appraised and they also seem worth pursuing. The advantage of these scales is that they can be scored both categorically, and thereby produce a diagnosis, and

dimensionally resulting in an estimate of the degree of deviance. Whatever course future developments take, at the very least, the next DSM committee must attempt to develop more precise criteria that can be interpreted with less ambiguity.

Hopefully, the comments and observations expressed in this paper will encourage future researchers, whether or not they support the views expressed here, to examine the possibilities and problems detailed in this paper. The primary value of a review, such as that presented here, is to encourage empirically based challenges.

## Compliance with Ethical Standards

**Conflict of Interest** William L. Marshall is a section editor for *Current Psychiatry Reports*.

Drew A. Kingston declares no potential conflicts of interest.

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

## References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
  - Of major Importance
1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Washington, DC: Author; 1952.
  2. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 2nd ed. Washington, DC: Author; 1968.
  3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3rd ed. Washington, DC: Author; 1980.
  4. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Washington, DC: Author; 2013.
  5. Laws DR, O'Donohue WT, editors. Sexual deviance: theory, assessment, and treatment. New York: Guilford Press; 1997.
  6. Laws DR, O'Donohue WT. Sexual deviance: theory, assessment, and treatment. 2nd ed. New York: Guilford Press; 2008. **Provides a comprehensive survey of the assessment and treatment of most of the paraphilias.**
  7. Fazekas D. Pedophilia: more than a moral dilemma. *Arch Sex Behav (Special Edition)*. 2002;31:483–4.
  8. Marshall WL. Pedophilia: psychopathology and theory. In: Laws DR, O'Donohue WT, editors. Sexual deviance: Theory and application. New York: Guilford Press; 1997. p. 152–74.
  9. Marshall WL. Diagnostic problems with sexual offenders. In: Marshall WL, Fernandez YM, Marshall LE, Serran GA, editors. Sexual offender treatment: Controversial issues. Chichester, UK: John Wiley & Sons; 2006. p. 33–43.
  10. Marshall WL. Diagnostic issues, multiple paraphilias, and comorbid disorders in sexual offenders: their incidence and treatment. *Aggress Violent Behav*. 2007;12:16–36.
  11. Kingston DA, Yates PM. Sexual sadism: assessment and treatment. In: Laws DR, O'Donohue WT, editors. Sexual deviance: theory, assessment, and treatment. 2nd ed. New York: Guilford Press; 2008. p. 231–49.

12. O'Donohue WT, Regev LG, Hagstrom A. Problems with the DSM diagnosis of pedophilia. *Sex Abuse*. 2000;12:95–105.
13. Kingston DA, Firestone P, Moulden HA, Bradford JMW. The utility of the diagnosis of pedophilia: a comparison of various classification procedures. *Arch Sex Behav*. 2007;36:423–36.
14. Moulden HM, Firestone P, Kingston DA, Bradford JMW. Recidivism in pedophiles: an investigation using different methods of defining pedophilia. *Foren Psychiat & Psychol*. 2009;20:680–701.
15. Wilson RJ, Abracen J, Looman J, Pichea JE, Ferguson M. Pedophilia: an evaluation of diagnostic and risk prediction methods. *Sex Abuse*. 2011;23:260–74.
16. Kendell RE. *The role of diagnosis in psychiatry*. Oxford, UK: Blackwell; 1975.
17. Klerman GL. The evolution of the scientific nosology. In: Shershow JD, editor. *Schizophrenia: science and practice*. Cambridge, MA: Harvard University Press; 1978. p. 99–121.
18. Mezzich JE. Comprehensive diagnosis: a conceptual basis for future diagnostic systems. *Psychopathology*. 2002;35:162–5.
19. Beech AR, Ward T, editors. *The Wiley handbook on the theories, assessment, and treatment of sexual offending: Vol 1: Theories*. Chichester, UK: John Wiley & Sons; 2017. **This paper book provides a description of a variety of current theories of the etiology of sexual offending.**
20. Cantor JM, Blanchard R, Christensen BK, Dickey R, Klassen PE, Beckstead AL, et al. Intelligence, memory, and handedness in pedophilia. *Neuropsychology*. 2004;18:3–14.
21. Cantor JM, Blanchard R, Robichaud LK, Christensen BK. Quantitative reanalysis of aggregate data on IQ in sexual offenders. *Psychol Bull*. 2005;131:555–68.
22. Blanchard R, Christensen BK, Strong SM, Cantor JM, Kuban ME, Klassen P, et al. Retrospective self-reports of childhood accidents causing unconsciousness in phallometrically diagnosed pedophiles. *Arch Sex Behav*. 2002;31:511–26.
23. Blanchard R, Kuban ME, Klassen P, Dickey R, Christensen BK, Cantor JM, et al. Self-reported head injuries before and after age 13 in pedophilic and nonpedophilic men referred for clinical assessment. *Arch Sex Behav*. 2003;32:573–81.
24. Cantor JM, Kabani N, Christensen BK, Zipursky RB, Barbaree HE, Dickey R, et al. Cerebral white matter deficiencies in pedophilic men. *J Psychiatr Res*. 2008;42:167–83.
25. Marshall WL, Barbaree HE. An integrated theory of sexual offending. 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. 257–75.
26. Marshall WL, Marshall LE. An attachment-based theory of the aetiology of affiliative child molestation: resilience/vulnerability factors across the life-span. In: Ward T, Beech A, editors. *The Wiley handbook on the assessment, treatment and theories of sexual offending, vol. Vol.1: Theories*. Chichester, UK: John Wiley & Sons; 2017. p. 3–24.
27. Mann RE, Hanson RK, Thornton D. Assessing risk for sexual recidivism: some proposals on the nature of psychologically meaningful risk factors. *Sex Abuse*. 2010;22:191–217.
28. Marshall WL. Are pedophiles treatable? Evidence from North American studies. *Seksulogija Polska*. 2008;6:1–6.
29. Berlin FS. Pedophilia: when is a difference a disorder? *Arch Sex Behav (Special Edition)*. 2002;31:479–80.
30. Gaither GA. Pedophilia as a sexual orientation? *Arch Sex Behav (Special Ed)*. 2002;31:485–6.
31. Langevin R. Yes, Virginia, there are real pedophiles: a need to revise and supervise, not eliminate, DSM. *Arch Sex Behav (Special Ed)*. 2002;31:488–9.
32. Seto MC. Is pedophilia a sexual orientation? *Arch Sex Behav*. 2012;41:231–6.
33. Berner W. Pedophilic sexual orientation: a fuzzy expression. *Arch Sex Behav (Special Edition)*. 2002;31:480–1.
34. Dixon AF. Abnormal erotosexual preferences in human beings: the nature of pedophilia. *Arch Sex Behav (Special Ed)*. 2002;31:482.
35. Muller K, Curry S, Ranger R, Briken P, Bradford JMW, Fedoroff JP. Changes in sexual arousal as measured by penile plethysmography in men with pedophilic sexual interests. *Int Soc Sex Med*. 2014;11:1221–9.
36. Bailey JM. A failure to demonstrate changes in sexual interest in pedophilic men: comment on Muller et al.. (2014). *Arch Sex Behav*. 2015;44:249–52.
37. Cantor JM. Purported changes in pedophilia as statistical artefacts: comment on Muller et al. (2014). *Arch Sex Behav*. 2015;44:253–4.
38. Lalumière ML. The lability of pedophilic interests as measured by phallometry. *Arch Sex Behav*. 2015;44:255–8.
39. Mokros A, Habermeyer E. Regression to the mean mimicking changes in sexual arousal to child stimuli in pedophiles. *Arch Sex Behav*. 2016;45:1863–7. **Outlines the various influences that can distort assessments of sexual deviance.**
40. Abel GG, Blanchard EB, Barlow DH. The effects of stimulus modality, instructional set and stimulus content on the objective measurement of sexual arousal in several paraphilias. *Behav Res Ther*. 1981;19:25–33.
41. 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.
42. Quinsey VL, Carrigan WF. Penile responses to visual stimuli. *Crim Justice Behav*. 1978;5:333–42.
43. Babchisin KM, Curry SD, Fedoroff JP, Bradford J, Seto MC. Inhibiting sexual arousal to children: correlates and its influence on the validity of penile plethysmography. *Arch Sex Behav*. in press; <https://doi.org/10.1007/s10508-017-0925-2>.
44. Nelson-Gray RD. Empirical guidelines from psychometrics. *J Abnorm Psychol*. 1991;100:308–15.
45. Cicchetti DV, Sparrow SS. Developing criteria for establishing interrater reliability of specific items: applications of assessment of adaptive behavior. *Am J Ment Defic*. 1981;86:127–37.
46. Cohen F. Introduction to legal issues: how the legal framework developed. In: Schwartz BK, Cellini HR, editors. *The sex offender: corrections, treatment and legal practice*. Kingston, NJ: Civic Research Institute; 1995. p. 22.1–22.12.
47. Doren DM. *Evaluating sex offenders: a manual for civil commitments and beyond*. Thousand oaks, CA: Sage publications; 2002.
48. Janus ES. Sexual predator commitment laws: lessons for law and the behavioral sciences. *Behav Sci Law*. 2000;18:5–21.
49. Levenson JS. Reliability of sexually violent predator civil commitment criteria in Florida. *Law Hum Behav*. 2004;28:357–68.
50. Marshall WL, Kennedy P. Sexual sadism in sex offenders: an elusive diagnosis. *Aggress Violent Behav*. 2003;8:1–22.
51. Marshall WL, Kennedy P, Yates PM. Issues concerning the reliability and validity of the diagnosis of sexual sadism applied in prison settings. *Sex Abuse*. 2002;14:301–11.
52. Marshall WL, Kennedy P, Yates PM, Serran GA. Diagnosing sexual sadism in sexual offenders: reliability across diagnosticians. *Int J Offender Ther Comp Criminol*. 2002;46:668–77.
53. Marshall WL, Hucker SJ. Issues in the diagnosis of sexual sadism. *Sex Off Treat*. 2006;1:1–5.
54. Marshall WL, Fernandez YM. *Phallometric testing with sexual offenders: limits to its value*. Brandon, VT: Safer Society Press; 2003.
55. Freund K. Erotic preference in paedophilia. *Behav Res Ther*. 1987;25:339–48.
56. Freund K, Blanchard R. Phallometric diagnosis of pedophilia. *J Consult Clin Psychol*. 1989;57:1–6.
57. Proulx J, Blais E, Beaugard E. Sadistic sexual aggressors. In: Marshall WL, Fernandez YM, Marshall LE, Serran GA, editors.

- Sexual offender treatment: controversial issues. Chichester, UK: John Wiley & Sons; 2006. p. 61–77.
58. Thornton D, Laws DR, editors. Cognitive approaches to the assessment of sexual interest in sexual offenders. Chichester, UK: John Wiley & Sons; 2009.
  59. Abel GG, Jordan A, Hand C, Holland L, Phipps A. Classification models of child molesters utilizing the Abel Assessment for Sexual Interest. *Child Abuse Negl.* 2001;25:703–18.
  60. Gress CLZ, Laws DR. Cognitive modelling of sexual arousal and interest: choice reaction time measures. In: Thornton D, Laws DR, editors. Cognitive approaches to the assessment of sexual interest in sexual offenders. Chichester UK: John Wiley & Sons; 2009. p. 85–99.
  61. Babchishin KM, Nunes K, Kessous N. A multimodal examination of sexual interest in children: a comparison of sex offenders and non-sex offenders. *Sex Abuse.* 2014;26:343–74.
  62. Smith P. Assessing sexual interest with the emotional Stroop test. In: Thornton D, Laws DR, editors. Cognitive approaches to the assessment of sexual interest in sexual offenders. Chichester UK: John Wiley & Sons; 2009. p. 159–76.
  63. Bartels RM, Gray NS, Snowden RJ. Indirect measures of deviant sexual interest. In: Craig LA, Rettenberger M, editors. *The Wiley handbook on the theories, assessment, and treatment of sexual offending*, vol. Vol II Assessment. Chichester UK: John Wiley & Sons; 2017. p. 965–93.
  64. Price SA, Beech AR, Mitchell IJ, Humphrey GW. Measuring deviant sexual interests in adolescents using the emotional Stroop task. *Sex Abuse.* 2014;26:450–71.
  65. Price SA, Hanson RK. A modified Stroop task with sexual offenders: a replication study. *J Sex Aggress.* 2007;13:203–16.
  66. Smith P, Waterman M. Processing bias for sexual material: the emotional Stroop and sexual offenders. *Sex Abuse.* 2004;16:167–71.
  67. Seto MC, Lalumière ML. A brief screening scale to identify pedophilic interests among child molesters. *Sex Abuse.* 2001;13:15–25.
  68. Freund K, Watson R. Assessment of the sensitivity and specificity of a phallometric test: an update of “Phallometric diagnosis of pedophilia”. *Psychol Assess.* 1991;3:254–60.
  69. Seto MC, Harris GT, Rice ME, Barbaree HE. The screening scale for pedophilic interests predicts recidivism among adult sex offenders with child victims. *Arch Sex Behav.* 2004;33:455–66.
  70. Nitschke J, Osterheider M, Mokros A. A cumulative scale of severe sexual sadism. *Sexual Abuse: A Journal of Research and Treatment.* 2009;21:262–78.
  71. Mokros A, Osterheider M, Schilling F, Eher R, Nitschke J. *The Severe Sexual Sadism Scale: cross validation and scale properties.* *Psychol Assess.* 2012;24:764–9.
  72. Pflugradt DM, Bradley PA. Evaluating female sex offenders using the Cumulative Scale of Severe Sexual Sadism, Poster presented at the 30<sup>th</sup> Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers. Toronto: Canada; 2011.
  73. Wilson, R.J., Pake, D.R., & Duffee, S. (2011) DSM-5 pedohebephilia, PCD, and sadism diagnoses: reliability in Florida. Paper presented at the 30<sup>th</sup> Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Toronto, Canada.
  74. Kendell RE, Jablensky A. Distinguishing between the validity and utility of psychiatric diagnoses. *Am J Psychiatr.* 2003;160:4–12.
  75. Williamson P, Allman J. *The human illnesses.* Oxford: Oxford University Press; 2011.
  76. Barbaree HE, Marshall WL, Lanthier R. Deviant sexual arousal in rapists. *Behav Res Ther.* 1979;17:215–22.
  77. Lalumière ML, Quinsey VL. The discriminability of rapists from non-sex offenders using phallometric measures: a meta-analysis. *Crim Justice Behav.* 1994;21:150–74.
  78. Nitschke J, Marshall WL. An evaluation of assessments of sexual sadism. In: Proulx J, et al., editors. *International handbook of sexual homicide.* London: Routledge; in press.
  79. Knight RA. Issues in sexual sadism. Berlin: Paper presented at the 12th Conference of the International Association for the Treatment of Sexual Offenders; 2012.
  80. Knight RA, Sims-Knight JE, Guay JP. Is a separate diagnostic category defensible for paraphilic coercion? *J Crim Just.* 2013;41:90–9.
  81. Mokros A, Schilling F, Weiss K, Nitschke J, Eher R. Sadism in sexual offenders: evidence for dimensionality. *Psychological Assessment.* 2014;26:138–47. **Offers a dimensional alternative to current categorical diagnostic approach.**
  82. O’Meara A, Davies J, Hammond S. The psychometric properties and utility of the *Short Sadistic Impulse Scale* (SSIS). *Psychol Assess.* 2011;23:523–31.
  83. Hanson RK, Bussière MT. Predicting relapse: a meta-analysis of sexual offender recidivism studies. *J Consult Clin Psychol.* 1998;66:123–42.
  84. Hanson RK, Morton-Bourgon KE. The characteristics of persistent sexual offenders: a meta-analysis of recidivism studies. *J Consult Clin Psychol.* 2005;73:1154–63.
  85. Mann RE, Ainsworth F, Al-Attar Z, Davies M. Voyeurism: assessment and treatment. In: Laws DR, O’Donohue WT, editors. *Sexual deviance: theory, assessment, and treatment.* 2nd ed. New York: Guilford Press; 2008. p. 320–35.