2008

Professional Perspectives on Sexual Sadism

Julia E McLawsen, *University of Nebraska-Lincoln*
Rebecca L Jackson
Steven D Vannoy, *University of Washington, Seattle*
Gregg J Gagliardi
Mario J Scalora, *University of Nebraska-Lincoln*

Sexual Abuse: A Journal of Research and Treatment
Volume XX Number X
Month XXXX xx-xx

10.1177/1079063208320250
http://sajrt.sagepub.com
hosted at
http://online.sagepub.com

Professional Perspectives on Sexual Sadism

Julia E. McLawsen
University of Nebraska-Lincoln
Rebecca L. Jackson
Pacific Graduate School of Psychology, Palo Alto, California, and Special Commitment Center, Steilacoom, Washington
Steven D. Vannoy
University of Washington, Seattle
Gregg J. Gagliardi
Western State Hospital, Steilacoom, Washington
Mario J. Scalora
University of Nebraska-Lincoln

Significant controversy surrounds the diagnosis of sexual sadism. Research suggests that many characteristics attributed to sexual sadists fail to differentiate sexual offenders with and without this diagnosis. Furthermore, when there are differences between sadists and nonsadists, "sadistic" features are frequently associated with nonsadists. Finally, diagnosticians appear to use idiosyncratic methods to diagnose sexual sadism. These findings raise concerns about the reliability and validity of a diagnosis of sexual sadism, particularly with respect to how professionals conceptualize this diagnosis. This study examines how professionals understand the relative importance of behaviors associated with sadistic versus nonsadistic sexual offending. Professionals rated behaviors according to their "essentialness" for this diagnosis. Results show that professionals rated behaviors associated with three out of four conceptualizations of sexual sadism as significantly more essential to making a diagnosis of sexual sadism, compared to behaviors associated with nonsadistic sexual offending. Results suggest that professionals reliably discriminate between sadistic and nonsadistic offense behaviors.

Keywords: *sexual sadism; sexual offenders; behavioral indicators; diagnostic validity; mental health professionals*

Significant controversy surrounds the diagnosis of sexual sadism. The most recent edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–IV–TR*; text revision, American Psychiatric Association, 2000) specifies diagnostic criteria that are vague (e.g., "suffering") and difficult to systematically measure, such as "suffering (including humiliation) of the victim is sexually exciting to the person" (p. 574). Attempts to differentiate offenders with and without diagnoses of sexual sadism have yielded mixed results.The overwhelming majority of sexual

sadism research indicates that a diagnosis of sexual sadism fails to distinguish groups of sexual offenders. One study found that sexually sadistic[1] murderers expressed higher mean phallometric responses to the Humiliation Rape Index when compared to nonsadistic sexual murderers, although there were no between-group response differences to the Nonsexual Physical Violence Index (Proulx, Blais, & Beauregard, 2003, as cited in Proulx, Blais, & Beauregard, 2006). Other researchers, however, have criticized the diagnostic procedures that Proulx and colleagues used to define their study's sexually sadistic sample (Marshall & Hucker, 2006b), calling into question the generalizability of their study's results.

Marshall and Kennedy (2003) note the difficulty of operationalizing the *DSM–IV–TR* diagnostic criteria for sexual sadism, attributing the "elusive nature" (p. 15) of this diagnosis to diagnosticians' need to infer that sexual motivation drives the offender's infliction of suffering. In addition, Marshall, Kennedy, and Yates (2002) found that features described by the sexual sadism literature as characteristic of sexually sadistic offenses failed to differentiate sex offenders with and without diagnoses of sexual sadism. The authors observed that when there were group differences between the sadists and nonsadists, results associated the "sadistic" features more frequently with nonsadists. Other research has demonstrated that diagnosticians use idiosyncratic methods to diagnose sexual sadism (Levenson, 2004; Marshall, 2006; Marshall & Hucker, 2006b; Marshall, Kennedy, Yates, & Serran, 2002). To be useful, a classification system such as *DSM* diagnoses must use categories that can be reliably applied (Nelson-Gray, 1991). With regard to sexual sadism, however, Marshall, Kennedy, and colleagues' findings indicate a diagnosis that neither communicates consistent information nor points to effective mental health interventions. If these findings are accurate, then they imply that sexual sadism fails to fulfill the two goals of psychiatric diagnosis.

The term "diagnosis" originates from the Greek for through/between (*dia)* and knowledge (*gnosis*). Diagnosis refers to a process through which one person discerns specific knowledge about another person. First, a diagnosis represents a standardized label that communicates information about the patient (*DSM–IV-TR*). However, if a diagnosis of sexual sadism communicates information that depends on the idiosyncratic methods that the diagnostician used in rendering the diagnosis, it fails to represent a standardized label. Second, diagnosis represents "the first step in a comprehensive evaluation" (*DSM–IV-TR*, p. xxxiv) that guides mental health professionals to select specific interventions that are likely to ameliorate the symptoms warranting the diagnosis. If a diagnosis of sexual sadism lacks a standardized meaning, research studies that focus on this diagnosis cannot be compared to one another. Hence, researchers and clinicians may be prevented from accumulating a body of knowledge that could inform the development of mental health interventions and management strategies. Although evidence casts doubt on whether this diagnosis communicates meaningful information that points to effective mental health interventions, a diagnosis of sexual sadism can greatly influence decisions based on an

offender's perceived recidivism risk (Marshall & Hucker, 2006a, b; Marshall, Kennedy, et al., 2002; Marshall, Kennedy, Yates, et al., 2002). Sexually violent predator (SVP) civil commitment judgments represent the most notable of these risk-based decisions.

The Importance of a Diagnosis of Sexual Sadism

In a recent study of factors predicting the civil commitment of SVPs, Levenson and Morin (2006) observed that evaluators recommended civil commitment for 9 out of 10 individuals diagnosed with sexual sadism. Based on this finding, Levenson and Morin hypothesized that evaluators strongly rely on the presence of a diagnosis of sexual sadism when formulating SVP civil commitment recommendations. This hypothesis seems sensible considering the common elements of SVP civil commitment laws.

Although SVP laws differ from state to state, they share a similar set of criteria (Jackson, Rogers, & Shuman, 2004; Janus, 2000). First, the individual must have at least one conviction for a sexually violent offense (as defined by state penal code).[2] Second, the individual must have a mental abnormality or personality disorder. Third, it must be "more likely than not" that the individual will commit another sexual offense, if released to a community setting. Finally, the mental abnormality must increase the likelihood that the individual will commit another sexual offense. Although a diagnosis of sexual sadism does not automatically fulfill the legal elements for SVP commitment, the behaviors that form the basis for this diagnosis are often relevant to the elements. For example, a sexually violent act results in a criminal conviction (element 1). That same act may form the basis for a diagnosis of sexual sadism (element 2). And because past behavior strongly predicts future behavior (Andrews & Bonta, 2003; Thorndyke, 1911), the diagnosis can also suggest an increased likelihood that a given individual will continue to engage in sexually violent acts (elements 3 and 4).

A diagnosis of sexual sadism bears serious implications for decisions related to an offender's confinement and release, which makes it important to ensure that professionals understand and apply this diagnosis appropriately. High levels of agreement (e.g., a *kappa* coefficient of 0.9) should be the norm for data used to inform such important decisions (American Educational Research Foundation, 1999; McDermott, 1988; Murphy & Davidshofer, 1998). However, research indicates that diagnosticians diagnose sexual sadism with poor interrater reliability (Marshall & Hucker, 2006b; Marshall & Kennedy, 2003), with *kappas* ranging from 0.14 (Marshall, Kennedy, Yates, et al., 2002) to 0.3 (Levenson, 2004). The following section describes research that has raised questions about the reliability of a sexual sadism diagnosis, and provides a chronological review of philosophical and psychological conceptualizations of sexual sadism.

Past and Present Conceptualizations of Sexual Sadism

Recent studies investigating the discriminant validity of sexual sadism are equivocal. A fundamental requirement for diagnostic validity is that individuals with different diagnoses differ from one another in discernable and clinically meaningfully ways. However, comparisons of individuals with and without diagnoses of sexual sadism do not reliably differentiate the two groups on offense characteristics, psychological data, and phallometric responses (Marshall & Kennedy, 2003; Marshall, Kennedy, et al., 2002). Despite serious concerns of reliability and validity, sexual sadism has a rich and colorful history that still shapes contemporary understandings of the diagnosis. This section begins by reviewing historical literature relevant to sexual sadism to illustrate the construct that contemporary diagnostic systems sought to operationalize. Next, this section examines contemporary research studies that have raised questions about the meaningfulness of a diagnosis of sexual sadism. Finally, the article describes the present study that sought to determine how professionals who regularly assign and/or make decisions based on a diagnosis of sexual sadism understand behavioral features associated with this diagnosis.

The term *sadism* originates from the writings of Donatien Alphonse Francois de Sade, more commonly known as the Marquis de Sade (Lever, 1993). Sade authored most of his sexually explicit tales during the 27 years he spent in prison—a consequence of his scandalous lifestyle. Although Sade did not *create* sadism, his life and works provide the first significant[3] illustration of a phenomenon that still inspires inquiry. Consideration of Sade's work has been integral to the production of an ensuing body of literature that continues to parse out the meanings of sadistic behavior and sadism itself. During the last two centuries, inquiry into sadism has bifurcated into artificially distinct domains of philosophical reflection and psychiatric categorization.

Richard von Krafft-Ebing, a German psychiatrist, introduced the term "sadism" to the medical sphere. His magnum opus *Psychopathia Sexualis* [Psychopathy of Sex] (1886/1965, p. 109) contains numerous case studies that ground his formulation of rudimentary behavioral standards for sadism:

> The experience of sexual, pleasurable sensations (including orgasm) produced by acts of cruelty, bodily punishment afflicted on one's person or when witnessed in others, be they animals or human beings. It may also consist of an innate desire to humiliate, hurt, wound, or even destroy others in order, thereby, to create sexual pleasure in one's self.

In addition to describing its basic manifestation, Krafft-Ebing (1886/1965) organizes sadism into a number of subclassifications (see also Hucker, 1997, p. 196):

> Acts that involve a connection between sexual arousal and killing (this category is also referred to as "lust-murder"). "Jack the Ripper" is a particularly notorious—and exemplary—case study in this section.

Mutilation of corpses and/or necrophilia.
Injury inflicted on women.
Defilement of women.
Symbolic assaults on women (i.e., the perpetrator cuts his[4] victim's hair rather than inflicting overt physical pain).
Sadistic fantasies that are not acted out.
Sadistic acts with objects other than women (i.e., men or boys).
Sadistic acts with animals.

Krafft-Ebing (1886/1965) divides these eight subtypes into two groups. The first group pertains to mild sadism in a consensual sexual relationship (e.g., the sadistic component of contemporary "S&M" subcultural activities). The second group involves injury or death and usually refers to actions that occur within a nonconsensual relationship. In both, the element of inflicting pain on the victim is the sexual stimulus for the sadistic actor. Krafft-Ebing's second category of sadism endures in the world of psychiatry, and is now reflected in the contemporary diagnostic criteria for sexual sadism (*DSM–IV–TR*). In the century or so between Krafft-Ebing and the *DSM,* several other theoreticians have contributed to the present understanding of sadism.

Albert von Schrenck-Notzing (1895), a German physician, coined the term *algolagnia,* "pain craving." Within this category, he distinguishes active and passive manifestations of this pain craving as two opposing poles (sadism and masochism) of the same disorder. Albert von Eulenberg (1911), a German sexologist, suggests that the term sadism should encompass psychological, as well as physical, pain. Wilhelm Stekel (1929), a colleague of Krafft-Ebing, emphasized the difference between masochism and sadism, and helped introduce both constructs to clinical work. Benjamin Karpman (1954), a psychoanalyst who worked with sexual psychopaths at St. Elizabeth's Hospital, argues that pain (physical and/or emotional) in and of itself is not of great importance to the construct of sadism. Rather, pain becomes significant insofar as it represents the sadist's power and control over his victim. Likewise, Erich Fromm (1977) writes that the "core of sadism . . . is the passion to have absolute and unrestricted control over living beings. . . . The person who has complete control over another living being makes this being into his thing, his property, while he becomes the other being's god" (pp. 383-384). In an essay titled "Must We Burn Sade?" (1953), Simone de Beauvoir observes that Sade's portrayal of sadism charges sex with a significance that transcends its typical significance. An article by MacCulloch, Snowden, Wood, & Mills (1983) echoes Beauvoir's observation, commenting that the sadist sexualizes his experience of power and control. These formulations of sadism have been interwoven to create the two conventional descriptions of sadism, found in the *DSM–IV–TR* and the *International Classification of Diseases* (ICD; World Health Organization [WHO], 1992).

The *DSM–IV–TR* codifies sexual sadism as a disorder in which the following two criteria are met: (1) over a period lasting at least 6 months the person has recurrent

sexual fantasies in which the physical or psychological suffering of a victim produces sexual excitement; and (2) the person has acted on these desires with a nonconsenting person, or the sexual urges or fantasies caused marked distress or interpersonal difficulty. The ICD description of sadomasochism refers to a preference for sexual activities that impose pain, humiliation, or bondage, without exhibiting anger or cruelty in a sexual context. The inflicting role is referred to as sadism, and the receiving role is referred to as masochism. A diagnosis of sadomasochism is made only when such activities are the main or exclusive source of a person's sexual gratification (WHO, 1992). There have been two contemporary descriptive studies of behavioral features associated with sexual sadism. Dietz, Hazelwood, and Warren (1990) conducted a study to illustrate behavioral features associated with a diagnosis of sexual sadism. Using a sample of individuals identified as "possible sexually sadistic criminals" (p. 166) drawn from the National Center for the Analysis of Violent Crime (NCAVC) database, they confirmed diagnoses of sexual sadism in 30 of these individuals by finding that a subject "had been sexually aroused in response to images of suffering or humiliation on two or more occasions spanning an interval of at least six months" (p. 166). Dietz and colleagues rejected cases in which there was no clear evidence of sexual sadism even though "others had believed [the rejected cases] to be possible sexual sadists" (p. 166), based on the presence of torture, mutilation, and murder. A descriptive analysis of personal and offense characteristics revealed that all subjects were white males who intentionally tortured their victims to arouse themselves. Other common offense characteristics included careful planning of their offenses (93.3%), detached affect during the offense (86.6%), approaching victims under pretext (90%), participation of coperpetrators (36.7%), physical abuse (60%), leading victims to a preselected location (76.7%), restraint of victims (e.g., bound, blind-folded, or gagged victim; 86.7%), sexual bondage (77%), anal (73.3%) and vaginal rape (56.7%), forced fellatio (70%), forcing victims to say degrading and/or humiliating phrases (23.3%), recording offenses (53.3%), and keeping victims' personal items (40%). This study did not use a comparison group, resulting in the major limitation that all personal and offense characteristics showed, to varying degrees, an association to sexual sadism.

Gratzer and Bradford (1995) used Dietz and colleagues' (1990) NCAVC sample as a comparison group for 29 sexual sadists and 28 nonsadistic sexual offenders incarcerated at the Royal Ottawa Hospital (ROH). They noted several characteristics that occurred with relatively equal frequencies between the two groups of sexual sadists, including careful planning of the offense, detached affect during the offense, physical abuse, and leading victims to a preselected location. Gratzer and Bradford also noted a number of differences between the NCAVC and ROH sadists' offense characteristics. Compared to the ROH sadists' offense characteristics, the NCAVC sadists' offenses were more likely to have included the participation of coperpetrators, restraint of victims, sexual bondage, anal and vaginal rape, forced fellatio, telling the victim what to say, recording the offense, and keeping the victim's personal items.

Gratzer and Bradford attributed the between-group differences to the fact that the NCAVC group was part of a national database designed to profile some of the most heinous sexual offenses in United States history, whereas the ROH group consisted of sexual offenders who happened to reside at a Canadian forensic hospital at the time of this study. Finally, Gratzer and Bradford reported several differences between the ROH sadists' offenses and ROH nonsadists' offenses. Compared to the nonsadists' offense characteristics, the sadists' offenses were more likely to have involved careful planning, taking the victim to a preselected location, detached affect, intentional torture, beating the victim, and sexual dysfunction. These characteristics relate to premeditation, as well as controlling and degrading the victim. Gratzer and Bradford hypothesized that sexual dysfunction occurred more frequently among sadists because, rather than becoming aroused by sexual acts in and of themselves, sadists use sex as a "[vehicle] for the degradation and control of the victims" (p. 452). However, it is uncertain how sex, if viewed as a means of degrading and controlling a victim, would fail to arouse a sexually sadistic offender. Although based on small sample sizes, these results illustrate that differences in offense characteristics can exist not only between sadistic and nonsadistic offenders, but also between different groups of sadistic offenders.

The earliest interdiagnostician reliability studies for psychosexual diagnoses occurred during *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-III;* third edition; American Psychiatric Association, 1980) field trials. Unfortunately, these studies neither looked at sexual sadism in isolation from other sexual dysfunctions and paraphilias, nor recorded the individual diagnoses (O'Donohue, Regev, & Hagstrom, 2000). Thus, the resultant *kappa* of 0.92, based on seven cases (O'Donohue & Geer, 1993), does not reflect interdiagnostician agreement specific to sexual sadism. Development of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IV;* fourth edition; American Psychiatric Association, 1994) did not inspire further field trials, even though diagnostic criteria for most of the paraphilias changed substantially between each subsequent version of the *DSM* (Marshall, 2006).

Seven years after the Gratzer and Bradford (1995) study, Marshall, Kennedy, and colleagues (2002) investigated the reliability and validity of sexual sadism diagnoses in Canadian prison settings. They collected archival data from 59 sexual offenders, among whom 41 had received a diagnosis of sexual sadism and 18 had received other diagnoses, based on the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-III-R;* third edition, revised; American Psychiatric Association, 1987) or *DSM-IV* criteria. Marshall, Kennedy, and colleagues concluded that the sadists did not differ from the nonsadists, with a few noteworthy exceptions. First, there were no significant differences in the rates of substance abuse disorders, paraphilias, or antisocial personality disorder between the two groups. There was, however, a significantly higher rate of other personality disorders among the nonsadists. Second, contrary to what theory would predict, nonsadists beat and tortured their victims with greater frequency than did sadists. For example, 61.6% of nonsadists, but only

24.4% of sadists, violently beat their victims. Likewise, 38.9% of nonsadists, but only 9.8% of sadists, tortured their victims. Third, there were no between-group differences regarding self-reported fantasies or acts. Surprisingly, within the subcategories of sexually violent fantasies and prior rapes, "the direction of the [nonsignificant] difference suggested that the nonsadists were more deviant than the sadists" (p. 309), and only 2 out of the 41 sadists "reported fantasies of sexually controlling another person" (p. 309). Finally, nonsadists showed greater phallometric arousal to nonsexual violent stimuli, whereas sadists showed greater phallometric arousal to consenting adult stimuli.[5]

Based on their observations of sadistic offense behavior and sadistic phallometric responses, Marshall, Kennedy, and colleagues (2002) concluded that the nonsadists appeared more deviant than the sadists. This study points to the unreliability of a diagnosis of sexual sadism—raising questions about the methods diagnosticians use to arrive at this diagnosis—and suggests that behavioral features thought to be indicative of sexual sadism (Dietz et al., 1990; Gratzer & Bradford, 1995) may not actually distinguish between sadists and nonsadists.

Marshall, Kennedy, Yates, and colleagues (2002) conducted a follow-up to the previous study (Marshall, Kennedy, et al., 2002) by examining the more specific question of how reliably diagnosticians diagnosed sexual sadism. They invited 24 psychiatrists, "deemed to be expert in forensic diagnoses" (p. 669), to complete two questionnaires based on 12 vignettes. The vignettes described 12 offenders who, in the previous study (Marshall, Kennedy, et al., 2002), had been diagnosed as either a sadist or nonsadist. The first questionnaire instructed the participants to decide whether each of the 12 vignettes described a sadist or a nonsadist, and rate their confidence in this diagnosis on a 5-point scale. The second questionnaire listed 26 offense and offender features, and instructed the participants to use a 5-point scale to rate the relevance of each feature to a diagnosis of sexual sadism, and denote which features were necessary for a diagnosis of sexual sadism.

Diagnostic agreement resulted in a *kappa* of 0.14, or an agreement rate of 21.7%, well below what the authors consider an acceptable level of agreement (90%). Regarding diagnostic importance, the authors found that all participants rated the following features as relevant: (1) control, domination, or power; (2) humiliation or degradation; (3) cruelty or torture; (4) deviant sexual arousal; and (5) sexual mutilation of victims. The authors mention that most features that the participants rated as important for a diagnosis of sexual sadism did not correspond to features that the sexual sadism literature describes as important (e.g., ritualism violence, strangulation, abduction or confinement, anal sex, keeping trophies or records, sexual bondage). Although only 4 participants identified features that they considered necessary to make a diagnosis of sexual sadism, the features they selected did correspond to features that other participants had rated as important for making a diagnosis of sexual sadism (control, domination, or power; humiliation or degradation; cruelty or torture; and sexual mutilation). Still, the authors note that the participants appeared to apply these criteria idiosyncratically as they rendered diagnoses.

The authors concluded that the level of diagnostic agreement for sexual sadism falls short of the appropriate standards for a diagnosis with such serious consequences. Because this study focused on "expert" forensic diagnosticians, it remains unclear how the results would generalize to a more mainstream group of professionals whose work requires an understanding of sexual sadism.

Levenson (2004) addressed the main limitation of Marshall, Kennedy, Yates, and colleagues' (2002) work by examining the agreement of psychiatric diagnoses between two or more independent psychological evaluations of Florida inmates under consideration for SVP status. Based on 277 inmates drawn from a sample of 450, the *kappa* for sexual sadism was 0.30. Rather than suggesting that "experts" diagnose sexual sadism with less reliability than Floridian evaluators, this finding further supports the claim that regardless of the circumstances (vignettes and questionnaires as opposed to "real life" court-ordered psychological evaluations), the reliability of sexual sadism diagnoses falls well below acceptable standards.

Packard and Levenson (2006) reanalyzed Levenson's (2004) data using raw proportions of agreement, odds and risk ratios, and conditional probability estimates as alternative methods to examine diagnostic agreement. They reported odds ratio results indicating that two diagnosticians were 67 times more likely to agree with each other about a sexual sadism diagnosis than disagree. Using relative risk ratios, Packard and Levenson found that a second diagnostician was 53 times more likely to make the same diagnostic decision as the first diagnostician. However, they found low positive predictive power for sexual sadism. Specifically, when a subject received a diagnosis of sexual sadism, there was a .20 probability that a second diagnostician would also diagnose that same subject with sexual sadism. Although Packard and Levenson's results suggest that a diagnosis of sexual sadism has higher interrater reliability than previously accepted (Levenson, 2004; Marshall, Kennedy, Yates, et al., 2002), they still found lower levels of agreement for sexual sadism than for other diagnoses.

More recently, researchers affiliated with Washington State's Special Commitment Center (Jackson, Richards, McCraw, & Koenen, 2006) investigated whether offense behaviors could be parsed out in a way that would reveal reliable differences between sadists and nonsadists. They used features from historical and contemporary sexual sadism literature to construct four categories of offense conduct: (1) planning, onset, and rumination (e.g., careful planning, kept trophies from offense); (2) sexual offense conduct (e.g., vaginal penetration, anal penetration); (3) offense violence other than sexual acts (e.g., struck victim during sex act, use of restraints); and, (4) threats and material exploitation (threats to evoke fear, stealing money from victim). Using a sample of 78 civilly committed SVPs, half of whom had diagnoses of sexual sadism, they found that only 6 of the 47 offense conduct items were reliably related to a diagnosis of sexual sadism. Specifically, results showed that sexual sadists were *less* likely to experience sexual dysfunction during the offense ($p = .007$), and more likely to use threats to evoke fear in their in their victims ($p = .018$), move their victims by force to another location ($p = .040$), and use physical restraints on

their victims during the offense ($p = .047$). Finally, results indicated that only sadists used violence during a sexual act (7 out of 39 sadists, as opposed to none of the 39 nonsadists), and cut or stabbed a victim at any time during an offense (5 out of 39 sadists, as opposed to none of the 39 nonsadists). Results demonstrated that several offense behaviors effectively discriminated between sadists and nonsadists.

Based on these findings, Jackson and colleagues (2006) hypothesized that if violence represents a necessary precursor to a sadist's sexual arousal, then it would make sense for the violence to occur before or during sexual acts. Violence occurring after sexual climax would not serve the purpose of sexual arousal and therefore should not be relevant to a diagnosis of sexual sadism. They contended that because some preoffense violence could be committed to gain the victim's compliance, however, only violence committed after gaining the victim's compliance and prior to the offender's sexual climax could be attributed to the offender's sexual arousal. Granted, this "arousal hypothesis" does raise practical concerns. For example, determining when a certain behavior occurred in the offense sequence (e.g., in relationship to an offender's sexual climax) can prove difficult. Furthermore, behaviors may fulfill two purposes simultaneously (e.g., the same behavior could lead to an offender's sexual arousal *and* secure a victim's compliance).

Present Study

This study sought to determine how professionals who regularly apply and/or make decisions based on a diagnosis of sexual sadism understand behavioral features associated with this diagnosis. Behavioral features, as opposed to inference-based judgments, were selected because preliminary findings hint at their potential to somewhat reliably discriminate between sadistic and nonsadistic offenses (Marshall & Hucker, 2006b; Marshall, Kennedy, Yates, et al., 2002). Based on the literature's diverse characterizations of sexual sadism, we identified four conceptualizations. The first conceptualization includes behaviors found to occur "frequently" during sexually sadistic offenses (Dietz et al., 1990; Gratzer & Bradford, 1995). "Frequently" was operationalized as either "more than 50% of the time" for offenses within Dietz et al.'s (1990) sample, or "significantly more likely to have occurred in an offense perpetrated by a known sexual sadist" for offenses within Gratzer and Bradford's (1995) sample. There were two exceptions to these rules, "forced fellatio" and "vaginal rape." Dietz et al. found that forced fellatio and vaginal rape occurred in more than 50% of offenses committed by sexual sadists, and Gratzer and Bradford found that forced fellatio and vaginal rape occurred at significantly higher frequencies in offenses committed by sexual sadists as compared to nonsadists. Nevertheless, these two behaviors were excluded from the sexual sadism conceptualization and considered "general sexual offending" behaviors for the purpose of analyses. The second conceptualization operationalizes *DSM* (1980-2000) diagnostic criteria wherein an offender considers a

victim's suffering as sexually exciting. The third conceptualization draws from Marshall, Kennedy, Yates, and colleagues' (2002) finding that expert diagnosticians consistently rated themes of control/domination/power, cruelty/torture; sexual mutilation; and humiliation/degradation as highly relevant to diagnosing sexual sadism. The fourth conceptualization stems from Jackson and colleagues' (2006) arousal hypothesis, and was operationalized as six offense behaviors specified as occurring *during* the sexual act. These conceptualizations were tested by four exploratory analyses designed to describe how professionals conceptualize sexual sadism.

In consideration of the amount of research that has found homogeneity between sadists and nonsadists, and low interrater reliability for sexual sadism diagnoses (Levenson, 2004; Marshall, Kennedy, Yates, et al., 2002), our first hypothesis was that ratings for behaviors associated with sexual sadism would, on average, not significantly differ from ratings for behaviors associated with general sexual offending. However, it was anticipated that the six behaviors indicative of Jackson and colleagues' (2006) arousal hypothesis would receive significantly higher-than-neutral ratings. It was also expected that "sexual dysfunction during the offense" would receive ratings significantly below the neutral value.

With regard to the four conceptualizations of sexual sadism, it was hypothesized that participants' ratings would support at least one of these conceptualizations. Because of the exploratory nature of this question, no predictions were made about which conceptualization would receive the highest or lowest overall ratings.

Method

Participants

We used a purposive sampling frame that included individuals who subscribed to either the Association for the Treatment of Sexual Abusers' (ATSA) e-mail list or the American Psychology-Law Society's (AP-LS) e-mail list. This sampling frame closely matches the target population of professionals who have worked with/made decisions based on a diagnosis of sexual sadism. Granted, not every single professional subscribes to one of these e-mail lists, nor does every single subscriber work with individuals who qualify for a diagnosis of sexual sadism. Still, these e-mail lists reach thousands of professionals with different levels of experience, some of whom are likely to have worked with individuals who qualify for a diagnosis of sexual sadism. The recruitment e-mail message contained an active link to the anonymous and confidential research survey, allowing interested individuals to access the survey at their convenience.

Sixty participants completed the survey. Thirty-six participants (60%) reported learning about the survey through the ATSA e-mail list. Six participants (10%) reported learning about the survey through the AP-LS e-mail list, and 18 participants

Table 1
Participants' Years of Experience and Diagnostic Considerations

	ATSA			AP-LS			Forwarded E-mail Message			Overall		
Participant Information	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Age	36	46.72	12.57	6	30.67	4.68	17	41.29	9.37	59	43.53	12.10
Years of experience	34	13.93	10.90	6	4.50	3.62	18	7.33	6.10	58	10.91	9.74
Considered making diagnosis	35	8.69	11.58	6	3.67	8.04	18	12.78	25.76	59	9.42	16.91
Diagnosed "yes"	35	3.46	5.18	6	0.83	2.04	18	1.33	2.68	59	2.54	4.41
Diagnosed "no"	34	4.50	7.18	6	2.83	6.01	17	10.12	25.41	57	6.00	15.02

Note: ATSA = Association for the Treatment of Sexual Abusers; AP-LS = American Psychology–Law Society.

(30%) reported learning about the survey from a recruitment e-mail forwarded to them by a colleague. Participants' mean ages are summarized in Table 1. There were significant mean differences in age among ATSA, AP-LS, and forwarded e-mail message participants, $F(2, 56) = 5.74$, $p = .005$. Pairwise comparisons using least significant difference follow-ups (using a minimum significant mean difference of 4.14) showed that ATSA participants reported a significantly higher mean age than AP-LS participants, although no other groups significantly differed agewise. ATSA participants did not differ significantly from ATSA's 2005 demographic information, $t(35) = -1.57$, $p > .05$ (J. Gruber, personal communication, June 1, 2005).

With regard to education, the majority (66.7%) of participants reported having either a master's degree ($n = 21$, 35.0%) or a doctoral degree ($n = 19$, 31.7%). Another 13% ($n = 8$) were enrolled in doctoral programs but had not yet completed their degree. The remaining participants reported either associate's degrees ($n = 3$, 5%), bachelor's degrees ($n = 4$, 6.7%), medical or juris doctorate degrees ($n = 4$, 6.7%), and one participant reported both a doctoral degree and a juris doctorate.

Participants were asked to specify the capacities in which they had worked with individuals potentially eligible for a diagnosis of sexual sadism. Response options included individual treatment, group treatment, evaluation, assessment, legal, and a write-in response. For the purpose of data analysis, participants' responses were recoded into a single variable that characterized professional experience as: therapy; evaluation/assessment; legal; therapy and evaluation/assessment; therapy, evaluation/assessment, and legal; and other. Table 2 displays participants' professional backgrounds.

Participants were asked to estimate the total number of times they had considered making a diagnosis of sexual sadism. On average, participants had considered making a diagnosis of sexual sadism 9.42 times ($n = 59$, $SD = 16.91$), with responses ranging

Table 2
Participants' Professional Backgrounds

Background	*N*	%
Therapy	7	11.7
Evaluation	3	5.0
Legal	5	8.3
Therapy/Evaluation	29	48.3
Therapy, Evaluation, and Legal	13	21.7
Other[a]	3	5

a. Two of these participants expressed that they thought they did not work with individuals who would qualify for a diagnosis of sexual sadism, and the third reported working as an administrator at a residential sex offender treatment facility.

from 0 to 100 (41.7% of participants [$n = 25$] had never made a diagnosis of sexual sadism, and 56.7% of participants [$n = 34$] had made a diagnosis of sexual sadism at least once). Subsequently, participants were asked to estimate the number of times they had made a diagnosis of sexual sadism, as well as the number of times they had decided against making a diagnosis. Participants had made an average of 2.54 diagnoses of sexual sadism ($n = 59$, $SD = 4.41$), with responses ranging from 0 to 25. Overall, participants had decided against making a diagnosis 6 times ($n = 57$, $SD = 15.02$), with responses ranging from 0 to 98.

Materials

A secure server, maintained by the University of Washington's Catalyst Group, hosted the online survey (see appendix for survey items from the "Sexual Offending Behaviors" section) and compiled responses. The online survey was available to participants for a time window of one year and seven months, opening on July 11, 2005 and closing on February 11, 2007. Peripheral materials included the participants' computers and internet connections, through which they accessed the survey. The study's only task was the anonymous and confidential online survey. The survey included a forced-choice informed consent page, 62 items relevant to the behavioral features of sexual sadism and general sexual offending (i.e., participation of a coperpetrator, digital penetration, use of threats to gain victim's compliance) and 13 demographic questions. Survey items were culled from an extensive literature review that included previous studies of behavioral features associated, to varying degrees, with sexual sadism (Dietz et al., 1990; Gratzer & Bradford, 1995; Jackson et al., 2006; Marshall, Kennedy, Yates, et al., 2002; H. Richards, personal communication, February 9, 2005), and presented as 62 different statements.

Participants read instructions to rate each statement on a seven-point Likert-type scale ranging from *not at all essential* to *absolutely essential* for making a diagnosis

of sexual sadism. Based on literature stating that Likert-type scale midpoints can represent either a true neutral or an undecided response (Raaijmakers, van Hoof, Hart, Verbogt, & Vollebergh, 2000), this study's Likert-type scale midpoint was not explicitly anchored as it was not critical to distinguish between participants' neutral and undecided responses.

Procedure

Informed consent procedures were approved by the governing institutional review board and obtained from all participants. The recruitment message described the inspiration for this study, outlined the sequential components of the survey, explained that all survey responses will be anonymous, emphasized the voluntary nature of participation, and provided an estimate of how long it will take to complete the survey (15-20 min). Investigators' contact information was provided. Individuals could access the survey by clicking on the live link embedded in the e-mail. When the link was clicked, the survey would open in a new Web browser window. The first page presented the survey's information statement. Below the text, participants were asked to select one of the following options: (1) I do not want to proceed any further. (If so, please click "Cancel" to exit the survey); or, (2) I would like to proceed and fill out the survey. (In the text box, please write "consent" and click "Next"). People had to select one of these two options before proceeding. If the first option was selected, the survey window closed and the individual exited the survey. If the second option was selected, the individual became a participant and proceeded to a page of instructions for completing the first section of the survey. From here, participants clicked "previous," next," or "cancel" buttons to navigate through the survey. After participants submitted their survey responses, a debriefing page automatically opened in a new Web browser window.

Analyses and Results

To determine whether behaviors that prior research and theory had associated with sexual sadism received significantly different ratings than behaviors associated with general sexual offending, items were first rationally divided into two mutually exclusive categories, sexual sadism (39 items) and general sexual offending (23 items), and compared to a neutral rating (4) using a one-sample *t* test. Results are presented in Table 3. Within the sexual sadism category, 15 items (38.4%) received ratings that were significantly higher than the neutral point (the hypothesized direction), whereas ratings for the remaining 24 items (61.5%) were not significantly different from the neutral point. Within the general offending category, 21 items (91.3%) received ratings significantly lower than the neutral point (the hypothesized direction), whereas ratings for the remaining 2 items (8.7%) were not significantly different

Table 3
Ratings of Behaviors Associated With Sexual Sadism Versus Behaviors Associated With General Sexual Offending

Sexual Sadism					General Sexual Offending				
Behavior	Difference From Neutral (4)	*t*	*df*	*p* (2-tailed)	Behavior	Difference From Neutral (4)	*t*	*df*	*p* (2-tailed)
Careful planning of the offense	–0.37	–1.41	59	.163	Offender under influence of drugs	–2.33	–14.07	59	.000
Rape kit	–0.27	–1.06	59	.293	Approached victim under pretense	–0.93	–3.97	59	.000
Forcefully moved victim to another location	–0.45	–1.78	59	.085	Used impersonation to access victim	–1.05	–4.20	59	.000
Recorded offenses in journal	–0.30	–1.21	59	.230	Used deception to lure victim	–0.85	–3.29	59	.002
Collected trophies of offense	0.17	0.64	59	.525	Participation of coperpetrator	–1.67	–7.86	59	.000
Slapped or punched victim prior to sexual act	0.25	0.97	58	.336	Victim forced to masturbate offender	–1.57	–8.41	59	.000
Slapped or punched victim after sexual act	0.22	0.86	58	.395	Digital penetration	–1.55	–8.19	59	.000
Strangled victim prior to sexual act	0.66	2.39	57	.020	Sexual dysfunction during offense[a]	–1.67	–8.34	59	.000
Strangled victim after sexual act	0.42	1.54	58	.130	Offender achieved orgasm (by any means)	–0.77	–2.26	59	.027
Cut or stabbed victim prior to sexual act	0.70	2.74	59	.008	Forced victim to say words or phrases, or show enjoyment	0.49	1.94	58	.057
Cut or stabbed victim after sexual act	0.57	2.13	59	.037	Forced victim to masturbate or sexually touch self	–0.69	–3.10	58	.003
Slapped or punched victim during sexual act[a]	0.98	3.90	59	.000	Performed oral sex on victim	–1.40	–6.84	58	.000

(continued)

Table 3 (continued)

Sexual Sadism					General Sexual Offending				
Behavior	Difference From Neutral (4)	*t*	*df*	*p* (2-tailed)	Behavior	Difference From Neutral (4)	*t*	*df*	*p* (2-tailed)
Cut, stabbed, strangled, bit, or beat victim during sexual act[a]	1.33	5.31	59	.000	Gently or affectionately touched victim	–2.02	-9.70	58	.000
Bit victim at any point	0.68	2.58	58	.013	Used threats to gain compliance or prevent future reporting	–0.37	–1.41	58	.164
Physical restraints used prior to sexual acts	0.60	2.26	59	.028	Attempted to verbally calm or comfort victim	–1.87	–10.48	59	.000
Physical restraints used after sexual act	0.43	1.61	59	.113	Indicated desire or interest in a future relationship with victim	–1.87	–9.86	59	.000
Physical restraints used during sexual act[a]	0.62	2.42	59	.019	Accommodated to victim's needs	–2.22	–12.80	59	.000
Maintained threat with a weapon during sexual acts[a]	0.53	1.94	48	.058	Took money (or purse, wallet) from victim	–1.72	–8.23	59	.000
Covered victim's face with hand/arm or blindfold prior to sexual act	–0.21	–0.74	47	.462	Took articles of marketable value from victim's home or car	–1.71	–8.02	58	.000
Covered victim's face with hand/arm or blindfold after sexual act	–0.06	–0.23	48	.820	No medical attention; postrape exam/testing only	–1.80	–8.54	58	.000
Covered victim's face with hand/arm during sexual act[a]	–0.13	–0.44	47	.663	ER/doctor's office (single or few visits)	–1.10	–4.92	58	.000

(continued)

Table 3 (continued)

Sexual Sadism					General Sexual Offending				
Behavior	Difference From Neutral (4)	*t*	*df*	*p* (2-tailed)	Behavior	Difference From Neutral (4)	*t*	*df*	*p* (2-tailed)
Covered victim's face with blindfold during sexual act[a]	–0.08	–0.29	48	.776	Forced fellatio	–1.07	–4.54	59	.000
Used threats to evoke fear (not to gain compliance)	1.48	7.04	59	.000	Vaginal rape	–0.97	–3.84	59	.000
Gleeful or triumphant affect, or arrogant statements	0.63	2.59	59	.012					
Took item(s) of personal significance	–.033	–1.32	59	.192					
Anal rape	–0.38	–1.43	59	.159					
Offender used physical objects (other than own body) in sexual manner toward victim	0.35	1.39	59	.171					
Used body secretions/excretions to humiliate victim	1.25	5.38	59	.000					
Used degrading or humiliating language toward victim	1.23	5.37	59	.000					
Serious injury requiring extended care (but not hospitalization)	0.02	0.06	59	.951					

(continued)

Table 3 (continued)

Sexual Sadism					General Sexual Offending				
Behavior	Difference From Neutral (4)	*t*	*df*	*p* (2-tailed)	Behavior	Difference From Neutral (4)	*t*	*df*	*p* (2-tailed)
Hospitalized for a day or more	0.20	0.72	59	.475					
Permanent disfigurement or disability	0.63	2.07	59	.043					
Death	0.40	1.31	59	.194					
Cut that required stitches	0.03	0.13	57	.899					
Broken bones	0.15	0.54	59	.593					
Any facial injury	0.32	1.12	58	.269					
Inflicted pain to sexual area from the use of a physical object	1.32	5.11	59	.000					
Inflicted pain to sexual area not because of use of physical object	0.91	3.54	57	.001					
Significant injury to sexual area because of penile penetration	0.33	1.30	59	.197					

a. Items indicative of Jackson, Richards, McCraw, & Koenen's (2006) arousal hypothesis.

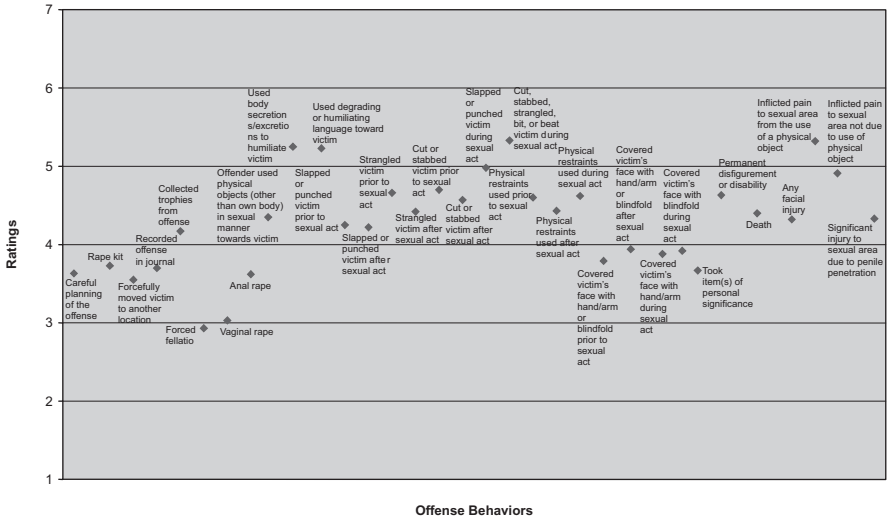
from the neutral point. Contrary to the research hypothesis, overall ratings for the two categories differed significantly, $F(1, 60) = 128.15$, $p < .001$. Specifically, items within the sexual sadism category received a mean rating of 4.38 ($SD = .52$), which was significantly above a neutral rating, $t(38) = 4.56$, $p < .001$, and items within the general sexual offending category received a mean rating of 2.67 ($SD = .66$), which was significantly below a neutral rating, $t(21) = -9.73$, $p < .001$.

Another "portion" of this hypothesis predicted that the six behaviors indicative of Jackson and colleagues' (2006) arousal hypothesis would receive ratings significantly above a neutral value, whereas "sexual dysfunction during offense" would receive ratings significantly below a neutral value. To test this hypothesis, *t* tests were used to compare ratings for each of these behaviors against a neutral rating. Three out of the six behaviors received ratings significantly above a neutral rating. The remaining three behaviors received higher-than-neutral ratings, but not at levels that reached statistical significance, As hypothesized, "sexual dysfunction during the offense" received ratings significantly below a neutral value, signifying that professionals do not consider the presence of this behavior as indicative of sexual sadism. Results, presented as the italicized portions of Table 3, partially support this research hypothesis.

This study's third hypothesis stated that participants' ratings would support at least one of the four conceptualizations of sexual sadism. To examine this hypothesis, items were categorized according to their fit with at least one of the four conceptualizations of sexual sadism. A few items contributed to all four conceptualizations, some items contributed to only a few conceptualizations, and other items were unique to just one conceptualization.

Next, the mean rating for each item within each conceptualization was computed. Finally, patterns of support for each of the four conceptualizations were modeled by plotting the mean ratings for each item, with items grouped according to their respective conceptualization (see Figures 1, 2, 3, and 4). Support for each conceptualization was examined by aggregating the means of individual items within a particular theory and comparing the mean to a neutral rating. Three of the four conceptualizations received support that significantly exceeded neutral ratings. Specifically, the 33 items within the Dietz and colleagues (1990)/Gratzer and Bradford (1995) conceptualization had an overall mean of 4.28 ($SD = .62$) and, overall, significantly differed from a neutral rating, $t(32) = 2.56$, $p = .015$. Likewise, the 26 items within the *DSM* (1980-2000) conceptualization had an overall mean rating of 4.376 ($SD = .56$) and were significantly different from a neutral rating, $t(26) = 3.51$, $p = .002$. The 25 items within the Marshall, Kennedy, Yates, and colleagues' (2002) conceptualization had an overall mean rating of 4.53 ($SD = .51$) and, overall, were significantly different when compared to a neutral rating, $t(24) = 5.12$, $p < .001$. The 6 items representing Jackson and colleagues' (2006) arousal hypothesis had an overall mean of 4.54 ($SD = .61$), which exceeded the neutral rating, but not sufficiently to achieve statistical significance, $t(5) = 2.30$, $p > .05$ ($p = .069$). We speculate that this mean value would have achieved significance had the survey had included more offense behaviors specified as occurring during a sexual act.

Figure 1
Mean Importance Ratings for Each Offense Behavior: Dietz, Hazelwood, & Warren (1990) and Gratzer & Bradford (1995) Variables

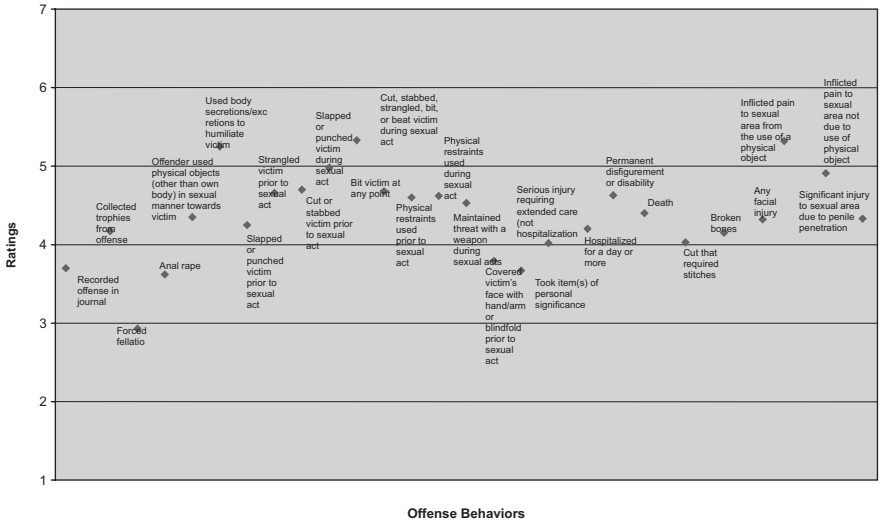


Discussion

Results from this study describe how professionals understand behavioral features associated with a diagnosis of sexual sadism. The first hypothesis, which anticipated that professionals would not reliably discriminate between behaviors traditionally associated with sexual sadism versus those traditionally associated with general sexual offending, was not supported. Rather, results indicated that professionals reliably distinguished between behaviors traditionally associated with sexual sadism versus those traditionally associated with general sexual offending. Approximately one third of behaviors traditionally associated with sexual sadism received significantly higher-than-neutral ratings. The remaining two thirds of behaviors traditionally associated with sexual sadism received ratings that either fell significantly below a neutral value, or else did not significantly differ from a neutral value. Nearly all of the behaviors traditionally associated with general sexual offending received ratings significantly below a neutral value. Of these 22 items, only 2 received ratings that did not significantly differ from a neutral value.

The second component of the present study revealed that three of the four conceptualizations received overall ratings that significantly exceeded a neutral value,

Figure 2
Mean Importance Ratings for Each Offense Behavior: *Diagnostic and Statistical Manual of Mental Disorders* (1980-2000) Criteria Variables

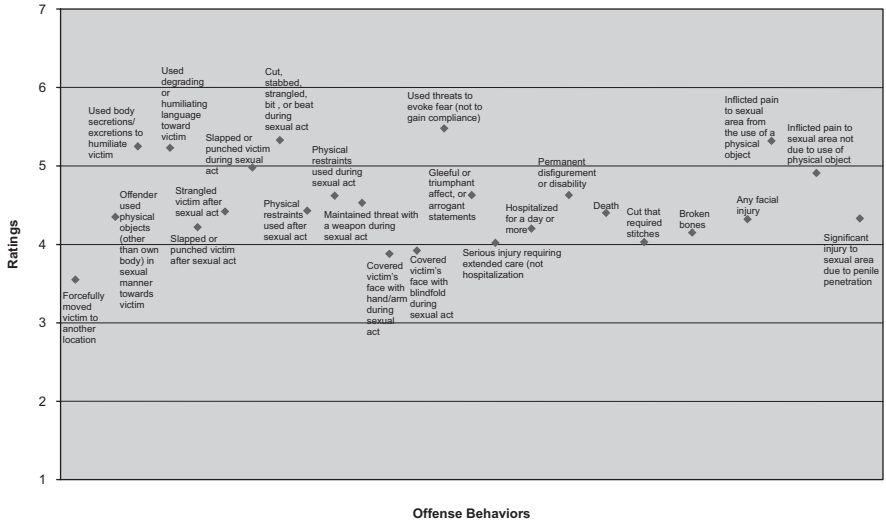


and out of those three conceptualizations, no one conceptualization received significantly more or less support than any other conceptualization. This finding is reminiscent of Rosenzweig's (1936) "Dodo Bird Verdict," [6] and raises the question of whether elements common to these conceptualizations may have led to the similar expressions of support. Future research could shed light on the identity of common factors and, presuming their existence, could operationalize these factors in ways that would facilitate diagnostic decision making and improve diagnosticians' interrater reliability. The following section represents a preliminary step in that direction by listing each of the 15 behaviors that professionals endorsed as indicative of sexual sadism, and exploring their relationships to one another.

Behaviors Associated With Sexual Sadism: Toward More Parsimonious Conceptualization?

The current study began by culling offense-related behaviors from 4 conceptualizations of sexual sadism. As Table 4 indicates, many of the 15 behaviors that professionals rated as important to diagnosing sexual sadism relate to multiple

Figure 3
Mean Importance Ratings for Each Offense Behavior: Marshall, Kennedy, Yates, & Serran (2002) Variables

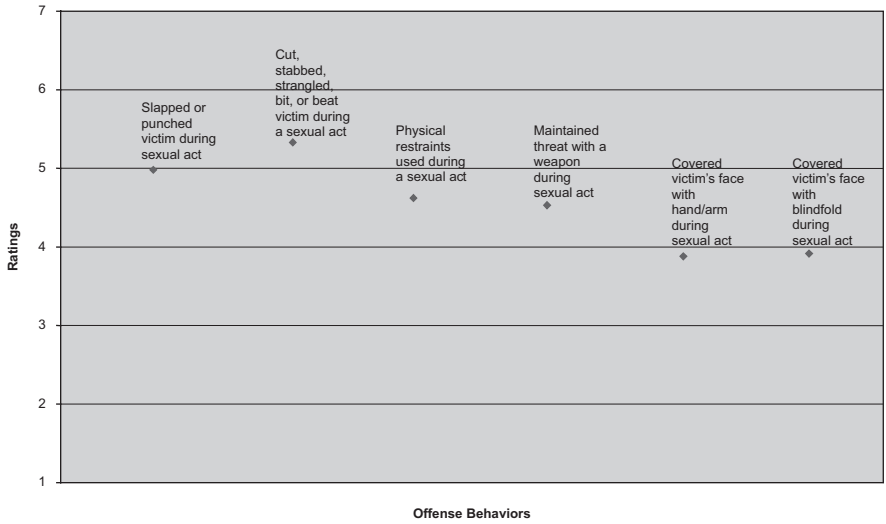


conceptualizations. The overlap between conceptualizations suggests the possibility of constructing a more parsimonious conceptualization of sexual sadism. This section represents a preliminary step in that process by examining the relationships between these 15 behaviors.

Myriad frameworks could offer an organizational structure for these behaviors. Therefore, this section offers ideas, rather than conclusions, about deriving common themes that may tie these behaviors together. The first attempt to derive common themes proceeds by examining behaviors in order of importance ratings (highest to lowest), and the second attempt proceeds by examining behaviors in the order of how strongly they relate to each of the four conceptualizations (beginning with behaviors that relate to each of the four conceptualizations).

To facilitate the first method, Table 4 presents the 15 behaviors in the order of importance ratings. However, aside from fitting the overinclusive category of inflicting physical or emotional suffering/humiliation to a victim, the first three behaviors do not appear to fit together in any meaningful way. Expanding this inquiry to include the first five behaviors yields little improvement, as the newly-included behaviors do not modify the themes of physical pain and emotional suffering/humiliation. Not

Figure 4
Mean Importance Ratings for Each Offense Behavior: Jackson, Richards, McCraw, & Koenen (2006) Variables—"Arousal Hypothesis"



even examining the first eight behaviors appears to modify these themes. This method of examining behaviors according to importance level may be better suited for constructing a checklist of behaviors indicative of sexual sadism, rather than a cohesive conceptualization of sexual sadism. A well-constructed checklist has the potential to assist diagnostic decision making, and represents a promising topic for future research.

Another promising construction method is to first examine behaviors that relate to all four conceptualizations, then proceed to examine behaviors that relate to three conceptualizations, and so on. The following behaviors relate to all four conceptualizations: slapped or punched victim during sexual act; cut, stabbed, strangled, bit, or beat victim during sexual act; and, physical restraints used during sexual act. Not only do these behaviors fit within each of the four conceptualizations, but they also represent the three supported behaviors within the arousal hypothesis. Thus, these behaviors—and possibly similar behaviors that were not included within the present study—may be necessary, albeit not sufficient, for a conceptualization of sexual sadism. These three behaviors share the common feature of occurring *during* sexual acts.

The following behaviors relate to three conceptualizations: used body secretions/excretions to humiliate victim; permanent disfigurement or disability; inflicted

Table 4
Relationships Between Behaviors Rated as Important to a Diagnosis of Sexual Sadism and the Four Conceptualizations

Behavior	Mean Importance Rating	Dietz, Hazelwood, & Warren (1990)/Gratzer & Bradford (1995)	*DSM* (1980-2000) Criteria	Marshall, Kennedy, Yates, & Serran (2002)	Jackson, Richards, McCraw, & Koenen (2006)
Used threats to evoke fear (not to gain compliance)	5.48			X	
Cut, stabbed, strangled, bit, or beat victim during sexual act	5.33	X	X	X	X
Inflicted pain to sexual area from the use of a physical object	5.32	X	X	X	
Used body secretions/excretions to humiliate victim	5.25	X	X	X	
Used degrading or humiliating language toward victim	5.23	X		X	
Slapped or punched victim during sexual act	4.98	X	X	X	X
Inflicted pain to sexual area not because of use of physical object	4.91	X	X	X	
Cut or stabbed victim prior to sexual act	4.70	X	X		
Bit victim at any point	4.68		X		
Strangled victim prior to sexual act	4.66	X	X		
Gleeful or triumphant affect, or arrogant statements	4.63			X	
Permanent disfigurement or disability	4.63	X	X	X	
Physical restraints used during sexual act	4.62	X	X	X	X
Physical restraints used prior to sexual act	4.60	X	X		
Cut or stabbed victim after sexual act	4.57	X			

Note: *DSM = Diagnostic and Statistical Manual of Mental Disorders*.

pain to sexual area from the use of a physical object; inflicted pain to sexual area not because of use of physical object. As it turns out, these behaviors relate to each of the three supported conceptualizations. Thus, these (and similar) behaviors likely represent important information for a conceptualization of sexual sadism. The last three behaviors share the common feature of inflicting lasting harm to the victim, whereas the first behavior conveys humiliation, a central theme within historical—and some present—conceptualizations of sexual sadism. So far, this approach to bottom-up construction has identified violence during sexual acts, humiliation, and inflicting lasting harm to the victim as potentially important themes underlying the list of behaviors.

Considering how to treat behaviors that relate to only one or two conceptualizations illustrates that such an extraction process would be better suited to a study that tested more than 4 conceptualizations. Still, this process has derived themes that underlie 7 of the 15 behaviors, and seem reasonably related to other conceptualizations within the relevant literature. However, it remains uncertain whether such an extraction process would yield a better, as opposed to just *another,* conceptualization. Future research that pursues the bottom-up construction of a conceptualization of sexual sadism should be sure to test the utility of any new conceptualization against extant conceptualizations. A new conceptualization will be only as good as its ability to improve the reliability and validity of diagnostic decision making.

This section has framed the present study as a preliminary step in the process of developing a more parsimonious conceptualization of sexual sadism that would, hopefully double as a means of improving diagnostic decision making. The following section returns to the central issue of diagnostic decision making, and examines how the present study relates to other investigations of how professionals diagnose sexual sadism.

Methodological Differences Between Studies

Several studies convincingly contend that professionals diagnose sexual sadism with surprisingly low levels of agreement (Levenson, 2004; Marshall, Kennedy, Yates, et al., 2002). In addition, Marshall, Kennedy, and colleagues (2002) found that features traditionally associated with sexually sadistic offenses failed to differentiate sexual sadists from nonsadistic sexual offenders. Still, a handful of studies support the presence of behaviors unique to sexual sadism, as opposed to nonsadistic sexual offenses (Gratzer & Bradford, 1995; Jackson et al., 2006). The present study most closely relates to the handful of diagnostic reliability studies, but yields surprising results—that professionals reliably distinguish between sadistic and nonsadistic behaviors. However, before concluding that professionals may, indeed, be able to reliably discriminate between sadistic and nonsadistic behaviors, it seems reasonable to compare and contrast this study's methodology to studies that have asked similar questions and found different results.

Each study that investigated how professionals understand diagnoses of sexual sadism used different samples and sampling procedures, had participants complete different research tasks under different conditions, and measured responses differently. Regarding the sample and sampling procedures and research tasks, Levenson (2004) collected archival data from the Florida correctional system, whereas Marshall, Kennedy, Yates, and colleagues (2002) used highly purposive sampling to invite forensic experts to complete a pencil/paper task. The present study used purposive sampling and invited professionals who reported a variety of experiences with sexual offenders to rate sexual offense behaviors regarding their importance for a diagnosis of sexual sadism.

Furthermore, the conditions under which participants completed the research tasks were substantially different. When the Florida psychological evaluators made diagnoses that Levenson's study used, they did not know that their diagnostic decisions would become data for a research study. On the other hand, the expert forensic evaluators who provided diagnoses for the vignettes in Marshall, Kennedy, Yates, and colleagues' (2002) study likely had some awareness of the study's purpose, especially given the lead author's reputation in this area of research. The present study's data came from participants who reported varied professional experiences with sexual offenders, and knew that the study aimed to examine how professionals understood sexual sadism.

Another difference between studies relates to how diagnostic agreement was calculated. Marshall, Kennedy, Yates, and colleagues (2002) and Levenson (2004) used *kappa,* which indicated low levels of agreement. On the other hand, Packard and Levenson (2006) reported higher levels of diagnostic agreement for sexual sadism using raw proportions of agreement, odds and risk ratios, and estimates of conditional probabilities. Packard and Levenson noted several limitations of using *kappa* to measure interrater reliability. First, *kappa* assumes statistical independence between raters, which would pose a problem if one diagnostician knew of another's diagnosis (Sim & Wright, 2005). In addition, *kappa* assumes approximately equal values in each diagnostic category of a contingency table, and yields biased results when proportions of agreement within the contingency table considerably differ (Feinstein & Cicchetti, 1990; Sim & Wright, 2005). Finally, different authors follow different guidelines for calculating (Uebersax, 1987) and interpreting (Bloom, Fischer, & Orme, 1999; Landis & Koch, 1977) *kappa.* Small values of *kappa* do not necessarily represent low interrater reliability, which challenges results that have interpreted low *kappa* values as support for the unreliability of a sexual sadism diagnosis.

With regard to other measurement differences, Levenson (2004) responded directly to the question of interrater reliability by using archival data to compute a *kappa* for sexual sadism diagnoses. Levenson's findings are relatively consistent with those of Marshall, Kennedy, Yates, and colleagues (2002), who had expert forensic evaluators decide whether to assign a diagnosis of sexual sadism to 12 vignettes (6 of which described sexual offenders with a diagnosis of sexual sadism, 6 of which described

sexual offenders without a diagnosis of sexual sadism). Unlike the previous 2 studies, the present study did not have professionals make actual diagnoses. Rather, professionals simply rated the diagnostic importance (or lack thereof) of sexual offense behaviors. Whereas the present study's task required making decisions along a graded continuous scale (a 1-7 Likert-type scale), rendering a diagnosis requires making decisions on a dichotomous yes/no scale. Compared to dichotomous yes/no questions, Likert-type scales yield increased scale reliability (Finn, 1972; Lissitz & Green, 1975; McMordie, 1979), albeit with diminishing returns beyond a 1 through 7 range (Nunnally, 1978). Perhaps the results from this study would have shown lower levels of reliability had it used dichotomous yes/no response options, rather than Likert-type scales. In short, the present study examined how professionals ascribe importance to behaviors associated with sexual sadism—a related, but different question than diagnostic interrater reliability.

In conclusion, these methodological differences make it difficult to directly compare the findings of the present study with the extant literature on how professionals understand diagnoses of sexual sadism. Still, the present study lends credence to the conclusion that under some conditions, professionals can reliably distinguish sadistic from nonsadistic behaviors, and raises the question of how this ability translates into making diagnoses.

Limitations

The present study included significant limitations related to its sample. First, the present study drew results from a small sample size that included graduate students and professionals who had never considered making a diagnosis of sexual sadism. Because many students and professionals may, one day, consider making a diagnosis of sexual sadism, there is value in understanding how such individuals understand the relative importance of behaviors associated with sexual sadism versus behaviors associated with general sexual offending. That said, these participants' lack of hands-on experience with this diagnosis may have affected how they responded to survey items. Future research may benefit from implementing more stringent inclusion criteria with regard to participants' diagnostic experience. Second, it is uncertain what led to such a low response rate. The possibility that those who did complete the survey differ from those who did not complete the survey cannot be ruled out. This means that the present study's results may not generalize to other groups of professionals. Future research that uses a similar methodology to investigate a similar question may benefit from implementing different recruitment methods. Finally, the small sample size means that this study lacked power to detect smaller effects. Replicating this study with a larger sample may reveal that professionals rate more than just 15 behaviors as indicative of sexual sadism, and may yield support for Jackson and colleagues' (2006) arousal hypothesis.

Implications and Directions for Future Research

Studies that investigate how professionals make diagnostic decisions based on offense behaviors, as well as those that examine differences between sadistic and nonsadistic sexual offense behaviors, constitute a small body of literature with mixed results. Some evidence suggests that there are few behavioral differences between sadistic and nonsadistic sexual offenders, a finding contradicted by Jackson and colleagues' (2006) research. Other evidence suggests that a diagnosis of sexual sadism shares low levels of agreement, a finding called into question—although not contradicted—by the present study.

The present body of literature does not yet contain enough evidence to support firm conclusions about the construct validity, or lack thereof, of sexual sadism. This means that professionals who make this diagnosis and/or use this diagnosis for decision-making purposes should, at the very least, proceed with caution. Furthermore, diagnosticians and professionals in a position to evaluate the suitability of a patient's sexual sadism diagnosis may wish to consider the relevance of violence committed for reasons other than gaining the victim's compliance. The presence of such behaviors would provide theory-based support for a sexual sadism diagnosis.

On a related note, a parsimonious and clearly defined conceptualization would improve diagnostic decision making and increase the reliability and validity of sexual sadism diagnoses. Although the present study lends credence to the conclusion that under some conditions, professionals can reliably distinguish sadistic from nonsadistic behaviors, it raises the question of how this ability translates into diagnostic decision making. To this effect, future research should continue to investigate the relevance of certain offense behaviors—individually as well as within the framework of theory-based conceptualizations—that received significant ratings in the expected directions. Furthermore, future studies should encourage participants to explain the rationales behind their ratings. Results from such research would equip diagnosticians with behavioral criteria that more closely track the construct of sexual sadism, leading to correspondingly increased levels of agreement between diagnosticians and more valid decisions.

Appendix
Professional Perspectives on Sexual Sadism Survey Items on Sexual Offending Behaviors

PLANNING, ONSET, and RUMINATION

1. Careful planning of the offense
2. Utilized a rape kit of prearranged implements and tools
3. Offender under influence of substance(s) at the time of the offense
4. Approached victim under pretext

(continued)

Appendix (continued)

5. Used impersonation to gain access to victim
6. Used deception to lure victim to another location
7. Moved victim to another location under force/coercion
8. Participation of a coperpetrator in the sexual assault
9. Recorded offenses in a journal or log
10. Kept trophies of the offense
11. Duration of the assault, *please explain*

OFFENSE SEXUAL CONTACT

12. Victim forced to masturbate offender
13. Digital penetration
14. Forced fellatio
15. Vaginal rape
16. Anal rape
17. Sexual dysfunction (inability to maintain erection, premature ejaculation) during offense
18. Achieved orgasm during the offense (by any means)
19. Used physical objects (other than own body) in a sexual manner toward victim
20. Used body secretions/excretions to humiliate victim (oral-anal contact included, fellatio and anal or vaginal intercourse not included)
21. Used degrading or humiliating language toward victim
22. Forced victim to say words or phrases, or to show enjoyment of the offense conduct
23. Forced victim to masturbate or sexually touch self
24. Performed fellatio on victim (or manually masturbated male victim)
25. Gentle or affectionate touching of the victim

OFFENSE VIOLENCE OTHER THAN SEXUAL ACTS

26a. Used a gun in the offense
26b. Used a knife in the offense
26c. Used a different weapon in the offense
27a. Slapped or punched victim prior to any sexual act
27b. Slapped or punched victim after any sexual act
28a. Strangled victim prior to any sexual act
28b. Strangled victim after any sexual act
29a. Cut or stabbed victim prior to any sexual act
29b. Cut or stabbed victim after any sexual act
30. Slapped or punched victim during a sexual act
31. Cut, stabbed, strangled, bit, or beat victim during a sexual act
32. Bit the victim at any point
33a. Physical restraints used prior to sexual acts
33b. Physical restraints used after sexual acts
34. Physical restraints used during a sexual act
35. Maintained threat with weapon during sexual assault
36a. Covered victim's eyes with hand/arm, blindfold prior to any sexual act

(continued)

Appendix (continued)

36b. Covered victim's eyes with hand/arm, blindfold after any sexual act
37a. Covered victim's eyes with hand/arm during sexual act
37b. Covered victim's eyes with blindfold during sexual act

THREATS and MATERIAL EXPLOITATIONS

38. Used threats to evoke fear (not to gain compliance)
39. Used threats to gain compliance or prevent future report of crime
40. Gleeful or triumphant affect, or arrogant statements
41. Attempted to verbally calm or comfort the victim
42. Indicated desire or interest in a future mutual relationship with victim
43. Accommodated to victim's needs
44. Took item(s) of personal significance
45. Took money (or purse) from the victim
46. Took article(s) of marketable value from home or car

VICTIM'S INJURIES

47. No medical attention, postrape examination/testing only
48. ER or doctor's office (outpatient, single/few visits) care for injuries
49. Serious injury requiring extended care (but no record of extended hospitalization)
50. Hospitalized for a day or more
51. Permanent disfigurement or disability
52. Death
53. Cut that required stitches
54. Broken bones
55. Any facial injury
56. Inflicted pain to sexual area from the use of a physical object
57. Inflicted injury to sexual area from the use of a physical object
58. Inflicted pain to sexual area not because of use of a physical object
59. Inflicted injury to sexual area not because of use of a physical object
60. Significant injury to sexual area because of penile penetration

Notes

1. This article will use labels of *sadist* and *nonsadist* to designate *an individual with a diagnosis of sexual sadism* and *an individual without a diagnosis of sexual sadism*. Although using sadist and nonsadist contradicts the article's investigation of the validity of a sexual sadism diagnosis, these terms are less cumbersome than behaviorally-based designations.

2. North Dakota is the only state requiring that the individual simply "be shown" to have engaged in sexually predatory behavior (Civil Commitment, 1997).

3. This article does not claim that Sade was the original sadist and/or the original author of sadistic stories. The significance of Sade's work and life (as evidenced through his work) lies in the fact that it remains the most comprehensive sympathetic account of sadistic behavior. Even today, it inspires reflection on and, at the same time, shapes the phenomenon of sadism.

4. This article will use masculine pronouns to refer to sadists. The decision is based on the objective of fostering a straightforward writing style, and is not meant to imply that only males can be sadists. According to research, however, men express sadistic preferences at a higher rate than women. Krafft-Ebing himself (1886/1965) asserted that sadistic acts were more common among males, under the

assumption that men were naturally inclined to subjugate women. Contemporary researchers have examined gender-specific rates of sadism. Kinsey, Pomeroy, Maratin, and Gebhard (1953) found that 3% to 12% of women and 10% to 20% of men reported experiencing sexual arousal in response to sadomasochistic narratives. Arndt, Foehl, and Good (1985) found that 50% of men versus roughly 33% of women reported sexual fantasies of tying up a partner. For a thorough review of the prevalence of sadism, see Steven J. Hucker's "Sexual Sadism: Psychopathology and Theory" (1997).

5. Several other studies that used phallometric assessment to examine arousal differences between sadistic and nonsadistic sexual offenders reported results similar to those obtained by Marshall, Kennedy, and colleagues (2002; cf., Langevin et al., 1985; Rice, Chaplin, Harris, & Coutts, 1994; Seto & Kuban, 1996).

6. Rosenzweig (1936) coined this term based on a scene in Lewis Carroll's *Alice's Adventures in Wonderland,* wherein the dodo bird orders a caucus race, only to later declare, "Everybody has won, and all must have prizes." Psychotherapy literature frequently employs this term to contend that factors common to all psychotherapeutic approaches produce the majority of positive treatment outcomes (cf. Luborsky et al., 2002).

References

American Educational Research Foundation. (1999). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., revised). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

Andrews, D. A., & Bonta, J. L. (2003). *The psychology of criminal conduct* (3rd ed.). Cincinnati, Ohio: Anderson.

Arndt, W., Foehl, J., & Good, F. (1985). Specific sexual fantasy themes: A multidimensional study. *Journal of Personality and Social Psychology, 48,* 472-480.

Beauvoir, S. de. (1953). *Must we burn Sade?* (A. Michelson, Trans.). London: P. Nevill.

Bloom, M., Fischer, J., & Orme, J. G. (1999). *Evaluating practice: Guidelines for the accountable professional* (3rd ed.). Boston: Allyn & Bacon.

Civil Commitment of Sexually Dangerous Individuals, N.D. Cent. Code § 25-03.3-0 et seq. (1997).

Dietz, P. E., Hazelwood, R. R., & Warren, J. (1990). The sexually sadistic criminal and his offenses. *Bulletin of the American Academy of Psychiatry and the Law, 18*(2), 163-178.

Eulenberg, A. von. (1911). *Sadism and masochism.* New York: Bell.

Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low kappa: I: The problems of two paradoxes. *Journal of Clinical Epidemiology, 43*(6), 543-549.

Finn, R. H. (1972). Effects of some variations in rating scale characteristics on the means and reliabilities of ratings. *Educational and Psychological Measurement, 32,* 255-265.

Fromm, E. (1977). *The anatomy of human destructiveness.* Markham, Ontario: Penguin.

Gratzer, T., & Bradford, J. M. W. (1995). Offender and offense characteristics of sexual sadists: A comparative study. *Journal of Forensic Sciences, 40*(3), 450-455.

Hucker, S. J. (1997). Sexual sadism: Psychopathology and theory. In D. R. Laws & W. O'Donohue (Eds.), *Sexual deviance: Theory, assessment, and treatment* (pp. 194-209). New York: Guilford.

Jackson, R. L., Richards, H. J., McCraw, J. E., & Koenen, M. (2006, March). *Sexual sadism: In word and deed*. Paper presented at the conference of the American Psychology-Law Society, St. Petersburg, FL.

Jackson, R. L., Rogers, R., & Shuman, D. W. (2004). The adequacy and accuracy of sexually violent predator evaluations: Contextualized risk assessment in clinical practice. *International Journal of Forensic Mental Health, 3*(2), 15-29.

Janus, E. S. (2000). Sexual predator commitment laws: Lessons for law and the behavioral sciences. *Behavioral Sciences and the Law, 18,* 5-21.

Karpman, B. (1954). *The sexual offender and his offenses: Etiology, pathology, psychodynamics and treatment*. New York: Julian.

Kinsey, A. C., Pomeroy, W. B., Maratin, C. E., & Gebhard, P. H. (1953). *Sexual behavior in the human female.* Philadelphia: Saunders.

Krafft-Ebing, R. von. (1965). *Psychopathia sexualis* [Psychopathy of sex]. New York: Stein & Day. (Original work published 1886)

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

Langevin, R., Bain, J., Ben-Aron, M. H., Coutlhard, R., Day, D., Handy, L., et al. (1985). Sexual aggression: Constructing a predictive equation: A controlled pilot study. In R. Langevin (Ed.), *Erotic preference, gender identity, and aggression in men: New research studies* (pp. 39-76). Hillsdale, NJ: Lawrence Erlbaum.

Levenson, J. S. (2004). Reliability of sexually violent predator civil commitment criteria. *Law and Human Behavior, 28,* 357-368.

Levenson, J. S., & Morin, J. W. (2006). Factors predicting selection of sexually violent predators for civil commitment. *International Journal of Offender Therapy and Comparative Criminology, 50*(6), 609-629.

Lever, M. (1993). *Sade: A biography* (A. Goldhammer, Trans.). New York: Farrar.

Lissitz, R. W., & Green, S. B. (1975). Effect of the number of scale points on reliability: A Monte Carlo approach. *Journal of Applied Psychology, 60,* 10-13.

Luborsky, K., Rosenthal, R., Diguer, L., Andrusyna, T. P., Berman, J. S., Levitt, J. T., et al. (2002). The Dodo Bird Verdict is alive and well—mostly. *Clinical Psychology: Science and Practice, 9*(1), 2-12.

MacCulloch, M., Snowden, P., Wood, P., & Mills, H. (1983). Sadistic fantasy, sadistic behavior, and offending. *British Journal of Psychiatry, 143,* 20-29.

Marshall, W. L. (2006). Diagnostic problems with sexual offenders. In W. L. Marshall, Y. M. Fernandez, L. E. Marshall, & G. A. Serran (Eds.), *Sexual offender treatment: Controversial issues* (pp. 33-43). Chichester, UK: Wiley.

Marshall, W. L., & Hucker, S. J. (2006a). Issues in the diagnosis of sexual sadism. *Sexual Offender Treatment, 1*(2). Available from http://www.sexual-offender-treatment.org**[TS; 2 hyphens in URL]**

Marshall, W. L., & Hucker, S. J. (2006b). Severe sexual sadism: Its features and treatment. In R. D. McAnulty and M. M. Burnette (Eds.), *Sex and sexuality: Sexual deviation and sexual offenses* (Vol. 3, pp. 227-250). Westport, CT: Praeger Perspectives.

Marshall, W. L., & Kennedy, P. (2003). Sexual sadism in sexual offenders: An elusive diagnosis. *Aggression and Violent Behavior, 8*(1), 1-22.

Marshall, W. L., Kennedy, P., & Yates, P. (2002). Issues concerning the reliability and validity of the diagnosis of sexual sadism applied in prison settings. *Sexual Abuse: A Journal of Research & Treatment, 14*(4), 301-311.

Marshall, W. L., Kennedy, P., Yates, P., & Serran, G. (2002). Diagnosing sexual sadism in sexual offenders: Reliability across diagnosticians. *International Journal of Offender Therapy and Comparative Criminology, 46*(6), 668-677.

McDermott, P. A. (1988). Agreement among diagnosticians or observers: Its importance and determination. *Professional School Psychology, 3*(4), 225-240.

McMordie, W. R. (1979). Improving measurement of death anxiety. *Psychological Reports, 44,* 975-980.

Murphy, K. R., & Davidshofer, C. O. (1998). *Psychological testing principles and applications* (4th ed.). Hillsdale, NJ: Prentice Hall.

Nelson-Gray, R. O. (1991). DSM-IV: Empirical guidelines from psychometrics. *Journal of Abnormal Psychology, 100,* 308-315.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

O'Donohue, W. T., & Geer, J. H. (1993). Research issues in sexual dysfunctions. In W. O'Donohue & J. H. Geer (Eds.), *Handbook of sexual dysfunctions* (pp. 1-14). Boston: Allyn & Bacon.

O'Donohue, W. T., Regev, L. G., & Hagstrom, A. (2000). Problems with the DSM-IV diagnosis of pedophilia. *Sexual Abuse: A Journal of Research and Treatment, 12,* 95-105.

Packard, R. L., & Levenson, J. S. (2006). Revisiting the reliability of diagnostic decisions in sex offender civil commitment. *Sexual Offender Treatment, 1*(3). Available from http://www.sexual-offender-treatment.org

Proulx, J., Blais, E., & Beauregard, É. (2006). In W. L. Marshall, Y. M. Fernandez, L. E. Marshall, & G. A. Serran (Eds.), *Sexual offender treatment: Controversial issues* (pp. 61-77). Chichester, UK: Wiley.

Raaijmakers, Q. A. W., van Hoof, A., Hart, H., Verbogt, T. F., & Vollebergh, W. A. M. (2000). Adolescents' midpoint responses on Likert-type scale items: Neutral or missing values? *International Journal of Public Opinion Research, 12*(2), 208-216.

Rice, M. E., Chaplin, T. C., Harris, G. T., & Coutts, J. (1994). Empathy for the victim and sexual arousal among rapists and nonrapists. *Journal of Interpersonal Violence, 8,* 435-449.

Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry, 6,* 412-415.

Schrenck-Notzing, A. von. (1895). *Therapeutic suggestion: Psychopathia sexualis, pathological manifestations of the sexual sense, with especial reference to contrary sexual instinct.* Philadelphia: Davis.

Seto, M. C., & Kuban, M. (1996). Criterion-related validity of a phallometric test for paraphilic rape and sadism. *Behaviour Research and Therapy, 34,* 175-183.

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy, 85*(3), 257-268.

Stekel, W. (1929). *Sadism and masochism: The psychology of hatred and cruelty* (L. Brink, Trans.). New York: Horace Liveright.

Thorndyke, E. L. (1911). *Animal intelligence.* New York: MacMillan.

Uebersax, J. S. (1987). Diversity of decision-making models and the measurement of interrater agreement. *Psychological Bulletin, 101*(1), 140-146.

World Health Organization. (1992) *The ICD-10 classification of mental and behavioral disorders*. Geneva: Author.