Adding Comments and Notes to Your PDF

To facilitate electronic transmittal of corrections, we encourage authors to utilize the comment/annotations features in Adobe Acrobat. The PDF provided has been *comment enabled,* which allows you to utilize the comment and annotation features even if using only the free Adobe Acrobat reader (see note below regarding acceptable versions). Adobe Acrobat's Help menu provides additional details on the tools. When you open your PDF, the annotation tools are clearly shown on the tool bar (although icons may differ slightly among versions from what is shown below).

Using the Javascript Window to save/publish and complete/finalize your corrections.

1. Use the "Publish" button on the "Javascript Window" to save any comments to the proof. You will be able to publish (or save) annotations, exit a proof, and return to make additional annotations at a later time
2. Upon completion of your review, in order to submit your annotations, select "Finalize (I'm Done)." When this action is taken, no additional corrections are accepted, though the window will remain open. After finalizing, your annotations will be delivered to our production system. No upload of the proof is needed.
3. You also have the ability to save the PDF proof to your local computer. The JavaScript window travels with the PDF and those buttons should be used when publishing and ultimately finalizing the article, sending it back along with all annotations into production.

For purposes of correcting the PDF proof of your journal article, the important features to know are the following:

- To **insert text,** place your cursor at a point in text and select the Insert Text tool () from the menu bar. Type your additional text in the pop-up box.
- To **replace text,** highlight the text to be changed, select the Replace Text tool () from the menu bar, and type the new text in the pop-up box. Do this instead of deleting and then reinserting.
- To **delete text,** highlight the text to be deleted and press the Delete button on the keyboard.
- Use the **Sticky Note tool** () to describe changes that need to be made (e.g., changes in bold, italics, or capitalization use; altering or replacing a figure) or to answer a question or approve a change from the editor. To use this feature, click on the Sticky Note tool in the menu bar and then click on a point in the PDF where you would like to make a comment. Then type your comment in the pop-up box.

- Use the **Callout tool** () to point directly to changes that need to be made. Try to put the callout box in an area of white space so that you do not obscure the text.
- Use the **Highlight and Add Note to Text tool** () to indicate font problems, bad breaks, and other textual inconsistencies. Select text to be changed, choose this tool, and type your comment in the pop-up box. One note can describe many changes.

- To view a list of changes to the proof or to see a more comprehensive set of annotation tools, select **Comment** from the menu bar.

As with hand-annotated proof corrections, the important points are to communicate changes clearly and thoroughly, to answer all queries and questions, and to provide complete information to allow us to make the necessary changes to your article so it is ready for publication. Do not use tools that incorporate changes to the text in such a way that no indication of a change is visible. Such changes will not be incorporated into the final proof.

To utilize the comments features on this PDF you will need Adobe Reader version 7 or higher. This program is freely available and can be downloaded from http://get.adobe.com/reader/

Subscriptions and Special Offers

In addition to purchasing reprints of their articles, authors may purchase an annual subscription, purchase an individual issue of the journal (at a reduced rate), or request an individual issue at no cost under special "hardship" circumstances. To place your order online, visit http://www.apa.org/pubs/journals/subscriptions.aspx; or you may fill out the order form below (including the mailing label) and send the completed form and your check or credit card information to the address listed on the order form.

For information about becoming a member of the American Psychological Association, visit http://www.apa.org/membership/index.aspx; or call the Membership Office at 1-800-374-2721.

2018 APA Journal Subscription Rates

Journal*	Non-Agent Individual Rate	APA Member Rate	Journal*	Non-Agent Individual Rate	APA Member Rate
American Psychologist	$ 458	$ 12	Jrnl of Family Psychology	$ 263	$ 127
Behavioral Neuroscience	$ 473	$ 194	Jrnl of Personality & Social Psychology	$ 806	$ 288
Developmental Psychology	$ 577	$ 205	Neuropsychology	$ 263	$ 127
Emotion	$ 198	$ 127	Professional Psych: Research & Practice	$ 216	$ 68
Experimental & Clinical Psychopharmacology	$ 145	$ 100	Psychological Assessment	$ 318	$ 169
History of Psychology	$ 146	$ 79	Psychological Bulletin	$ 318	$ 169
Jrnl of Abnormal Psychology	$ 273	$ 127	Psychological Methods	$ 159	$ 68
Jrnl of Applied Psychology	$ 545	$ 169	Psychological Review	$ 263	$ 109
Jrnl of Comparative Psychology	$ 159	$ 68	Psychology & Aging	$ 263	$ 127
Jrnl of Consulting & Clinical Psychology	$ 394	$ 169	Psychology of Addictive Behaviors	$ 263	$ 127
Jrnl of Counseling Psychology	$ 216	$ 95	Psychology, Public Policy, and Law	$ 159	$ 68
Jrnl of Educational Psychology	$ 263	$ 127	Rehabilitation Psychology	$ 159	$ 68
JEP: Animal Learning and Cognition	$ 159	$ 68	Clinician's Research Digest – Adult Populations	$ 159	$ 68
JEP: Applied	$ 159	$ 68			
JEP: General	$ 432	$ 169	Clinician's Research Digest – Child and Adolescent Populations	$ 159	$ 68
JEP: Human Perception, and Performance	$ 545	$ 205			
JEP: Learning, Memory, and Cognition	$ 545	$ 205			

*For journal descriptions, see APA's website: http://www.apa.org/pubs/journals

---

Instructions: Check the appropriate box, enter journal title and price information, and complete the mailing label in the right column. Enclose a check made out to the **American Psychological Association**, and mail it with the form to the APA Order Department or complete the credit card information below. Orders can also be placed online at http://www.apa.org/pubs/journals/subscriptions.aspx.

☐ Annual Subscription (available on January-December basis only). To subscribe, specify calendar year of the subscription. Refer to the Subscription Rates shown above.

Journal: _____

Calendar year of subscription: _____ Price: _____

Special Offers! If you are a journal article author, you may take advantage of two Special Offers.

☐ **Individual Copy.** You may order individual copies of the entire issue in which your article appears. As an author, you receive a special reduced rate of $5 per copy for up to a maximum of 25 copies. No phone requests accepted.

Journal: _____

Vol. no.: _____ Issue no.: _____ Issue month: _____

____ copies @ $5 a copy = $_____ (order amount)
+_____ (handling; see below)
TOTAL enclosed: $_____

Shipping & Handling Fees			
Order amount:	U.S. & Puerto Rico	Guaranteed Non-U.S.*	Economy Non-U.S.**
Up to $14.99	$5.00	$50.00	$15.00
$15 to $59.99	$6.00	$75.00	$16.00
$60.00+	10% Order Total	$125.00	$20.00

*International rates for guaranteed service are estimates.

**I agree that international economy service is non-guaranteed and does not provide tracking or date/time specific delivery. Delivery time for this level of service can take up to 8 weeks. If this level of service is selected, APA will not be held liable and will not honor any claims for undelivered, delayed, or lost shipments.

☐ **Hardship Request.** If you do not have a personal subscription to the journal and you do not have access to an institutional or departmental subscription copy, you may obtain a single copy of the issue in which your article appears at no cost by filing out the information below.

Journal: _____

Vol. no. : _____ Issue no.: _____

Issue month: _____

CREDIT CARD PAYMENT

___ VISA ___ MASTERCARD ___ AMERICAN EXPRESS

CARD NUMBER_____

Expire Date _____ Signature _____

PRINT CLEARLY – THIS IS YOUR MAILING LABEL

SHIP TO: Phone No. _____

Name _____

Address _____

_____

City _____ State _____ Zip _____

Expedited Service (enter service required): _____

Send the completed form and your check, made out to the **American Psychological Association**, or your credit card information to:

APA Order Department
750 First Street, NE
Washington, DC 20002-4242

All orders must be prepaid. Allow 4-6 weeks after the journal is published for delivery of a single copy or the first copy of a subscription.

tapraid5/z1t-assess/z1t-assess/z1t99918/z1t3414d18z | xppws | S=1 | 9/4/18 | 5:18 | Art: 2018-2088

American Psychological Association

Psychological Assessment

1040-3590/18/$12.00

http://dx.doi.org/10.1037/pas0000653

BRIEF REPORT

Sadism Among Sexual Homicide Offenders: Validation of the Sexual Sadism Scale

AQ: au

Ewa B. Stefanska
University of Roehampton

Joachim Nitschke
Ansbach District Hospital, Ansbach, Germany

Adam J. Carter
HM Prison and Probation Service, London, United Kingdom

Andreas Mokros
University Hospital of Psychiatry Zurich

Sexual sadism is assumed to be a crucial factor in sexual homicide. Prevalence estimates vary greatly due to differences in the definition of sexual sadism. A nationwide sample of 350 male perpetrators who had committed a sexual homicide offense against a female 14 years of age or above in England or Wales was assessed based on archival records. Sexual sadism was assessed using the Sexual Sadism Scale (SeSaS). Item response theory (IRT) analyses were conducted focusing on the 2-parameter logistic model. The single-factor structure of the SeSaS Part 1 was tested using confirmatory factor analysis. Estimates of both internal consistency and interrater agreement were satisfactory to substantial. IRT analysis showed that the Part 1 items captured moderate to severe levels of the latent construct (i.e., theta levels >0). Based on the Posterior Probability of Diagnosis index, the prevalence of the disorder was estimated at 37% in the sample. The substantial correlation between the SeSaS Part 1 total score and original clinical diagnoses of sadism confirms the criterion validity of the scale. Exertion of control and infliction of torture were among the more informative items. In sum, the results support the usefulness of the SeSaS instrument for assessing forensically relevant forms of sadism.

AQ: 1

Public Significance Statement
Sexual sadism is highly prevalent among the perpetrators of sexual homicide. In sexual homicide offenders, a checklist based on crime-scene behavior proves helpful to establish a tentative diagnosis.

Keywords: sadism, sexual homicide, prevalence, SeSaS, PPOD

Supplemental materials: http://dx.doi.org/10.1037/pas0000653.supp

Historically, sadism has been conceptualized as follows:

> Sadism is the experience of sexual pleasurable sensations (including orgasm) produced by acts of cruelty, bodily punishment afflicted on one's own 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. (von Krafft-Ebing, 1906, p. 80)

Since then, various definitions have been introduced with criteria that often disagree on the primary motivating force that drives sexual sadists (Marshall & Kennedy, 2003). Such assumptions on motivating forces include humiliation of the victim (e.g., Ressler, Burgess, & Douglas, 1988; Warren, Hazelwood, & Dietz, 1996), control of the victim (e.g., MacCulloch, Snowden, Wood, & Mills, 1983), the use of aggression (e.g., Myers, Burgess, Burgess, & Douglas, 1999), or the infliction of pain and victim's suffering (e.g., Seto & Kuban, 1996). Thus, as noted by Marshall and Kennedy (2003), the dispute does not revolve around the range of typical behaviors enacted by sadists but rather around what constitutes the key element that elicits their sexual excitement.

Definitional and diagnostic challenges have led to differing levels of agreement among professionals when assessing sexual sadism, ultimately impacting on the ability to diagnose sexual sadism reliably (Nitschke, Mokros, Osterheider, & Marshall, 2013). The prevalence of sexually sadistic behavior (not disorder)

AQ: 12 Ewa B. Stefanska, Department of Psychology, University of Roehampton; Joachim Nitschke, Forensic-Psychiatric Hospital, Ansbach District Hospital, Ansbach, Germany; Adam J. Carter, HM Prison and Probation Service, London, United Kingdom; Andreas Mokros, Department of Forensic Psychiatry, University Hospital of Psychiatry Zurich.

The research reported herein was presented at the 2nd International Convention of Psychological Science, Vienna, Austria, on March 25, 2017. We thank Dimiter M. Dimitrov, from George Mason University, for his help.

Correspondence concerning this article should be addressed to Andreas Mokros, who is now at the Department of Psychology, FernUniversität in Hagen, Postfach, 58084 Hagen, Germany. E-mail: andreas.mokros@fernuni-hagen.de

tapraid5/z1t-assess/z1t-assess/z1t99918/z1t3414d18z | xppws | S=1 | 9/4/18 | 5:18 | Art: 2018-2088

AQ: 2

in the population at large has been estimated at about 2% to 3% (Baur et al., 2016). According to current psychiatric classification, sexual sadism disorder requires not only extended duration of the condition (i.e., more than 6 months) but also one of two additional aspects: either sadistic acts against a nonconsenting individual or distress–impairment for the person afflicted (American Psychiatric Association, 2013). In samples of sexual offenders, the prevalence of sexual sadism disorder has been estimated somewhat higher, with percentages ranging up to 10% of rapists (Eher et al., 2016). Finally, among the perpetrators of sexual homicide, the prevalence of sexual sadism disorder has been reported at about one third in samples from Germany (36.7%, $N = 166$; Briken, Habermann, Berner, & Hill, 2005) and the United States (29.3%, $N = 232$; Geberth & Turco, 1997). Given the general uncertainty of clinical diagnoses in the forensic domain (Mokros, Habermeyer, & Küchenhoff, 2018) and in light of the variability of observer agreement on sexual sadism in particular (Nitschke et al., 2013), it remains unclear how high the prevalence of sexual sadism truly is among the perpetrators of sexual homicide. Although sadism was recently shown to be less relevant for offense recidivism than are customary indicators of risk (such as antisocial personality or psychopathy; Eher et al., 2016, Study 2), it might still be the case that sexual sadism is a primary force behind committing sexual offenses that are rare but most severe (i.e., sexual homicide). A meta-analysis of seven studies with a total sample comprising 2,169 individuals (Eher et al., 2016, Study 1) showed that sexual sadism was associated with a slightly higher risk of sexual recidivism (risk ratio = 1.38), yet not to a statistically significant degree ($p = .052$).

The current study was planned to assess the psychometric properties and test the applicability of an item response theory (IRT) model (two-parameter logistic model [2PLM], aka Birnbaum model) for a behavioral index of sexual sadism, the Sexual Sadism Scale (SeSaS; Mokros, Schilling, Weiss, Nitschke, & Eher, 2014). Second, the current study was meant to yield a robust estimate of the prevalence of sexual sadism among the perpetrators of sexual homicide offenses. For this purpose, a method was used that allows gauging prevalence from the minimum level of the latent trait associated with a given cutoff, the Posterior Probability of Diagnosis (PPOD) index (Lindhiem, Kolko, & Yu, 2013).

The SeSaS (Mokros et al., 2014) is a checklist of dichotomous (yes−no) items that consists of two parts: Part 1 contains 11 items that code for crime scene behavior, including aspects like the gratuitous exertion of violence or confinement of the victim. These behavioral indicators were derived empirically from a larger pool of items showing content validity according to a survey of experts in the area of sexual sadism (Marshall, Kennedy, Yates, & Serran, 2002). Part 2 of the SeSaS instrument comprises three biographical items (planful conduct, prior sadistic acts beyond listed offenses, and arousability through sadistic fantasies or acts). The composite score of the Part 1 items was shown to have excellent interrater agreement, with an intraclass correlation coefficient (ICC) of [2, 5] (i.e., average measure, absolute agreement = .91) in a sample of 20 cases assessed by five raters (Mokros et al., 2014). The Part 1 sum score showed a moderate to substantial correlation with clinical diagnoses of sadism ($r_{pc} = .55$ according to Eher et al., 2016; $r_{pc} = .46$ according to Longpré, Proulx, & Brouillette-Alarie, 2018; area under the curve = .87 according to Mauzaite, Sauter, Seewald, & Dahle, 2017). Furthermore, the Part 1 sum score was strongly correlated ($r = .66$) with the Massachusetts Treatment Center Sadism Scale (Longpré, Guay, & Knight, 2017).

For the predecessor of the SeSaS Part 1, a cutoff score of 4 points has been suggested as being indicative of sexual sadism (Nitschke, Osterheider, & Mokros, 2009). Across four samples of male (84.8%) and female (15.2%) offenders from Germany and the United States (total $N = 591$), the overall sensitivity of the cutoff regarding a diagnosis of sexual sadism according to the criteria of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; *DSM–IV–TR*; American Psychiatric Association, 2000) was estimated at 95% and the specificity at 99% (Nitschke et al., 2013). A prior study on the items now forming Part 1 of the SeSaS instrument showed good absolute model fit for a one-factorial structure in a confirmatory factor analysis (root-mean-square error of approximation [RMSEA] = .05), even though the incremental fit index (here, comparative fit index [CFI]) was below commonly accepted standards (.89; Mokros, Schilling, Eher, & Nitschke, 2012). Finally, previous analyses yielded support for scalability of the Part 1 items in terms of nonmetric item response theory (Nitschke et al., 2009) or the one-parameter logistic (aka Rasch) model (Mokros et al., 2012). Note that the SeSaS instrument was developed into a structured professional judgment instrument with more detailed item descriptions subsequently (Mokros et al., 2014).

Method

The sample used in the study comprised 350 male sexual killers who perpetrated against female victims 14 years of age[1] or above and served a custodial sentence within HM Prison Service in England and Wales. Homicides were nonserial, with the majority of offenders killing a single victim and six cases having two victims (with the maximum time frame between killing the two victims established as 3 hr). The criteria for sexual homicide included offenses where a sexual element in the killing was evidenced, suspected, or admitted. The sample represented a full data search of all cases stored electronically in the Offender Assessment System in England and Wales captured from the beginning of its existence in the early 2000s (i.e., from that date, the offender was still serving a prison sentence). The actual time frame of the index offenses committed by the perpetrators ranged from the 1950s to 2010s. Details of the offense events were collected from the Public Protection Unit Database.

Fn1

The analyses reported herein were focused on the 11 dichotomous indicator variables of the SeSaS coding for crime-scene behavior (i.e., the SeSaS Part 1). The presumed unidimensional structure of these 11 items was assessed by confirmatory factor analysis (CFA) using the program Mplus, Version 6.12, for Mac (●●●). A robust estimator that is suitable for categorical items was chosen for the CFA (i.e., weighted least squares means and variance adjusted). IRT analyses based on the two-parameter logistic model (2PLM) were conducted in Mplus, Version 6.12 for Mac, as well. The 2PLM was obtained through maximum likelihood estimation with robust standard errors.

AQ: 3

The internal consistency of the SeSaS Part 1 items was assessed at both the factor level (in terms of MacDonald's omega, ω) and

[1] The age of the victim was set at 14 to offer consistency with previous research (Carter & Hollin, 2010).

tapraid5/z1t-assess/z1t-assess/z1t99918/z1t3414d18z | xppws | S=1 | 9/4/18 | 5:18 | Art: 2018-2088

the manifest level (in terms of the 2PLM reliability estimate for dichotomous data, rho [ρ], developed by Dimitrov, 2003b). For ω, a 95% bootstrap confidence interval (bCI) was obtained based on 1,000 bootstrap draws. Rho was estimated using the program IRT-TRUE (Dimitrov, 2003a).

According to Bayes's theorem, the individual level on the latent trait being measured by the IRT model can be gauged through expected a posteriori (EAP) scores. For EAP estimates, a density distribution is obtained for the latent trait of an individual based on prior information (e.g., the individual response pattern); the expected value of said distribution is used as the person parameter of the person in question (Walter & Rost, 2011). Unlike the case in maximum likelihood estimation, EAP estimates are also available for individuals for whom none of the items or all items were coded as present (Muraki & Engelhard, 1985).

Finally, the EAP person parameter estimates derived from the 2PLM were analyzed in terms of the Posterior Probability of Diagnosis (PPOD) index (Lindhiem et al., 2013). Due to the differential weighting of items in terms of their discrimination parameter within the 2PLM, the same total score may reflect different levels of the underlying trait, depending on the combination of items coded as present in a given case. Thus, all item profiles that occurred in the sample and equaled the cutoff score of 4 points (e.g., 01110100000 or 11110000000) were identified; the minimum EAP person parameter associated with any of these profiles was determined; all individuals whose posterior probability of their EAP person parameters being equal to or above that minimum level (i.e., the PPOD index) was at least .5 (regardless of the actual sum score) were considered as tentatively diagnosed as sadists based on the SeSaS; then, the agreement between those with manifest scores ≥4 and those with a PPOD index >.5 was checked, also in terms of sensitivity and specificity. For the calculation of the PPOD index, Method A from Lindhiem et al. (2013) was used (i.e., based on a normal cumulative distribution function).

The research plan was reviewed by the National Research Committee and found to comply with ethical standards. Moreover, the authors complied with American Psychological Association ethical standards in collecting, analyzing, and interpreting the data for the current study (●●●).

AQ: 4

Results

To establish interrater agreement of the instrument, two raters who were chartered forensic psychologists with the British Psychological Society and were registered with the Health and Care Professions Council independently blind-coded 28 cases. The ICC [2, 1] (i.e., single measure, absolute agreement) on the total score for Part 1 of the SeSaS was calculated to .80 (95% CI [.58, .90]). Adopting the rules of thumb suggested by Cicchetti and Sparrow (1981; cf. Fleiss, 1981), we found the agreement on the individual items to be excellent for three items (1, 5, and 8), good for another three items (3, 6, and 7), fair for two items (4 and 9), and poor for one item (2; see Table 1 for kappa estimates). Note that kappa could not be computed for two individual items (10 and 11), due to perfect agreement (joint absence).

T1, AQ:5

The sample mean of the SeSaS Part 1 sum score was 2.67 (SD = 1.71, Mdn = 2), with values ranging from 0 to 10. The distribution of the SeSaS Part 1 total score was skewed to the right (skewness = .95; i.e., had a longer right tail), leptokurtic (kurtosis = 1.71), and unimodal (mode = 2). Out of 350 individuals, 94 (26.9%) had been assigned a SeSaS Part 1 sum score of 4 or above.

Within CFA, a single-factor solution with 22 free parameters had the following model fit properties: CFI = .87, RMSEA = .054, 90% CI [.037, .070], and $\chi^2(44, N = ●●●) = 88.70$, $p < .001$. Thus, the absolute fit index (RMSEA) was indicative of good fit (<.05), whereas the incremental fit index (CFI) was below the commonly accepted standard of .95 for good fit. In the present case, the magnitude of the CFI is likely not informative, however, because it critically hinges on the suitability of the null (or baseline) model.[2] Therefore, the 90% CI for the RMSEA coefficient (i.e., [.037, .070]) is more informative presently and indicates good model fit. The fully standardized factor loadings ranged from .29 (Item 7) to .80 (Item 3; all ps < .01, two-sided). For the latent factor based on the Part 1 items, ω was calculated at .84 (95% bCI [.80, .89]). Hence, the internal consistency of the Part 1 items was good.

AQ: 6

Fn2

AQ: 7

The 2PLM model comprises 22 free parameters. Estimating the model with the current data yielded a log-likelihood of −1,381.87. The corresponding value for the Akaike information criterion and the Bayesian information criterion were 2,807.73 and 2,892.60, respectively. Model fit was tested through bivariate item comparisons. There were three occurrences of significant item misfit (|z| > 1.96) among 220 bivariate item comparisons (11 × 10/2 = 55 nonredundant item pairs with four possible numerical codings each, namely 0/0, 0/1, 1/0, and 1/1). Given that one should expect 5% (i.e., about 11 such violations) under a Type I error rate of .05, the observed rate points toward superior goodness of fit (p = .999) in a cumulative binomial test. Moreover, there was not a single occurrence of significant misfit for the 22 univariate item fit statistics (11 items × 2 possible codings [0/1]).

Table 1 shows the item parameters (discrimination, a_i, and difficulty, b_i) for the 11 items of the SeSaS Part 1 according to the 2PLM, along with the corrected polychoric part−whole correlations. At 3.14, Item 3 (torture) had the highest discrimination parameter (a_i) estimate. Thus, Item 3 afforded the maximum of information on the latent trait (θ) of sexual sadism within the sample. The lowest a_i estimate was .43 (for Item 7, excessive violence). Consequently, excessive violence does not distinguish well among those with lower or higher levels of θ. Thus, except for Item 7, none of the items had an a_i estimate below the minimum value of .5 usually observed in 2PLM applications (Reeve & Fayers, 2005). The different gradients of the a_i estimates are reflected by the slopes of the item characteristic curves in Figure 1 (Panel a) in the online supplemental materials, with higher a_i values equaling steeper slopes. Item 3 (torture) yields the maximum information (see the online supplemental materials, Figure 1, Panel b) but differentiates only within a narrow spectrum of the latent trait θ.

[2] As Kenny (2015) pointed out, the comparative fit index (CFI) should not be computed if the root-mean-square error of approximation (RMSEA) of the null model is smaller than .158 (cf. Rigdon, 1996). For the data at hand, the RMSEA of the null model is .133. This means that the null model (without any intercorrelations) already describes the data quite well. Consequently, there is little to be gained in terms of an incremental fit index such as the CFI or the Tucker−Lewis index (Kenny, 2015).

tapraid5/z1t-assess/z1t-assess/z1t99918/z1t3414d18z | xppws | S=1 | 9/4/18 | 5:18 | Art: 2018-2088

Table 1
Corrected Polychoric Item Part−Whole Correlations (R_{Pc}), Estimates of Item Parameters According to the Two-Parameter Logistic Model, Cohen's κ Coefficients, and Percentage of Items Coded as Present

SeSaS Part 1 items	r_{pc} (*p*)	Discrimination (a_i)	Difficulty (*bi*)	Cohen's κ	%
1. Sexual arousal during the crime scene behaviors	.22 (.009)	.889	−2.458	.78	87.1
2. Exertion of power, control, or dominance	.54 (<.001)	2.134	−.408	.39	62.6
3. Torturing the victim	.65 (<.001)	3.142	1.344	.65	12.0
4. Degrading or humiliating behavior directed toward the victim	.59 (<.001)	2.249	1.385	.51	13.7
5. Mutilation of sexual areas of the victim's body	.58 (<.001)	1.175	1.911	1.0	14.0
6. Mutilation of other parts of the victim's body	.51 (<.001)	1.116	2.652	.65	7.7
7. Excessive physical violence	.19 (.004)	.434	1.813	.71	32.0
8. Insertion of objects into the victim's bodily orifices	.33 (<.001)	.701	2.939	.87	13.1
9. Ritualistic behavior	.48 (<.001)	1.348	1.553	.42	16.9
10. Confinement of the victim	.57 (<.001)	1.891	2.702	—	2.6
11. Taking trophies	.29 (<.01)	.827	3.955	—	4.9

Note. $N = 350$. Dashes indicate that data could not be computed, due to perfect agreement (joint absence). SeSaS = Sexual Sadism Scale (Mokros, Schilling, Weiss, Nitschke, and Eher (2014)).

The estimates for item difficulty (b_i) ranged from −2.46 (for Item 1, sexual arousal) to 3.96 (for Item 11, taking trophies−keeping records). Thus, most b_i estimates were in the range from −3 to 3 commonly encountered in 2PLM modeling, with the exception of one item (11). In looking at the test information function (see the online supplemental materials, Figure 1, Panel c), it becomes clear that the maximum total information is conveyed at $\theta = 1.39$. Thus, within the sample analyzed, the SeSaS items conveyed the most information at an elevated trait level, which is similar (in terms of difficulty) to Items 3 (torture), 4 (degradation−humiliation), and 9 (ritualistic behavior), with b_i estimates of 1.34, 1.39, and 1.55, respectively.

In looking at the test characteristic curve (see the online supplemental materials, Figure 1, Panel d), it becomes clear that the association between the latent trait, θ, and the expected score is most reliable at medium to high trait levels (i.e., for $\theta > 0$). This is concomitant with the focus on the severe (or forensically relevant) variant of sexual sadism. A global 2PLM estimate of scale reliability (ρ) was estimated at .76, somewhat lower than ω (.84).

A kernel density estimate for the distribution of the EAP person parameter estimates is provided as Figure 2 in the online supplemental materials. The distribution is bimodal, with a local maximum at approximately −1 and a global maximum at 0. For each possible manifest score on the SeSaS Part 1 (s_k), we checked the minimum and maximum θ levels associated with profiles affording the total score in question within the sample (i.e., $min[\hat{\theta} \mid s_k]$ and $max[\hat{\theta} \mid s_k]$ for $k = 1.12$; see Table 2 in the online supplemental materials). For a SeSaS Part 1 total score of 1, for instance, the minimum θ level associated with this manifest score was estimated at −1.10, whereas $max(\hat{\theta} \mid 1)$ was estimated at −.37. For the cutoff score of 4 points recommended for the SeSaS Part 1, the minimum estimate was .14.

Next, we calculated the PPOD index, that is, the posterior probability of having a θ level of .14 or higher with any pattern of items coded as present within the sample. When Items 1, 2, and 8 were coded as present, θ was estimated at .178, for example. Thus, a case with the item profile of 11000001000 surpassed the latent trait level minimally implied by the manifest cutoff score of 4. Finally, all individuals whose PPOD index (i.e., the posterior probability of their $\hat{\theta}_j$ being at least .14, given their item profile) was .5 or higher were assigned to the PPOD ≥ .5 group.

The cell entries for a 2 × 2 contingency table (SeSaS Part 1 total score ≥4/PPOD ≥ .5: no−no, no−yes, yes−no, and yes−yes) were 218, 38, 1, and 93, respectively. In other words, only a single individual would not be considered sadistic based on the PPOD index but rather regarded as sadistic based on the SeSaS Part 1 sum score. Conversely, 38 individuals had PPOD index values indicative of sadism despite SeSaS Part 1 sum scores of <4. Nevertheless, the agreement between the two modes of assessment was high ($z = 10.78$, $p < .001$).

Based on the PPOD index >.5 criterion, sensitivity of the SeSaS Part 1 cutoff score (4 points) was estimated at 71.0% (95% CI [62.4, 78.6]), whereas the specificity was estimated at 99.5% (95% CI [97.5, 100]). Consequently, the cutoff score of 4 points is relatively conservative, maximizing specificity rather than sensitivity.

Furthermore, based on the PPOD index >.5 criterion, the prevalence of sexual sadism was estimated at 37.4% (95% CI [32.0, 42.7]). Hence, at least one third of sexual homicide offenders can be expected to be sexual sadists. Finally, the correlation between the SeSaS Part 1 total score and a dichotomous variable coding for whether the offender at hand had been diagnosed as a sadist according to his files at some point was $r_{pb} = .57$ ($p < .001$). Despite the variability of the methods and criteria used by clinicians and expert witnesses to reach such a diagnosis, the strength of the association attests to the criterion validity of the SeSaS for sexual homicide offenders.

Discussion

The current study assessed sexual sadism in a nationwide sample of men who had committed sexual homicide offenses in England and Wales. Using CFA, we corroborated the factorial structure of a file-based assessment instrument for sexual sadism, the SeSaS. By focusing on the items that code for crime-scene behavior (i.e., Part 1 of the SeSaS), both interrater agreement and internal consistency could be ascertained. Moreover, the corresponding sum score was shown to be associated with clinical diagnoses of sadism derived from the files.

Within the framework of IRT, the SeSaS Part 1 items were concomitant with the 2PLM. That is, the association between the latent trait of sexual sadism and the occurrence of behavioral

tapraid5/z1t-assess/z1t-assess/z1t99918/z1t3414d18z | xppws | S=1 | 9/4/18 | 5:18 | Art: 2018-2088

indicators could be described by logistic functions with two parameters (discrimination and difficulty). The results of the 2PLM modeling imply that the SeSaS captures moderate to severe levels of the latent trait of sexual sadism. This extends earlier research on selective (Nitschke et al., 2009) or smaller (Mokros et al., 2012) samples testing nonmetric IRT or Rasch models for the predecessor of the SeSaS Part 1, respectively.

More specifically, the 2PLM model opens up the possibility of assessing the uncertainty at the latent trait level that is associated with recommended cutoffs at the observed level. In the case of the SeSaS, the recommended cutoff score of 4 points represents a conservative threshold, compared with the so-called Posterior Probability of Diagnosis (PPOD) index. Moreover, the prevalence of sexual sadism was estimated based on the PPOD index. According to the 95% CI of the prevalence estimate, at least one third of sexual homicide offenders are sexual sadists. This estimate accords well with earlier findings at the manifest level (Briken et al., 2005; Geberth & Turco, 1997). In applying the SeSaS Part 1 items it should be noted, however, that the diagnostic usefulness of Items 11 and 8 may be limited due to high estimates of the difficulty parameters. Put differently, these items concern only a minor fraction of individuals. Similar reservations apply to Item 2 but for another reason (i.e., suboptimal interrater agreement).

Recently, Eher et al. (2016) showed that the *DSM–IV–TR* diagnosis of sexual sadism was only moderately related to violent or sexual offense recidivism. In addition, Eher and colleagues demonstrated that the diagnosis of sexual sadism does not add incremental validity for assessing the risk of reoffending once customary risk factors like antisocial personality or psychopathy have been controlled for. The current results, however, show that sexual sadism is a relevant condition in the most grievous (i.e., lethal) forms of sexual aggression. Therefore, it might turn out in further studies that the SeSaS, although not predictive for general violent reoffense (Eher et al., 2016), might contribute to predicting at least the most grievous forms of (sexual) aggression. Berner, Hill, and Briken (2018) emphasized the importance of diagnosing sexual sadism reliably regarding treatment planning. We might add that delineating the diagnosis based on behavioral indicators (cf. Kingston & Yates, 2008) is particularly important with individuals who are likely motivated to deny or downplay sadistic urges or fantasies, such as sexual homicide offenders. Furthermore, an operational definition of sexually sadistic conduct against nonconsenting individuals (as provided by the SeSaS) could further the understanding of the commonalities and differences with sadomasochistic role play.

AQ: 8

The current study dealt with male nonserial offenders who perpetrated against female victims 14 years old and over. This selection criteria allowed us to (a) focus on the most prevalent group of offenders in the correctional facilities that will most likely require the SeSaS assessment (Proulx, Cusson, & Beauregard, 2007), (b) examine a complex group given the lack of victim statements available and limited evidence behavioral patterns due to the nonserial nature of the offense, and (c) investigate a group more likely eligible for parole than are serial homicide offenders. However, such sample restrictions also create limitations, because the results are applicable to only the type of offenders included. Thus, the present research does not generalize to the rare group of serial sexual homicide offenders for whom the prevalence of sexual sadism is presumably even higher (Warren et al., 1996). Moreover, testing the criterion validity of instruments like the SeSaS with clinical diagnoses of sadism is a somewhat suboptimal strategy given the concerns about the reliability of clinician judgments for this diagnosis (Nitschke et al., 2013). Therefore, physiological measurement may provide further evidence of criterion validity (see, e.g., Seto, Lalumière, Harris, & Chivers, 2012, for a useful stimulus set) even though extant results using phallometry yielded nil correlations (Longpré et al., 2018).

In sum, the analyses presented herein confirm the appropriateness of the SeSaS as an assessment instrument for forensically relevant sexual sadism in English and Welsh offenders, extending its validity beyond the U.S., Canadian, and German samples scrutinized so far.

References

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

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Baur, E., Forsman, M., Santtila, P., Johansson, A., Sandnabba, K., & Långström, N. (2016). Paraphilic sexual interests and sexually coercive behavior: A population-based twin study. *Archives of Sexual Behavior, 45,* 1163–1172. http://dx.doi.org/10.1007/s10508-015-0674-2

Berner, W., Hill, A., & Briken, P. (2018). Sexual sadism: Pharmacological and psychotherapeutic approaches. In J. Proulx, E. Beauregard, A. J. Carter, A. Mokros, R. Darjee, & J. James (Eds.), *Routledge international handbook of sexual homicide studies* (pp. 317–325). London, United Kingdom: Routledge.

Briken, P., Habermann, N., Berner, W., & Hill, A. (2005). The influence of brain abnormalities on psychosocial development, criminal history and paraphilias in sexual murderers. *Journal of Forensic Sciences, 50*(5), JFS2004472. http://dx.doi.org/10.1520/JFS2004472

Carter, A. J., & Hollin, C. R. (2010). Characteristics of non-serial sexual homicide offenders: A review. *Psychology, Crime & Law, 16,* 25–45. http://dx.doi.org/10.1080/10683160802621933

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency, 86,* 127–137.

Dimitrov, D. M. (2003a). IRT-TRUE: A computer program for estimation of true-score parameters and reliability from IRT estimates of item parameters [Computer software]. Unpublished manuscript, Graduate School of Education, George Mason University, Fairfax, VA.

AQ: 9

Dimitrov, D. M. (2003b). Marginal true-score measures and reliability for binary items as a function of their IRT parameters. *Applied Psychological Measurement, 27,* 440–458. http://dx.doi.org/10.1177/0146621603258786

Eher, R., Schilling, F., Hansmann, B., Pumberger, T., Nitschke, J., Habermeyer, E., & Mokros, A. (2016). Sadism and violent reoffending in sexual offenders. *Sexual Abuse, 28,* 46–72. http://dx.doi.org/10.1177/1079063214566715

AQ: 10

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York, NY: Wiley.

Geberth, V. J., & Turco, R. N. (1997). Antisocial personality disorder, sexual sadism, malignant narcissism, and serial murder. *Journal of Forensic Sciences, 42,* 49–60. http://dx.doi.org/10.1520/JFS14067J

Kenny, D. A. (2015). *Measuring model fit*. Retrieved from http://davidakenny.net/cm/fit.htm

Kingston, D. A., & Yates, P. M. (2008). Sexual sadism: Assessment and

tapraid5/z1t-assess/z1t-assess/z1t99918/z1t3414d18z | xppws | S=1 | 9/4/18 | 5:18 | Art: 2018-2088

treatment. In D. R. Laws & W. O'Donohue (Eds.), *Sexual deviance: Theory, assessment, and treatment* (2nd ed., pp. 231–249). New York, NY: Guilford Press.

Lindhiem, O., Kolko, D. J., & Yu, L. (2013). Quantifying diagnostic uncertainty using item response theory: The Posterior Probability of Diagnosis index. *Psychological Assessment, 25,* 456–466. http://dx.doi.org/10.1037/a0031392

AQ: 11 Longpré, N., Guay, J.-P., & Knight, R. A. (2017). MTC Sadism Scale: Toward a dimensional assessment of severe sexual sadism with behavioral markers. *Assessment*. Advance online publication. http://dx.doi.org/10.1177/1073191117737377

Longpré, N., Proulx, J., & Brouillette-Alarie, S. (2018). Convergent validity of three measures of sexual sadism: Value of a dimensional measure. *Sexual Abuse, 30,* 192–208. http://dx.doi.org/10.1177/1079063216649592

MacCulloch, M. J., Snowden, P. R., Wood, P. J., & Mills, H. E. (1983). Sadistic fantasy, sadistic behaviour and offending. *British Journal of Psychiatry, 143,* 20–29. http://dx.doi.org/10.1192/bjp.143.1.20

Marshall, W. L., & Kennedy, P. (2003). Sexual sadism in sexual offenders. An elusive diagnosis. *Aggression and Violent Behavior, 8,* 1–22. http://dx.doi.org/10.1016/S1359-1789(01)00052-0

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,* 668–677. http://dx.doi.org/10.1177/0306624X02238161

Mauzaite, A., Sauter, J., Seewald, K., & Dahle, K. P. (2017, September). *Tatbildbasierte Screening-Instrumente für die Diagnosen Pädophilie und Sexuellen Sadismus in der Gruppe der Hochrisikotäter* [Crime-scene based screening instruments for the diagnoses of pedophilia and sexual sadism within the group of high-risk offenders]. Paper presented at the 17th Meeting of the Chapter on Legal Psychology of the German Psychological Society, Jena, Germany.

Mokros, A., Habermeyer, E., & Küchenhoff, H. (2018). The uncertainty of psychological and psychiatric diagnoses. *Psychological Assessment, 30,* 556–560. http://dx.doi.org/10.1037/pas0000524

Mokros, A., Schilling, F., Eher, R., & Nitschke, J. (2012). The Severe Sexual Sadism Scale: Cross-validation and scale properties. *Psychological Assessment, 24,* 764–769. http://dx.doi.org/10.1037/a0026419

Mokros, A., Schilling, F., Weiss, K., Nitschke, J., & Eher, R. (2014). Sadism in sexual offenders: Evidence for dimensionality. *Psychological Assessment, 26,* 138–147. http://dx.doi.org/10.1037/a0034861

Muraki, E., & Engelhard, G., Jr. (1985). Full-information item factor analysis: Applications of EAP scores. *Applied Psychological Measurement, 9,* 417–430. http://dx.doi.org/10.1177/014662168500900411

Myers, W. C., Burgess, A. W., Burgess, A. G., & Douglas, J. E. (1999). Serial murder and sexual homicide. In V. van Hasselt & M. Herson (Eds.), *Handbook of psychological approaches with violent offenders* (pp. 153–172). http://dx.doi.org/10.1007/978-1-4615-4845-4_9

Nitschke, J., Mokros, A., Osterheider, M., & Marshall, W. L. (2013). Sexual sadism: Current diagnostic vagueness and the benefit of behavioral definitions. *International Journal of Offender Therapy and Comparative Criminology, 57,* 1441–1453. http://dx.doi.org/10.1177/0306624X12465923

Nitschke, J., Osterheider, M., & Mokros, A. (2009). A cumulative scale of severe sexual sadism. *Sexual Abuse, 21,* 262–278. http://dx.doi.org/10.1177/1079063209342074

Proulx, J., Cusson, M., & Beauregard, E. (2007). Sexual murder: Definitions, epidemiology and theories. In J. Proulx, E. Beauregard, M. Cusson, & A. Nicole (Eds.), *Sexual murderers: A comparative analysis and new perspectives* (pp. 9–28). Chichester, United Kingdom: Wiley.

Reeve, B. B., & Fayers, P. (2005). Applying item response theory modeling for evaluating questionnaire item and scale properties. In P. Fayers & R. Hays (Eds.), *Assessing quality of life in clinical trials: Methods and practice* (2nd ed., pp. 55–73). Oxford, United Kingdom: Oxford University Press.

Ressler, R. K., Burgess, A. W., & Douglas, J. E. (1988). *Sexual homicide: Patterns and motives*. New York, NY: Lexington.

Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling, 3,* 369–379. http://dx.doi.org/10.1080/10705519609540052

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. http://dx.doi.org/10.1016/0005-7967(95)00056-9

Seto, M. C., Lalumière, M. L., Harris, G. T., & Chivers, M. L. (2012). The sexual responses of sexual sadists. *Journal of Abnormal Psychology, 121,* 739–753. http://dx.doi.org/10.1037/a0028714

von Krafft-Ebing, R. (1906). *Psychopathia sexualis, with especial reference to the antipathic sexual instinct: A medico-forensic study* (12th ed.). Retrieved from https://archive.org/details/psychopathiasexu00krafuoft

Walter, O., & Rost, J. (2011). Psychometrische Grundlagen von Large Scale Assessments [Psychometric foundations of large-scale assessments]. In L. F. Hornke, M. Amelang, & M. Kersting (Eds.), *Methoden der psychologischen Diagnostik* [Methods of psychological diagnostics] (pp. 88–150). Göttingen, Germany: Hogrefe.

Warren, J. I., Hazelwood, R. R., & Dietz, P. E. (1996). The sexually sadistic serial killer. *Journal of Forensic Sciences, 41,* 970–974. http://dx.doi.org/10.1520/JFS14033J

Received January 8, 2018
Revision received July 24, 2018
Accepted August 1, 2018 ■

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES 1

AQau—Please confirm the given-names and surnames are identified properly by the colors. = Given-Name, = Surname
The colors are for proofing purposes only. The colors will not appear online or in print.

AQ1—Author: In the "Sexual sadism was assessed" sentence and elsewhere in the abstract, citations were omitted per abstract guidelines.

AQ2—Author: In the "According to current" sentence, the citation seemed to be referring to the fifth edition of the *DSM,* so that reference was added to the reference list. Is this correct?

AQ3—Author: In the "The presumed unidimensional structure" sentence, please provide a regular reference for Mplus, following the format of Example 56 on page 211 of the sixth edition of the manual if there are no authors.

AQ4—Author: In the "Moreover, the authors" sentence, please provide the reference where the bullets appear.

AQ5—Author: In Table 1: (1) Because tables must stand on their own, apart from the text, please provide a more complete title (self-explanatory column headings don't need to be listed in the title, and if necessary, others can be explained in the table note), if possible keeping it to 15 words or fewer, and if abbreviations are used, define them in a table note. (2) "two-sided" was omitted for the *p* values because it seemed to mean the same as "two-tailed," which is assumed, if one-tailed is not also mentioned.

AQ6—Author: In the "Within CFA" sentence, please insert the *N* value for the chi-square statistic.

AQ7—Author: In the "The fully standardized" sentence, again, if "two-sided" means the same as "two-tailed," it can be deleted, because we assume two-tailed unless one-tailed is also discussed.

AQ8—Author: In the "Furthermore, an operational definition" sentence and elsewhere, instances of the editorial "we" were rephrased. Here, "our" was changed to "the."

AQ9—Author: Dimitrov, D. M. (2003a). Reference was amended to indicate this is unpublished computer software. Is this correct?

AQ10—Author: Eher, R., Schilling, F., Hansmann, B., Pumberger, T., Nitschke, J., Habermeyer, E., & Mokros, A. (2016). Title was amended per online article.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

AQ11—Author: Longpré, N., Guay, J.-P., & Knight, R. A. (2017). Is there any update?

AQ12—Author: In the author notes: (1) Because the author notes stated that Dr. Mokros "is now at the" Department of Psychology, FernUniversität in Hagen, this affiliation was omitted from the byline and first paragraph, which should show only the affiliations during conduct of the study. Is this correct? (2) If any other author has changed affiliation since the article was written, please provide the new department and institution in a new second paragraph in the format "[author's full name as in the byline] is now at the [department and institution]," and if it is nonacademic or no affiliation, provide the city and country. (Note that Dr. Mokros's change is explained in the correspondence paragraph.) (3) Is there any funding information to add? (4) If Dr. Dimitrov was involved in reviewing the article for acceptance for publication, journal policy prohibits thanking him in the author notes. He could be thanked for particular contributions in footnotes in appropriate locations in text if you like. Just provide the footnote text and indicate where to footnote it. (5) "Postfach" was added to the correspondence address from the department's website. Is this correct?