Revista Brasileira de
Sexualidade Humana ISSN 2675-1194
DOI: https://doi.org/10.35919/rbsh.v37.1383

SISTEMÁTICA, INTEGRATIVA E DE ESCOPO

BETWEEN INTEREST AND DISORDER: WHAT DO INSTRUMENTS ACTUALLY MEASURE IN THE ASSESSMENT OF PARAPHILIAS? A SCOPING REVIEW

ENTRE INTERESSE E TRANSTORNO: O QUE OS INSTRUMENTOS REALMENTE MEDEM NA AVALIAÇÃO DE PARAFILIAS? UMA REVISÃO DE ESCOPO

ENTRE INTERÉS Y TRASTORNO: ¿QUÉ MIDEN REALMENTE LOS INSTRUMENTOS EN LA EVALUACIÓN DE LAS PARAFILIAS? UNA REVISIÓN DE ALCANCE

Felício de Freitas Netto[1], Larissa Boiko[2], Tatiana Menezes Garcia Cordeiro[3], Ricardo Zanetti Gomes[4]

Abstract: Introduction: The distinction between paraphilia and paraphilic disorder, consolidated in the DSM-5, requires careful assessment to avoid both undue pathologization and clinical neglect. However, available instruments differ in terms of construct, context, and psychometric robustness. Objective: To map and characterize standardized instruments used in the assessment of paraphilias and paraphilic disorders in adults. Methods: A scoping review was conducted, guided by the PCC framework, with searches performed in the PubMed/MEDLINE and BVS platforms, without time restriction and including studies in Portuguese, English, and Spanish. Studies on the development, cross-cultural adaptation, or psychometric validation of standardized instruments focused on measuring paraphilias or paraphilic disorders in adults were eligible. Exclusively psychophysiological or indirect measures were excluded. Study selection was performed in two stages by independent reviewers, with descriptive synthesis organized by individual instrument. Results: Of the 1,696 records identified, 10 studies were included, all retrieved from PubMed/MEDLINE. Two main groups of instruments were identified: self-report measures focused on the dimensional assessment of interests and traits; and record-based scales, predominantly applied in forensic contexts. The most frequently reported psychometric evidence included internal structure analysis, measurement invariance, and criterion validity, with variability in the reporting of reliability and clinical applicability. No eligible instruments indexed in BVS were identified. Conclusion: The assessment of paraphilias is organized in a complementary model involving self-report and record-based scales, with persistent gaps in standardization and in the availability of instruments adapted to the Brazilian context.

Keywords: Paraphilic disorders; Psychometrics; Scoping review.

Resumo: Introdução: A distinção entre parafilia e transtorno parafílico, consolidada no DSM-5, exige avaliação criteriosa para evitar tanto patologização indevida quanto negligência clínica. Contudo, os instrumentos disponíveis diferem quanto ao constructo, contexto e robustez psicométrica. Objetivo: Mapear e caracterizar instrumentos padronizados utilizados na avaliação de parafilias e transtornos parafílicos em adultos. Métodos: Realizou-se revisão de escopo, orientada pelo modelo PCC, com busca nas bases PubMed/MEDLINE e BVS, sem restrição temporal e com inclusão de estudos em português, inglês e espanhol. Foram elegíveis pesquisas de desenvolvimento, adaptação ou validação de instrumentos padronizados voltados à mensuração de parafilias ou transtornos parafílicos em adultos. Excluíram-se medidas exclusivamente psicofisiológicas ou indiretas. A seleção ocorreu em duas etapas, por avaliadores independentes, com síntese descritiva por instrumento único. Resultados: Dos 1.696 registros identificados, 10 estudos foram incluídos, todos provenientes do PubMed/MEDLINE. Identificaram-se dois grupos principais: instrumentos de autorrelato, voltados à mensuração dimensional de interesses e traços; e escalas baseadas em registros, aplicadas, sobretudo, em contextos forenses. As evidências psicométricas mais frequentes envolveram análise de estrutura interna, invariância e validade por critério, com variabilidade na descrição de confiabilidade e aplicabilidade clínica. Observou-se ausência de instrumentos elegíveis indexados na BVS. Conclusão: A mensuração de parafilias organiza-se em modelo complementar entre autorrelato e escalas baseadas em registros, persistindo lacunas quanto à padronização e disponibilidade de instrumentos adaptados ao contexto brasileiro.

Palavras-chave: Transtornos Parafílicos; Psicometria; Revisão de Escopo.

[1]Especialização. Universidade Estadual de Ponta Grossa, Departamento de Medicina, Ponta Grossa, Brasil. feliciofnetto@gmail.com
[2]Graduação. Universidade Estadual de Ponta Grossa, Departamento de Medicina, Ponta Grossa, Brasil.
[3]Mestrado. Professora Assistente. Universidade Estadual de Ponta Grossa, Departamento de Medicina, Ponta Grossa, Brasil. tatimenezes@gmail.com
[4]Doutorado. Professor Associado. Universidade Estadual de Ponta Grossa, Departamento de Medicina, Ponta Grossa, Brasil. zanetticons@uol.com.br

DOI: https://doi.org/10.35919/rbsh.v37.1383

Resumen: Introducción: La distinción entre parafilia y trastorno parafílico, consolidada en el DSM-5, exige una evaluación rigurosa para evitar tanto la patologización indebida como la negligencia clínica. Sin embargo, los instrumentos disponibles difieren en cuanto al constructo, el contexto de aplicación y la solidez psicométrica. Objetivo: Mapear y caracterizar los instrumentos estandarizados utilizados en la evaluación de parafilias y trastornos parafílicos en adultos. Métodos: Se realizó una revisión de alcance, guiada por el marco PCC, mediante búsquedas en las bases de datos PubMed/MEDLINE y BVS, sin restricción temporal e incluyendo estudios en portugués, inglés y español. Fueron elegibles las investigaciones de desarrollo, adaptación transcultural o validación psicométrica de instrumentos estandarizados destinados a la medición de parafilias o trastornos parafílicos en adultos. Se excluyeron las medidas exclusivamente psicofisiológicas o indirectas. La selección de estudios se realizó en dos etapas por evaluadores independientes, con síntesis descriptiva organizada por instrumento individual. Resultados: De los 1.696 registros identificados, se incluyeron 10 estudios, todos recuperados de PubMed/MEDLINE. Se identificaron dos grupos principales de instrumentos: medidas de autoinforme orientadas en la evaluación dimensional de intereses y rasgos; y escalas basadas en registros, aplicadas predominantemente en contextos forenses. Las evidencias psicométricas reportadas con mayor frecuencia incluyeron análisis de estructura interna, invariancia de medida y validez de criterio, con variabilidad en la descripción de la confiabilidad y la aplicabilidad clínica. Se observó la ausencia de instrumentos elegibles indexados en la BVS. Conclusión: La medición de las parafilias se organiza en un modelo complementario entre autoinformes y escalas basadas en registros, con persistentes vacíos en la estandarización y en la disponibilidad de instrumentos adaptados al contexto brasileño.

Palabras clave: Trastornos parafílicos; Psicometría; Revisión de alcance.

Introduction

Atypical sexual interests constitute a spectrum historically described in sexology and psychiatry, requiring conceptual distinctions to avoid both undue pathologization and the neglect of conditions that, in fact, require clinical attention. With the publication of the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5), the distinction between paraphilia, defined as an atypical sexual interest, and paraphilic disorder, when there is clinically significant distress, functional impairment, and/or involvement in behavior with the potential to harm others, especially in the context of non-consent, became more explicitly established. This distinction has direct implications for assessment, management, and clinical communication, including through its effects on stigma, access to care, and professional decision-making (Lucena; Abdo, 2014).

Although the DSM-5 constitutes the dominant reference in the contemporary literature regarding the distinction between paraphilia and paraphilic disorder, this framework does not exhaust the classificatory and interpretive possibilities of the phenomenon. Different classification systems, clinical traditions, and sociocultural contexts may modulate how atypical sexual interests, distress, impairment, and risk are defined, assessed, and communicated (Barbieri Filho, 2017).

In clinical setting, the assessment of paraphilic disorders is particularly challenging because it involves heterogeneous phenomena, with different levels of insight, comorbidities, and patterns of treatment seeking or avoidance. In addition, the boundary between interest, fantasy, behavior, and distress may be modulated by individual and contextual factors, making diagnostic inference based solely on the presence of an atypical interest insufficient. For this reason, an assessment that integrates clinical history, psychosocial functioning, the presence of distress and impairment, and careful risk analysis when applicable is recommended (Lucena; Abdo, 2014; Barbieri Filho, 2017).

The international literature describes a broad set of strategies for measuring sexual interest and its clinical correlates, including: (i) structured interview and clinical examination; (ii) self-report scales and questionnaires; (iii) indirect measures, for example, post-viewing reaction time; and (iv) psychophysiological measures, such as penile plethysmography (phallometry). Reviews and syntheses on phallometry highlight that, despite its long-standing use, questions remain regarding the standardization of stimuli, interpretation parameters, sensitivity, specificity, and applicability across different services, reinforcing the need for caution and contextualized interpretation of findings (Bickle et al., 2021; Marshall et al., 2014).

Even in subareas in which objective measurement has been more extensively explored, such as investigations of sexual interest related to risk behaviors, there are relevant debates regarding incremental validity, generalizability of findings, and ethical limitations. Critical studies point to inconsistencies across studies, methodological variability, and the risk of divergent conclusions when protocols and interpretive criteria differ,

which may compromise comparisons between samples and services (Bickle et al., 2021; Clegg; Fremouw, 2009).

In parallel, indirect measures and performance-based protocols have been proposed as alternatives or complements to phallometry. Classic studies compared post-viewing reaction time measures with plethysmography, seeking greater objectivity in the assessment of sexual interest, but such approaches also depend on validation, bias control, and an understanding of psychometric limitations (Abel et al., 1998; Carvalho et al., 2020). Therefore, critical analyses of the Abel Assessment for Sexual Interest (AASI), a computerized screening instrument designed to measure sexual interests, including paraphilias and pedophilia, through analysis of the viewing time of 160 images, created by American psychiatrist Gene Abel, and related instruments emphasize that clinical and forensic decisions should be cautious when evidence regarding statistical adequacy and interpretation parameters in diverse populations is insufficient (Fischer; Smith, 1999).

More recently, efforts have been observed to develop and validate self-report instruments with a scope aligned with DSM-5 categories, aiming to expand measurement in community and clinical samples and reduce exclusive dependence on laboratory methods. The Paraphilic Interests and Disorders Scale (PIDS) exemplifies this trend by proposing a systematic assessment of paraphilic interests or disorders and reporting stages of content validation and pilot testing (Winters; Jeglic; Kaylor, 2023).

Similarly, the Questionnaire for Evaluation of Paraphilic Disorders (CETP) was presented with validation and standardization results in its context of application, reinforcing the contemporary movement toward structuring measures with explicit psychometric evidence (Delcea; Siserman, 2020). In addition, there are initiatives to develop more general paraphilia scales, with particular interest in standardization and measurement in specific contexts (Wakeling et al., 2021).

Despite this progress, the landscape remains fragmented: instruments differ in terms of the construct assessed (interest, disorder, compulsivity, risk, and distress), context (clinical, community, research, and forensic interface), reported outcomes, and degree of validation. In practical terms, this heterogeneity complicates the selection of appropriate measures, comparison of studies, and development of recommendations for professional training and services. Reviews focused on specific components of assessment, such as the measurement of pedophilic interest, clearly illustrate the multiplicity of available techniques and the absence of a universally accepted standard, highlighting the usefulness of systematically mapping methods and instruments (Carvalho et al., 2020).

In the Brazilian context, the need for systematization is particularly relevant. There is national literature addressing clinical debate, management, and reflections focused on clinical practice and cognitive-behavioral functioning in patients with paraphilias, but the availability, adaptation, validation, and routine use of specific instruments still tend to appear in a dispersed manner across fields and services. This scenario may result in assessment practices that are highly dependent on individual experience, with substantial variations in quality and comparability, in addition to hindering applied research in human sexuality (Barbieri Filho, 2017; Rodrigues, 2009).

Therefore, this review is justified to map which assessment instruments and methods have been used in the literature on paraphilic disorders, in which populations and contexts, which psychometric evidence and limitations are reported, and which gaps persist, particularly regarding the applicability and availability of measures in Portuguese and in Brazil. From this perspective, the objective of this study is to characterize standardized instruments used to assess paraphilias and paraphilic disorders in adults. Additionally, it seeks to discuss the extent to which such instruments measure atypical sexual interests, indicators of disorder, behavioral markers, risk, distress, or severity, considering that these dimensions are not necessarily equivalent from clinical, psychometric, and nosological perspectives.

Method

Type of study and reporting guidelines

This is a scoping review, selected because it allows the systematic mapping of assessment instruments,

their uses, and gaps in a field characterized by diversity in definitions, populations, and study designs. Accordingly, the reporting of the study was guided by the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) and by the Joanna Briggs Institute (JBI) methodological guidance for scoping reviews (Peters et al., 2015).

Research question and PCC framework

The protocol for this scoping review was retrospectively registered in the Open Science Framework (OSF) under DOI: 10.17605/OSF.IO/G8RYJ. The research question was structured according to the PCC framework, where (P) Population; (C) Concept; (C) Context.

P (Population): adults (≥ 18 years);

C (Concept): standardized assessment instruments (scales, questionnaires, inventories, diagnostic checklists, and structured or semi-structured interviews) intended to measure paraphilias and/or paraphilic disorders, including diagnostic indicators, intensity, severity, distress, impairment, and dimensions directly related to the construct;

C (Context): clinical practice and research in sexuality and mental health; studies at the interface with forensic services were eligible when the primary focus was the instrument rather than forensic procedures or psychophysiological measures.

Guiding question: "Which psychometric instruments and standardized interviews have been used to assess paraphilias and paraphilic disorders in adults, and which characteristics and psychometric evidence are reported?"

Databases

The search was conducted in PubMed/MEDLINE and BVS because of their broad biomedical coverage and the ease of traceability of search strategies. To reduce the risk of omitting relevant instruments not retrieved through the main strategy, reference tracking (bibliographic lists) of the included studies and searches for related studies were performed.

Search strategy

The strategy was constructed using two blocks of terms combined with Boolean operators: (i) terms related to paraphilias and paraphilic disorders; and (ii) terms related to psychometric instruments or interviews and measurement properties. No time restrictions were applied to avoid the loss of classic instruments. Texts in Portuguese, English, and Spanish were considered eligible (Table 1).

Eligibility criteria

Inclusion

Studies were included if they:

1. Presented the development, cross-cultural adaptation, or psychometric validation of an instrument focused on paraphilias or paraphilic disorders in adults;
2. Applied a standardized instrument for this purpose and described, at minimum, the instrument name, purpose, structure, and method of interpretation.

Exclusion

The following were excluded:

1. Studies in which the central assessment was psychophysiological (plethysmography or phallometry) or based on indirect tasks (reaction time, viewing time, eye-tracking);

DOI: https://doi.org/10.35919/rbsh.v37.1383

2. Intervention or treatment studies without a standardized instrument for assessing the paraphilic construct;
3. Opinion pieces, editorials, comments, non-systematized reviews, and clinical reports without a reproducible standardized instrument;
4. Instruments focused exclusively on risk or general sexual violence recidivism without specific measurement of interest, paraphilia, or paraphilic disorder.

Study selection

Selection occurred in two stages:

1. Title and abstract screening: identification of potentially eligible records according to the defined criteria;
2. Full-text review: confirmation of eligibility and extraction of the information required for synthesis. Title and abstract screening and full-text review were performed by two independent reviewers. Disagreements were resolved by consensus, and consultation with a third reviewer was not necessary.

Single-instrument restriction strategy and selection of index articles

Considering that broad searches retrieve a large volume of records, a synthesis strategy centered on the objective of the review was adopted: to "map instruments," rather than necessarily all studies that use them. Thus, after eligibility assessment, the articles were grouped by unique instrument, that is, for each eligible instrument, index articles were selected according to a fixed and reproducible rule:

1. Index article 1: publication describing instrument development;
2. Index article 2: main validation publication, preferably in an independent sample and with greater psychometric completeness.

When multiple options were available, tie-breaking criteria were applied in the following order: (i) larger sample size; (ii) broader range of reported psychometric properties; (iii) greater clarity regarding clinical application and interpretation; and (iv) priority given to Portuguese-language adaptations when available. This procedure allowed the maintenance of a comparable final corpus without losing representation of the available instruments.

Therefore, the adoption of index articles served an organizational purpose and was not intended to exhaust all empirical applications of each instrument. To reduce the loss of relevant information, index articles were selected by prioritizing studies of development, initial validation, or more comprehensive psychometric validation. Additional application studies, when they did not add new evidence regarding development, adaptation, or validation, were not included in the main corpus. This decision was made to preserve comparability among instruments and avoid overrepresentation of measures with a greater number of application studies.

Data extraction

Data extraction was performed independently by two reviewers using a previously tested standardized form. The extracted information was compared between reviewers, and discrepancies were resolved by consensus. When necessary, the studies were reread to confirm data related to instrument structure, sample type, application context and psychometric properties. A formal coefficient was not calculated, since disagreements were resolved by consensus before the final synthesis.

Critical appraisal of methodological quality

No formal assessment of risk of bias or methodological quality of the included studies was performed, in accordance with the exploratory purpose of scoping reviews, whose primary objective is to map the extent, nature, and characteristics of the available evidence. Nevertheless, information related to the psychometric evidence reported in each study was extracted, including internal structure, reliability, criterion validity, measurement invariance, sensitivity, specificity, and properties based on item response theory. Thus, the methodological quality of the instruments was discussed descriptively, without assigning a formal quality score.

Revista Brasileira de
Sexualidade Humana ISSN 2675-1194
DOI: https://doi.org/10.35919/rbsh.v37.1383

Synthesis of results

The findings were synthesized descriptively and in tabular form, organized by instrument. The presentation included: (i) a map of the identified instruments; (ii) a comparative table of structure and psychometric properties; and (iii) a narrative synthesis highlighting gaps and implications for research and clinical practice.

Results

The database search identified 1,658 records in PubMed/MEDLINE and 38 records in BVS, totaling 1,696. After applying the eligibility criteria defined for this review, restricted to standardized psychometric instruments and structured or semi-structured interviews specifically aimed at the assessment of paraphilias and/or paraphilic disorders in adults, none of the 38 records retrieved from BVS were found to meet the objective of the research question (Figure 1). In PubMed/MEDLINE, after the screening and eligibility process, 1,648 records were excluded, resulting in the final inclusion of 10 studies, encompassing publications between 2004 and 2025, as presented in Table 2.

The identified instruments were distributed into two main groups: "self-report instruments," applied primarily in community samples and aimed at the dimensional measurement of related interests and traits; and "record-based scales," constructed from behavioral markers and documentary information, with predominant application in forensic or correctional settings.

Among the self-report instruments with broader scope, the Paraphilic Interests and Disorders Scale was developed to measure paraphilic interests and disorders aligned with DSM-5 diagnostic categories. The Paraphilia Scale, in turn, showed a multifactorial structure in a large community sample, with investigation of measurement invariance across gender and sexual orientation groups, including initial stages of content validation with experts and pilot testing in the general population. In addition, the Massachusetts Treatment Center (MTC) Sexual Sadism Scale, a dimensional paraphilia scale, was subjected to internal structure analysis in a large community sample, supporting a multifactorial model and assessment of measurement equivalence across subgroups. A specific scale for somnophilia, the Somnophilia Interest and Proclivity Scale, was also identified, comprising three dimensions, distinguishing consensual and non-consensual components of interest, and applied in an online sample.

In the domain of screening instruments related to pedophilic interest, the Screening Scale for Pedophilic Interests (SSPI) and the Revised Screening Scale for Pedophilic Interests (SSPI-2) were identified. Both use victim characteristics and objective case information to generate a screening score. The SSPI was evaluated in two independent samples, with associations with recidivism outcomes in part of the analyses. The revised instrument, SSPI-2, included the additional item of child pornography and was tested in large development and validation samples, with evidence of criterion validity.

Regarding the construct of sexual sadism, record-based measures and self-report measures were identified. Among the record-based scales, an 11-item dichotomous scale, the Severe Sexual Sadism Scale (SSSS), showed support for a unidimensional structure and definition of a cutoff point with sensitivity and specificity estimates. Another scale developed for a sexual homicide sample, the Sexual Sadism Scale (SeSaS), showed a single-factor structure and precision analysis according to trait severity, as well as prevalence estimates in the studied sample. A third measure of sexual sadism (MTC Sexual Sadism Scale) was presented as a dimensional approach based on behavioral markers and with psychometric support reported in the study.

Among the self-report measures associated with sadism, an instrument focused on consensual sadism was identified, with evidence of invariance across groups and low influence of social desirability (Index of Consensual Sexual Sadism, ICSS), as well as an agonistic continuum scale (The Agonistic Continuum Scale, TACS), with a multifactorial structure and different model-fit results when between-group comparisons were performed.

Overall, the psychometric evidence most frequently reported in the included index articles comprised analyses of internal structure, particularly confirmatory factor analysis and, in some cases, item response theory modeling, as well as cutoff point definition with accuracy estimates. In contrast, variability was observed in the level of detail regarding reliability and in the availability of complete information for clinical application

and interpretation, particularly for more recent instruments or in studies whose primary focus was structural validation in specific populations (Table 3).

Discussion

The findings of this review show that the literature eligible for the proposed research question was concentrated entirely in PubMed/MEDLINE, with exclusion of all records retrieved from BVS. In addition to describing an evidence gap, this result suggests that, within the Latin American literature indexed in BVS, production is predominantly focused on clinical, ethical, and nosological discussion of the construct, which is important for conceptual framing but insufficient to support standardization of measurement and comparability across studies and services. This asymmetry is consistent with national publications that emphasize the DSM-5 distinction between paraphilic interest and paraphilic disorder and the ethical and clinical implications of this differentiation, but do not necessarily present standardized instruments to operationalize it in practice (Lucena; Abdo, 2014; Barbieri Filho, 2017).

This absence of eligible studies in BVS should not be interpreted as indicating a lack of Latin American literature on paraphilias or paraphilic disorders, but rather as the absence, within the criteria of this review, of studies indexed in this database that presented the development, adaptation, or psychometric validation of standardized instruments for adults. Moreover, this finding may also reflect indexing limitations, terminological variations, dispersion of publications across journals not indexed in the databases consulted, and a predominance of publications focused on clinical, ethical, or forensic discussion without a psychometric focus.

When synthesizing the eligible instruments, two measurement ecosystems with distinct rationales emerged: (i) self-report instruments, aimed at the dimensional measurement of interests/traits and applied mainly in community samples; and (ii) record-based instruments, constructed from behavioral/documentary markers and applied primarily in forensic/correctional settings (Winters; Jeglic; Kaylor, 2023; Seto et al., 2025; Deehan; Bartels, 2021; Kinrade et al., 2025; Davidson; Healey, 2024; Seto et al., 2004; Seto et al., 2017; Mokros et al., 2012; Stefanska et al., 2019; Longpré; Guay; Knight, 2019).

The distinction between self-report instruments and record-based instruments should be understood as a heuristic organization of predominant poles rather than as a rigid or exhaustive classification of the field. Some instruments may combine subjective, behavioral, and contextual information, whereas others align more clearly with one of these poles. Nevertheless, this division helps clarify that instruments differ in terms of data source, operationalized construct, context of use, and the type of inference they permit.

Within the self-report domain, a relevant methodological aspect is that psychometric robustness depends not only on good factorial indices but also on measurement equity across subgroups. In the case of the Paraphilia Scale, the authors tested factorial structure and invariance, identifying limitations at metric levels in comparisons between groups, which implies caution when making comparative inferences based on scores. This issue is particularly sensitive in human sexuality, as instruments may simultaneously capture individual traits and sociocultural contextual effects (norms, social desirability, self-censorship, item wording), producing apparent differences between groups that may reflect differential instrument functioning (Seto et al., 2025).

Also within the self-report domain, the Paraphilic Interests and Disorders Scale (PIDS) was proposed as an instrument targeting the eight DSM-5 categories of paraphilic disorders, based on content validation with experts and pilot testing, indicating potential for population-based research when the objective is to estimate dimensions of interest with nosological alignment. However, because this instrument is still undergoing consolidation, its usefulness for clinical decisions and comparisons between subgroups depends on continued studies to validate its internal structure, invariance, and criterion validity, which is consistent with general recommendations for the assessment of paraphilias that emphasize integration of multiple sources and caution regarding self-report biases (Winters; Jeglic; Kaylor, 2023; Carvalho et al., 2020).

In self-report instruments, interpretation of scores requires additional caution, because respondents may underreport interests, fantasies, or behaviors due to shame, fear of sanctions, social desirability, or limited perception of a problem. Moreover, interest, fantasy, arousal, behavior, and distress are not equivalent dimensions. An instrument may capture a greater willingness to report certain sexual content without necessarily indicating a paraphilic disorder, behavioral risk, or functional impairment. Thus, self-report measures tend

Revista Brasileira de
Sexualidade Humana ISSN 2675-1194
DOI: https://doi.org/10.35919/rbsh.v37.1383

to be more useful for dimensional investigation and population-based research and should be integrated with clinical interviews and other sources of information when used in diagnostic settings.

Among specific instruments, the scale for somnophilia exemplifies a relevant conceptual advance: separating consensual and non-consensual dimensions of interest. This distinction is operationally important in human sexuality because it prevents the instrument from conflating "atypical interest" with conditions involving risk, aligning with the DSM-5 principle that not every paraphilia constitutes a disorder. From the perspective of the scientific agenda, instruments with subscales that discriminate consensual and non-consensual components may support research seeking to define more precisely what constitutes "interest" and what constitutes clinical risk (Deehan; Bartels, 2021; Lucena; Abdo, 2014).

Within the domain of record-based instruments, the findings show a concentration on pedophilic interest and sexual sadism. The Screening Scale for Pedophilic Interests (SSPI-1) and the Revised Screening Scale for Pedophilic Interests (SSPI-2) were developed as brief scales derived from case characteristics, including, in the revised version (SSPI-2), an item related to child pornography, with criterion validation in forensic samples. This type of measure is consistent with literature discussing the assessment of pedophilic interest as a complex task, often dependent on integration between clinical interview, objective data, and, in some contexts, physiological methods (Seto et al., 2004; Seto et al., 2017; Carvalho et al., 2020; Marshall, 2014).

With regard to sexual sadism, the group comprising the Severe Sexual Sadism Scale (SSSS), Sexual Sadism Scale (SeSaS), and MTC Sadism Scale reflects a tradition of measurement centered on behavioral markers and dimensional modeling. The SSSS describes a unidimensional structure in a forensic sample and provides a cutoff point with sensitivity and specificity parameters; the SeSaS supports criterion validity in a sexual homicide sample and includes item response theory analysis; and the MTC Sadism Scale is presented as a dimensional measure with good properties in sexual offenders assessed in a correctional institution. Taken together, these instruments appear more appropriate for screening in contexts with complete case documentation, but the very "file-based" nature of these measures imposes a limitation: score quality depends on what was recorded, coded, and interpreted consistently (Mokros et al., 2012; Stefanska et al., 2019; Longpré; Guay; Knight, 2019).

Record-based instruments offer pragmatic advantages in forensic contexts, particularly when self-report is limited or unreliable. However, their validity depends on the quality, completeness, and purpose of the available records. Judicial, clinical, or correctional records are not neutral sources: they may reflect institutional priorities, investigative biases, inequalities in access to legal defense, penal selectivity, and categories previously used by professionals within the system. Therefore, there is a risk of circular validation when an instrument is developed and validated from records that were themselves produced under assumptions similar to the criteria the instrument is intended to measure. This limitation does not invalidate such scales, but it requires cautious interpretation and attention to interrater reliability, standardization of coding, and independence of external validation criteria.

An additional and clinically sensitive point is that recent literature has explicitly emphasized the need to delimit the spectrum of sexual sadism, including distinguishing consensual components, such as "bondage, discipline, sadism, and masochism" (BDSM), from coercive forms. In this regard, the ICSS provides evidence of a unidimensional structure, invariance across tested groupings, and a trivial relationship with social desirability, which favors research on consensual sexual sadism without automatically collapsing it into forensic categories. The TACS, in turn, as a measure of the agonistic continuum, emphasizes the challenge of stabilizing structure and comparability in heterogeneous samples, reinforcing the idea that measurement in this domain is highly sensitive to sample composition and measurement equivalence (Kinrade et al., 2025; Davidson; Healey, 2024).

The distinction between consensual and non-consensual dimensions has relevant clinical and ethical implications. Instruments that differentiate consensual practices, fantasies, or atypical interests from coercive behaviors contribute to reducing undue pathologization and stigma while preserving the ability to identify situations associated with distress, functional impairment, or risk to others. This separation is particularly important in clinical and forensic assessments, in which conflating consensual atypical interest with disorder may lead to excessive diagnostic inferences, whereas neglecting non-consensual components may compromise risk assessment.

From a broader methodological perspective, the synthesis obtained in this study converges with reviews

DOI: https://doi.org/10.35919/rbsh.v37.1383

and reference articles describing the assessment of sexual interest as a field in which no single source is sufficient: self-report has limitations and may be affected by desirability, whereas objective and indirect measures face operational, acceptability, and clinical interpretation challenges. The literature on plethysmography and phallometric assessment discusses both measurement potential and limitations and challenges to widespread adoption; and reviews on the assessment of pedophilic interest reiterate that clinical standardization remains a persistent problem despite the availability of multiple techniques (Marshall, 2014; Bickle et al., 2021; Clegg; Fremouw, 2009; Carvalho et al., 2020).

In the Brazilian context, the absence of eligible instruments in BVS should be interpreted as a sign that locally indexed production, within this scope, still prioritizes nosological discussion and analyses of broader forensic assessment practices. There are national publications that directly address diagnosis and ethics in paraphilic disorders and others that address forensic psychological assessment and the risk of sexual violence recidivism, revealing the applied relevance of the topic. However, this body of literature does not necessarily translate into specific validated instruments for measuring paraphilic disorders, thereby hindering comparability (Lucena; Abdo, 2014; Barbieri Filho, 2017; Lira-Cardoso et al., 2020; Fonseca; Setubal; Costa, 2019; Meneses et al., 2016).

This review was restricted to the PubMed/MEDLINE and BVS databases, which may have excluded psychometric studies indexed in psychology and behavioral science databases. In addition, the synthesis was organized by unique instrument, a strategy appropriate for mapping the "universe of instruments," but not intended to exhaustively represent all empirical applications of each measure.

The main practical implication, and the most productive point for the Brazilian field, is that the "gap" observed in BVS does not mean an absence of clinical debate, but rather an absence of instruments that can be locally integrated into service routines and cumulative research. Therefore, a realistic agenda is not to "create from scratch" a large national inventory, but rather to critically import: to select a small set of instruments already established in PubMed/MEDLINE, adapt them, and test them with emphasis on invariance, to avoid biased interpretations across groups, and on reproducibility, in the case of file-based measures.

Rather than merely indicating that there are few instruments, the results allow a more specific interpretation: the measurement of paraphilias and paraphilic disorders appears to be organizing around a triangulation model, in which the same construct is approached through three complementary pathways: (i) dimensional self-report (more sensitive to the subjective phenomenon and sociocultural context), (ii) behavioral markers in records (more useful when self-report is unreliable or when there are clear forensic objectives), and (iii) objective and/or indirect methods discussed in the literature, although not included within this scope by criterion (for example, plethysmography), which function as external anchors when available.

Regarding the included studies, the answer to the question "what do the instruments actually measure?" is not singular. The identified instruments do not directly measure "paraphilia" or "paraphilic disorder" as homogeneous clinical entities; rather, they operationalize specific indicators that depend on the method and data source. Self-report measures tend to capture interests, fantasies, subjective dispositions, traits, and, in some cases, associated distress. Record-based scales capture documented behavioral markers, victim characteristics, offense patterns, or retrospective indicators of severity. Thus, the instruments measure operational approximations of the construct rather than the construct in its entirety.

Conclusion

This scoping review identified standardized instruments designed to assess paraphilias and paraphilic disorders in adults, organized predominantly into two groups: self-report measures and record-based scales. The findings indicate that these instruments differ substantially concerning the operationalized construct, data source, context of application, and robustness of the available psychometric evidence.

The central research question should be answered with caution: the instruments do not directly capture the full clinical spectrum of paraphilic disorder, but rather specific indicators of interest, fantasy, behavior, severity, risk, distress, or documented markers. Therefore, their use requires contextualized interpretation and integration with clinical, psychosocial, and, when relevant, forensic assessment.

In the Brazilian context, the results indicate a relevant gap in the availability of adapted and validated

instruments. Future studies should prioritize cross-cultural adaptation, assessment of measurement invariance, reliability, criterion validity, clinical applicability, and interrater reliability, particularly for record-based instruments.

References

ABEL, G. G. et al. Visual reaction time and plethysmography as measures of sexual interest in child molesters. Sexual Abuse: A Journal of Research and Treatment, v. 10, p. 81-95, 1998. Available from: https://pubmed.ncbi.nlm.nih.gov/15326884/. Accessed: 18 jan. 2026.

BARBIERI FILHO, A. Tratamento medicamentoso dos transtornos parafílicos. Revista Brasileira de Sexualidade Humana, v. 28, n. 2, 2017. Available from: https://www.rbsh.org.br/revista_sbrash/article/view/21. Accessed: 8 jan. 2026.

BICKLE, A. et al. International overview of phallometric testing. BJPsych International, v. 18, n. 4, e11, 2021. Available from: https://pubmed.ncbi.nlm.nih.gov/34747938/. Accessed: 19 jan. 2026.

CARVALHO, J. et al. Measuring pedophilic sexual interest. The Journal of Sexual Medicine, v. 17, n. 3, p. 378-392, 2020. Available from: https://pubmed.ncbi.nlm.nih.gov/31932255/. Accessed: 8 jan. 2026.

CLEGG, C.; FREMOUW, W. Phallometric assessment of rapists: a critical review of the research. Aggression and Violent Behavior, v. 14, n. 2, p. 115-125, 2009. Available from: https://psycnet.apa.org/record/2009-03426-005. Accessed: 18 jan. 2026.

DAVIDSON, M.; HEALEY, J. Assessing the Agonistic Continuum Scale as a Measure of Sexual Sadism in a Sample of Community Members and BDSM Practitioners. Sexual Offending: Theory, Research, and Prevention, v. 19, e13829, 2024. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC11789441/. Accessed: 19 jan. 2026.

DEEHAN, E. T.; BARTELS, R. M. Somnophilia: Examining Its Various Forms and Associated Constructs. Sexual Abuse, v. 33, n. 2, p. 200-222, 2021. Available from: https://pubmed.ncbi.nlm.nih.gov/31729926/. Accessed: 18 jan. 2026.

DELCEA, C.; SISERMAN, C.-M. The Questionnaire for Evaluation of Paraphilic Disorders (CETP): development, validation and application. Romanian Journal of Legal Medicine, v. 28, n. 1, p. 14-20, 2020. Available from: https://www.academia.edu/43165305/VALIDATION_AND_STANDARDIZATION_OF_THE_QUESTIONNAIRE_FOR_EVALUATION_OF_PARAPHILIC_DISORDERS. Accessed: 9 jan. 2026.

FISCHER, L.; SMITH, G. Statistical adequacy of the Abel Assessment for Interest in Paraphilias. Sexual Abuse, v. 11, n. 3, p. 195-205, 1999. Available from: https://pubmed.ncbi.nlm.nih.gov/10497779/. Accessed: 18 jan. 2026.

FONSECA, M. C. F.; SETUBAL, C. B.; COSTA, L. F. Adulto autor de violência sexual: estudo exploratório de avaliação de risco de reincidência. Gerais: Revista Interinstitucional de Psicologia, Belo Horizonte, v. 12, n. 2, jul./dez. 2019. Available from: https://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1983-82202019000200013. Accessed: 19 jan. 2026.

KINRADE, C. et al. The Index of Consensual Sexual Sadism (ICSS): Scale development, validation, measurement invariance, and nomological network comparisons with everyday sadism. Psychological Assessment, v. 37, n. 4, p. 148-160, 2025. Available from: https://pubmed.ncbi.nlm.nih.gov/39869687/. Accessed: 18 jan. 2026.

LIRA-CARDOSO, Á. et al. Avaliação psicológica de agressores sexuais no contexto brasileiro: instrumentos e perspectivas. Revista Brasileira de Direito Processual Penal, v. 6, n. 1, 2020. Available from: https://revista.ibraspp.com.br/RBDPP/article/view/320. Accessed: 8 jan. 2026.

LONGPRÉ, N.; GUAY, J.-P.; KNIGHT, R. A. MTC Sadism Scale: Toward a Dimensional Assessment of Severe Sexual Sadism With Behavioral Markers. Assessment, v. 26, n. 1, p. 70-84, 2019. Available from:

https://pubmed.ncbi.nlm.nih.gov/29058955/. Accessed: 9 jan. 2026.

LUCENA, B. B. de; ABDO, C. H. N. Transtorno parafílico: o que mudou com o DSM-5. Diagnóstico & Tratamento, São Paulo, v. 19, n. 2, p. 94-96, 2014. Available from: https://docs.bvsalud.org/biblioref/2018/12/967431/de-lucena.pdf. Accessed: 8 jan. 2026.

MARSHALL, W. L. Phallometric assessments of sexual interests: an update. Current Psychiatry Reports, v. 16, n. 1, art. 428, 2014. Available from: https://pubmed.ncbi.nlm.nih.gov/24318320/. Accessed: 19 jan. 2026.

MENESES, F. F. F. et al. Intervenção psicossocial com o adulto autor de violência sexual intrafamiliar contra crianças e adolescentes. Contextos Clínicos, v. 9, n. 1, p. 98-108, 2016. Available from: https://periodicos.ufmg.br/index.php/contextosclinicos/article/view/60628. Accessed: 20 jan. 2026.

MOKROS, A. et al. The Severe Sexual Sadism Scale: cross-validation and scale properties. Psychological Assessment, v. 24, n. 3, p. 764-769, 2012. Available from: https://pubmed.ncbi.nlm.nih.gov/22142424/. Accessed: 19 jan. 2026.

PETERS, M.D. et al. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc, v. 13, n. 3, p. 141-6, 2015. Available from: https://pubmed.ncbi.nlm.nih.gov/26134548/. Accessed: 20 jan. 2026.

RODRIGUES JR, O. M. Distorções cognitivas nas parafilias. Revista Brasileira de Sexualidade Humana, v. 20, n. 1, 2009. Available from: https://www.rbsh.org.br/revista_sbrash/article/view/345. Accessed: 8 jan. 2026.

SETO, M. C. et al. Does the Paraphilia Scale Work for Everyone? Confirmatory Factor Analysis and Measurement Invariance Across Gender and Sexual Orientation Groups. The Journal of Sex Research, v. 62, n. 3, p. 411-420, 2025. Available from: https://pubmed.ncbi.nlm.nih.gov/38832846/. Accessed: 18 jan. 2026.

SETO, M. C. et al. The Revised Screening Scale for Pedophilic Interests (SSPI-2): Development and Criterion-Related Validation. Sexual Abuse, v. 29, n. 7, p. 619-635, 2017. Available from: https://pubmed.ncbi.nlm.nih.gov/26589444/. Accessed: 20 jan. 2026.

SETO, M. C. et al. The screening scale for pedophilic interests predicts recidivism among adult sex offenders with child victims. Archives of Sexual Behavior, v. 33, n. 5, p. 455-466, 2004. Available from: https://pubmed.ncbi.nlm.nih.gov/15305116/. Accessed: 18 jan. 2026.

STEFANSKA, E. B. et al. Sadism among sexual homicide offenders: Validation of the Sexual Sadism Scale. Psychological Assessment, v. 31, n. 1, p. 132-137, 2019. Available from: https://gala.gre.ac.uk/id/eprint/33545/. Accessed: 19 jan. 2026.

WAKELING, H. et al. The development of a scale for general paraphilia. Ministry of Justice Analytical Series. 2021. Available from: https://assets.publishing.service.gov.uk/media/6092a0d58fa8f51b95cc0aa4/the-development-of-a-scale-for-general-paraphilia.pdf. Accessed: 8 jan. 2026.

WINTERS, G. M.; JEGLIC, E. L.; KAYLOR, L. E. A new measure of paraphilic interests and disorders: development and initial validation of the Paraphilic Interests and Disorders Scale (PIDS). Sexual Abuse, v. 35, n. 2, p. 131-163, 2023. Available from: https://pubmed.ncbi.nlm.nih.gov/35400225/. Accessed: 8 jan. 2026.

Recebido em: 16/02/2026
Aprovado em: 27/05/2026

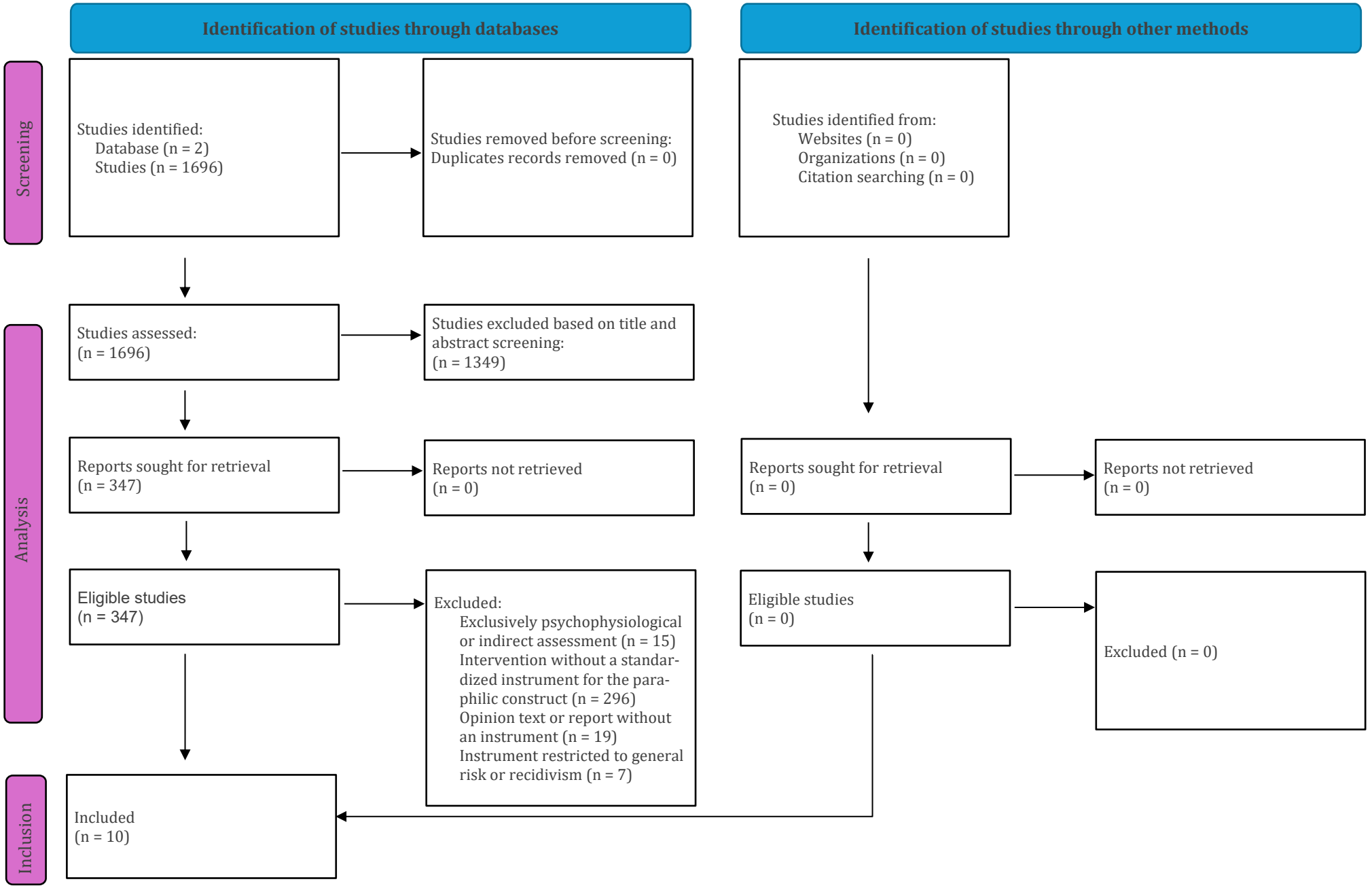
Table 1 - Characterization of the search strategies conducted in the PubMed/MEDLINE and BVS databases

Databse	Search date	Fields searched	Search strategy
PubMed/MEDLINE	January 2026	MeSH; title; abstract	(paraphili* OR "paraphilic disorder*" OR "paraphilic interest*" OR voyeurism OR exhibitionism OR frotteurism OR "sexual sadism" OR "sexual masochism" OR pedophil* OR fetishism OR transvestic*) AND (scale* OR questionnaire* OR inventory OR measure* OR psychometric* OR validation OR reliability OR validity OR "factor analysis" OR "structured interview" OR "semi-structured interview" OR "diagnostic interview" OR screening)
BVS	January 2026	DeCS; title; abstract; subject	(mh:"Transtornos Parafílicos" OR mh:"Voyeurismo" OR mh:"Pedofilia" OR parafili$ OR paraphili$ OR (transtorno parafílico) OR (transtornos parafílicos) OR (paraphilic disorder$) OR (paraphilic interest$) OR voyeurismo OR voyeurism OR exibicionismo OR exhibitionism OR exhibicionismo OR frotteurismo OR frotteurism OR froteurismo OR (sadismo sexual) OR (sexual sadism) OR (masoquismo sexual) OR (sexual masochism) OR pedofili$ OR pedophil$ OR pedofilia OR pedophilia OR fetichismo OR fetishism OR transvestic$ OR transvesti$ OR travestismo OR (transtorno transvéstico) OR (transvestic disorder$)) AND (mh:"Psicometria" OR mh:"Inquéritos e Questionários" OR mh:"Reprodutibilidade dos Testes" OR mh:"Estudo de Validação" OR escala$ OR scale$ OR questionário$ OR questionario$ OR questionnaire$ OR cuestionario$ OR inventário OR inventario OR inventory OR instrumento$ OR instrument$ OR medida$ OR measure$ OR mensuração OR measurement OR psicométric$ OR psicometric$ OR psychometric$ OR validação OR validacao OR validation OR validade OR validity OR confiabilidade OR reliability OR reprodutibilidade OR reproducibility OR (análise fatorial) OR (analise fatorial) OR (factor analysis) OR (entrevista estruturada) OR (structured interview) OR (entrevista semiestruturada) OR (entrevista semi-estruturada) OR (semi-structured interview) OR (entrevista diagnóstica) OR (diagnostic interview) OR rastreamento OR triagem OR screening)

MeSH: Medical Subject Headings; BVS: Virtual Health Library; DeCS: Health Sciences Descriptors.

DOI: https://doi.org/10.35919/rbsh.v37.1383

Figure 1 - PRISMA-ScR flowchart obtained from the bibliographic survey in the databases and other sources



Revista Brasileira de
Sexualidade Humana ISSN 2675-1194
DOI: https://doi.org/10.35919/rbsh.v37.1383

Table 2 - Summary of the instruments and main findings of the selected studies

Study	Construct	Instrument	Sample	Structure	Psychometrics	Notes
Winters, Jeglic & Kaylor (2023)	Paraphilic Interests and Disorders Scale (PIDS)	Self-report	n=122	Instrument for 8 paraphilic interests and disorders (DSM-5)	Initial evidence of content validity; need for additional psychometric support	Intended for population-based data and self-report; still undergoing consolidation
Seto et al. (2025)	Paraphilia Scale	Self-report	n=2310	Structured into 4 factors: coercion (biastophilia, sadism, masochism); chronophilia (pedophilia, hebephilia); courtship disorders (voyeurism, exhibitionism, and frotteurism); fetishism (transvestic, objects, urophilia-coprophilia)	Evidence of configural, but not metric, invariance across gender and sexual orientation groups, suggesting differences in factor loadings	Useful for general structural comparisons; caution is warranted when directly comparing scores across groups
Deehan & Bartels (2021)	Somnophilia Interest and Proclivity Scale	Self-report	n=437	3 subscales: active consensual; passive consensual; and active non-consensual	Empirical pilot study of the construct	Differentiates consensual and non-consensual components; relevant for delineating risk and severity. Men scored higher than women on all subscales except the passive consensual subscale
Seto et al. (2017)	Revised Screening Scale for Pedophilic Interests (SSPI-2)	Structured scoring scale	n=950	5 items 4 victim characteristics (number, age, sex, and relationship to the victims) and 1 child pornography item	Association with sexual arousal to children assessed by plethysmography; superior to the SSPI-1 in relation to the criterion	Recommended as a structured method based on behavior and offense; depends on reliable case data
Seto et al. (2004)	Screening Scale for Pedophilic Interests (SSPI-1)	Structured scoring scale	n=258	Based on characteristics of victims of pedophilia	Correlation with phallometric arousal; violent recidivism and sexual recidivism	Screening instrument with predictive utility in forensic settings; does not replace comprehensive clinical assessment
Mokros et al. (2012)	Severe Sexual Sadism Scale (SSSS)	File-based scale	n=105	11 dichotomous items (yes/no) on behavioral indicators of sexual sadism	Unidimensional structure; good criterion validity for clinical diagnosis; cutoff score of 7 points with S=56% and E=90%	Useful for screening in medical records or files; scale performance depends on the quality of the records

Study	Construct	Instrument	Sample	Structure	Psychometrics	Notes
Stefanska et al. (2019)	Sexual Sadism Scale (SeSaS)	File-based scale	n=350	IRT	CFA confirms a single-factor structure; satisfactory to substantial internal consistency and interrater agreement	Demonstrates applicability in extreme samples (sexual homicide offenders); useful for dimensional and forensic classification
Longpré, Guay & Knight (2019)	MTC Sexual Sadism Scale	Dimensional assessment	n=486	11 items to assess the severity of sadism in sexual offenders	Psychometric properties described as good; captures a wide range of intensity; mentions an IRT-based approach	Instrument designed to grade severity; useful in contexts with reliable behavioral data
Kinrade et al. (2025)	Index of Consensual Sexual Sadism (ICSS)	Self-report	n=1391	Unidimensional structure	Invariance by sample type, sex, and age; validity (distinction from everyday sadism); trivial relationship with social desirability; evidence of reliability	Guides the assessment of the consensual spectrum (important conceptual differentiation from coercive/forensic forms)
Davidson & Healey (2024)	Agonistic Continuum Scale	Self-report	n=246	30 items, 4 subscales: coercion and power; bondage and humiliation; physical violence; and killing	CFA favors a 4-factor model; multigroup analysis by sex worsens model fit	Relevant instrument for the dimensional debate, but caution is required when comparing men and women and when generalizing beyond community samples

DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition; CFA: confirmatory factor analysis; IRT: item response theory; S: sensitivity; Sp: specificity; MTC: Massachusetts Treatment Center

Revista Brasileira de
Sexualidade Humana ISSN 2675-1194
DOI: https://doi.org/10.35919/rbsh.v37.1383

Table 3 - Psychometric summary of the included instruments

Instrument	Internal structure	Confiability	Validity	Invariance	Accuracy	Notes
PIDS	NR	NR	CVI	NR	NR	Self-report aligned with the eight DSM-5 paraphilic disorders; proposed for community samples; requires additional psychometric studies
Paraphilia Scale	Comparative CFA: 13-factor model (CFI=0.95; RMSEA=0.063; SRMR=0.072) and hierarchical model (CFI=0.94; RMSEA=0.067; SRMR=0.083)	McDonald's omega: ω=0.78–0.92	Convergent validity with self-reported paraphilic behaviors; structural validity assessed by CFA, with models tested and replicated	Support for configural invariance; no support for metric and scalar invariance by sex and sexual orientation	NR	Self-report with an interest scale from -3 to +3; in the study, only interest responses were analyzed
SIPS	EFA (varimax): three factors (active consensual, passive consensual, and active non-consensual), explaining 85% of the variance	Cronbach's alpha: "active consensual" α=0.96; "passive consensual" α=0.97; "active non-consensual" α=0.83	Convergent validity, with subscales correlated with congruent fantasies and theoretically related variables, such as necrophilic, sadistic, masochistic, and dominance or submission fantasies	NR	NR	Scenario-based scale; rates arousal, behavioral propensity, and pleasure for each scenario; allows subscale and total scores
SSPI (SSPI-1)	NR	NR	Criterion validity: significant correlation with the phallometric index of sexual arousal to children. Predictive validity: association with violent recidivism in both samples and with sexual recidivism in the larger sample	NR	NR	4 dichotomous items: male victim, multiple victims, victim younger than 12 years, and extrafamilial victim; applied in two samples

SSPI-2	NR	NR	Criterion validity: significant association with phallometric arousal to children in development and validation samples; superior performance compared with the SSPI-1	NR	NR	Structured measure to infer pedophilic interest from behavior/offense; items can be coded from official records and/or self-reported offending history. Differs from the SSPI-1 by the addition of the child pornography item
SSSS	Unidimensional structure by CFA	Good reliability	Good criterion validity for the clinical diagnosis of sexual sadism	NR	S=56% E=90%	11 dichotomous items for file-based assessment of severe sexual sadism
SeSaS	Single-factor structure by CFA; items evaluated using IRT, 2-parameter logistic model	Internal consistency and interrater agreement classified as satisfactory to substantial	Criterion validity: substantial correlation between the total score and the original clinical diagnoses of sadism	NR	NR	National sample of male sexual homicide perpetrators; items assess moderate to severe levels of the trait ($\theta > 0$)
MTC ***Sexual Sadism Scale***	Unidimensional structure assessed for use in IRT: EFA (oblimin) with 7 factors (51.4%); first factor = 30.4%. Single-factor CFA: $\chi^2(299) = 529.55$; RMSEA = 0.04; CFI = 0.92; TLI = 0.92	Mean KR-20 = 0.72 (27 indicators); after excluding "impulsivity in the offense," KR-20 = 0.74 (26 indicators)	Good psychometric properties reported for dimensional assessment using behavioral markers	NR	NR	File-based scale for sexual offenders assessed in a correctional institution
ICSS	Unidimensional structure in initial psychometric testing	Good reliability	Construct validity: distinct from "everyday sadism"; associated with traits such as HEXACO, psychopathy, and sadistic fantasy; trivial relationship with social desirability	Invariance supported across sample type, sex, and age	NR	Self-report for consensual sexual sadism
TACS	CFA with 1-, 2-, and 4-factor models; best relative fit for the 4-factor model ($\chi^2 = 875.53$; df = 399; RMSEA = 0.07; SRMR = 0.07; CFI = 0.85; TLI = 0.84)	McDonald's omega: "total" $\omega = 0.94$; subscales "Coercion and Power" $\omega = 0.77$; "Bondage/Humiliation" $\omega = 0.79$; "Physical Violence/Beating" $\omega = 0.82$; "Killing" $\omega = 0.89$	Good validity explored through LPA (severity-based classes) and correlations with measures of paraphilic interest	The multigroup model by sex significantly worsened model fit; configural invariance was not established; subsequent tests were not conducted	NR	30 items across 4 subscales: "Coercion and Power," "Bondage/Humiliation," "Physical Violence/Beating," and "Killing"; total score ranging from 0 to 120; sample predominantly composed of BDSM practitioners

PIDS: Paraphilic Interests and Disorders Scale; SIPS: Somnophilic Interest and Proclivity Scale; SSPI: Screening Scale for Pedophilic Interests; SSPI-2: Screening Scale for Pedophilic Interests-2; SSSS: Severe Sexual Sadism Scale; SeSaS: Sexual Sadism Scale; MTC: Massachusetts Treatment Center; ICSS: Index of Consensual Sexual Sadism; TACS: The Agonistic Continuum Scale; NR: not reported; CFA: confirmatory factor analysis; EFA: exploratory factor analysis; IRT: Item Response Theory; RMSEA: Root Mean Square Error of Approximation; SRMR: Standardized Root Mean Square Residual; CFI: Comparative Fit Index; TLI: Tucker-Lewis Index; χ^2: model chi-square; df: degrees of freedom; α: Cronbach's alpha; ω: McDonald's omega; KR-20: Kuder-Richardson Formula 20; CVI: Content Validity Index; LPA: Latent Profile Analysis; DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition; BDSM: Bondage and Discipline, Dominance and Submission, Sadism and Masochism; θ: theta, parameter or latent trait in IRT; S: sensitivity; Sp: specificity; HEXACO: honesty-humility, emotionality, extraversion, agreeableness, conscientiousness, and openness to experience.