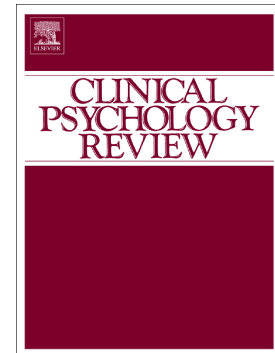# Journal Pre-proof

Screening and assessment tools for gaming disorder: A comprehensive systematic review

Daniel L. King, Samuel R. Chamberlain, Natacha Carragher, Joel Billieux, Dan Stein, Kai Mueller, Marc N. Potenza, Hans Juergen Rumpf, John Saunders, Vladan Starcevic, Zsolt Demetrovics, Matthias Brand, Hae Kook Lee, Marcantonio Spada, Katajun Lindenberg, Anise M.S. Wu, Tagrid Lemenager, Ståle Pallesen, Sophia Achab, Susumu Higuchi, Naomi A. Fineberg, Paul H. Delfabbro

Screening and assessment tools for gaming disorder: A comprehensive systematic review

Daniel L. King[1,*] daniel.king@adelaide.edu.au, Samuel R. Chamberlain[2], Natacha Carragher[3], Joel Billieux[4], Dan Stein[5], Kai Mueller[6], Marc N. Potenza[7], Hans Juergen Rumpf[8], John Saunders[9], Vladan Starcevic[10], Zsolt Demetrovics[11], Matthias Brand[12], Hae Kook Lee[13], Marcantonio Spada[14], Katajun Lindenberg[15], Anise M. S. Wu[16], Tagrid Lemenager[17], Ståle Pallesen[18], Sophia Achab[19], Susumu Higuchi[20], Naomi A. Fineberg[21], Paul H. Delfabbro[22]

[1]*College of Education, Psychology, and Social Work, Flinders University, Adelaide, Australia*

[2]*Department of Psychiatry, University of Cambridge, UK*

[3]*Office of Medical Education, University of New South Wales (UNSW), Sydney, Australia*

[4]*University of Lausanne, Lausanne, Switzerland*

[5]*SAMRC Unit on Risk & Resilience in Mental Disorders, Dept of Psychiatry and Neuroscience Institute, University of Cape Town*

[6]*Outpatient Clinic for Behavioral Addictions, Department of Psychosomatic Medicine and Psychotherapy at the University Medical Center Mainz, Germany*

[7]*Departments of Psychiatry and Neuroscience, Child Study Center, Yale University School of Medicine, New Haven, CT, USA*

[8]*Department of Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Germany*

[9]*Centre for Youth Substance Abuse Research, University of Queensland, Brisbane, Queensland, Australia*

[10]*Discipline of Psychiatry, Nepean Clinical School, Sydney Medical School, The University of Sydney, NSW, Australia*

[11]*Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary*

[12]*General Psychology: Cognition and Center for Behavioral Addiction Research (CeBAR), University Duisburg-Essen, Germany*

[13]*Department of Psychiatry, College of Medicine, The Catholic University of Korea, Seoul, South Korea*

[14]*Division of Psychology, School of Applied Sciences, London South Bank University, London, UK.*

[15]*Institute for Psychology, Heidelberg University of Education, Heidelberg, Germany*

[16]*Department of Psychology, Faculty of Social Sciences, University of Macau, Macao, China*

[17]*Department of Addictive Behavior and Addiction Medicine, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany*

[18]*Department of Psychosocial Science, University of Bergen, Bergen, Norway*

[19]*WHO Collaborating Center for Training and Research in Mental Health, University of Geneva, Switzerland*

[20]*National Hospital Organization Kurihama Medical and Addiction Center, Yokosuka, Japan*

[21]*University of Hertfordshire, University of Cambridge, UK*

[22]*School of Psychology, The University of Adelaide, Adelaide, Australia*

[*]Corresponding author at: School of Psychology, Flinders University, Adelaide, Australia. Tel: +61 8 82019733.

ABSTRACT

The inclusion of gaming disorder (GD) as an official diagnosis in the ICD-11 was a significant milestone for the field. However, the optimal measurement approaches for GD are currently unclear. This comprehensive systematic review aimed to identify and evaluate all available English-language GD tools and their corresponding evidence. A search of *PsychINFO, PsychArticles, ScienceDirect, Scopus, Web of Science*, and *Google Scholar* identified 32 tools employed in 320 studies (N=462,249

participants). The evaluation framework examined tools in relation to: (1) conceptual and practical considerations; (2) alignment with DSM-5 and ICD-11 criteria; (3) type and quantity of studies and samples; and (4) psychometric properties. The evaluation showed that GD instrumentation has proliferated, with 2.5 tools, on average, published annually since 2013. Coverage of DSM-5 and ICD-11 criteria was inconsistent, especially for the criterion of continued use despite harm. Tools converge on the importance of screening for impaired control over gaming and functional impairment. Overall, no single tool was found to be clearly superior, but the AICA-Sgaming, GAS-7, IGDT-10, IGDS9-SF, and Lemmens IGD-9 scales had greater evidential support for their psychometric properties. The GD field would benefit from a standard international tool to identify gaming-related harms across the spectrum of maladaptive gaming behaviors.

Keywords: Gaming disorder; Behavioral Addiction; Screening; Assessment; ICD-11; DSM-5

1. Introduction

Online gaming is a hundred-billion-dollar industry that continues to innovate and expand on a global scale (King & Gaming Industry Response Consortium, 2018). Individuals of all ages are motivated to play games recreationally for relaxation, challenge, and socialization (Yee, 2006). While there are benefits associated with gaming (Granic, Lobel, & Engels, 2014), unrestricted gaming may be highly absorbing and time-consuming, and may become addictive for vulnerable individuals (Brand et al., 2016; Higuchi et al., 2017; King et al., 2019; Sim et al., 2012). Over the last three decades, increasingly accumulated research and clinical evidence has supported recognizing the most severely maladaptive forms of gaming behavior as an addictive disorder (Baggio et al., 2016; Feng, Ramo, Chan, & Bourgeois, 2017; Khazaal et al., 2016; King & Delfabbro, 2019; Meng et al., 2015; Mihara & Higuchi, 2017; Müller et al., 2015; Rehbein & Baier, 2015; Rumpf et al., 2018; Scharkow, Festl, & Quandt, 2014; van Rooij et al., 2010; Yao et al., 2017). However, despite the recognition of gaming disorder (GD), there remains uncertainty regarding optimal approaches to screening and assessment. This uncertainty is due, in part, to the many available, similarly named but varied measurement tools for the condition. To help inform the next phase of research on this new disorder, wherein lies an opportunity for researchers to collaborate and adopt a more consistent approach, the aim of this review was to critically evaluate all available GD tools and their corresponding evidence.

Following a provisional status for 'internet gaming disorder' (IGD) in the DSM-5 (American Psychiatric Association, 2013), gaming disorder was officially adopted at the World Health Assembly in May 2019 as a diagnosis in the eleventh edition of the International Classification of Diseases (ICD-11; WHO, 2019). GD is characterized by persistent gaming behavior, impaired control over gaming, and functional impairment due to gaming for a period of at least 12 months in most instances (Saunders et al., 2017). Individuals with GD play games to the exclusion of other activities, resulting in missed life opportunities and interference with normal routine and basic self-care (i.e., sleep, eating, personal hygiene); real-world social interaction (i.e., meeting friends, visiting family); and important responsibilities (i.e., school, work, care of children) (Allison, Von Wahlde, Shockley, & Gabbard, 2006; Beranuy, Carbonell, & Griffiths, 2013; Griffiths, 2010). Individuals with GD often feel unable to regulate or cease their gaming behavior, and experience intense negative mood states (e.g., irritability, sadness, and boredom) when unable to play (Dong, Wang, Du, & Potenza, 2017; Kaptsis, King, Delfabbro, & Gradisar, 2016). Personal distress may also relate to a fear of missing out on the online game world, where the user feels a strong sense of personal identity and self-efficacy

(King & Delfabbro, 2014; Lemenager et al., 2013; Marino & Spada, 2017; Wegmann, Oberst, Stodt, & Brand, 2017). With the official inclusion of GD as a diagnostic category in the ICD-11, it was considered timely to evaluate the extent to which current instruments were consistent with current defining elements of GD.

Previous reviews and related articles on GD instrumentation have reported various inconsistencies and psychometric weaknesses (King et al., 2013; Griffiths, King, & Demetrovics, 2014; Lortie & Guitton, 2013; Petry et al., 2014; Starcevic, 2013). The most recent major systematic review examined 18 assessment tools employed in 63 studies and reported problems including inconsistent cut-off scores and symptom coverage, and inadequate data on predictive validity and inter-rater reliability (King et al., 2013). Uncertainty has also arisen due to the common research practice of adapting or developing new tools rather than using previous ones. Prior to the provisional DSM-5 criteria for GD, researchers would often adapt the criteria of other disorders (e.g., pathological gambling in the DSM-IV-TR) (Fisher, 1994; Griffiths, 1998). Over time, this practice evolved into adapting these criteria in new ways (e.g., word edits or substitutions, new response categories) and combining other previous items, sometimes sourced from three or more different scales, with new items to create composite measures (e.g., Groves, Gentile, Tapscott, & Lynch, 2015; Jap, Tiatri, Jaya, & Suteja, 2013; Peng & Liu, 2010).

Inconsistent and/or inadequate measurement of GD has major implications for the quality of its research base, as well as for the allocation of clinical and public health resources to address social problems arising from GD. Epidemiological studies that employ short screening tools are relied upon to inform policy decisions and therefore must provide a valid indication of the problem. Epidemiological research on problematic gaming and GD has often been criticized for its sampling approaches, such as recruiting gamers from online gaming forums or sampling a limited pool of students from local schools or universities (van Rooij et al., 2018). However, the psychometric properties of measurement tools in GD studies have received relatively less critical attention. In recent years, there has been an increase in population cohort studies of GD, as well as many large-scale studies (e.g., the Longitudinal Cohort Study on Substance Use Risk Factors [C-SURF] of young men in Switzerland) that have incorporated gaming-related questions. For example, Rehbein et al. (2015) conducted a state-representative school survey of 11,003 adolescents aged 13 to 18 years using the DSM-5 criteria for GD and reported a 1.2% prevalence of GD. A study by Müller et al. (2015) examined GD in seven European countries based on a representative sample of 12,938 adolescents between 14 and 17 years, reporting that 1.6% of the sample met the criteria for IGD, with a further 5.1% at risk for GD by meeting up to four criteria. Other studies have reported comparable figures, including: 0.6% of 816 Norwegian adolescents (Mentzoni et al., 2011); between 0.3 and 1.0% in four international cohorts totalling 18,932 people (Przybylski et al., 2016); 2.0% in a sample of 1,718 Chinese adolescents (Mak et al., 2014); 1.3% in a nationally representative panel of 902 Dutch gamers (Haagsma et al., 2012); 1.5% of Dutch adolescents (van Rooij et al., 2011); and 1.8% of 1,287 Australian adolescents (King et al., 2013). These figures appear to be comparable with prevalence estimates reported for other similar conditions, such as gambling disorder (Calado & Griffiths, 2015). However, there have also been numerous studies of GD that have reported much higher prevalence figures, including rates in excess of 15-20% which seem to defy logic (Seok & DaCosta, 2012; Wang et al., 2014; Xin et al., 2018) and raise concerns about the validity of instrumentation and associated risks such as false positives.

1.4. The present review

The present review aimed primarily to evaluate all available GD and related instruments and their associated empirical evidence base. Although numerous tools refer to IGD in their name and/or source publication, which is the construct used in the DSM-5, this review uses the abbreviation GD to encompass both the DSM-5 and ICD-11 classifications, including when referring to all tools and constituent items. A secondary aim of this review was to provide insights into the nature and quality of the overall evidence base on GD. This evaluation was guided by a similar previous major review conducted by King et al. (2013) prior to the inclusion of IGD as a condition for further study in the DSM-5. It was reasoned that the 2013 review should be updated given that new data may often inform a new consensus on a topic, particularly in a rapidly changing field. The Cochrane Collaboration, for example, recommends that systematic reviews are updated every two years (Moher et al., 2008).

The 2013 systematic review was informed by standards in psychological assessment (Cicchetti, 1994; Groth-Marnat, 2009), which were incorporated into the new framework for the present review. The 2013 review highlighted some of the conceptual inconsistencies across GD tools, as well as gaps in empirical evidence underlying available measures' psychometric properties. In addition to describing the conceptual and practical considerations of all GD tools, the present review sought to address basic questions of: (1) whether current tools were consistent with the DSM-5 and ICD-11 criteria; (2) which tools were being used in which specific research areas (i.e., epidemiological, neurobiological, interventions); and, (3) which tools had received relatively greater evidential support for their psychometric properties.

2. Methods

2.1. Identification and selection of tools

This review aimed to identify and examine all available instruments for screening or assessing problematic gaming and/or GD. Tools were selected for inclusion if they met the following criteria: (1) peer-reviewed and published in English language; (2) accessibility of all test items and response categories (e.g., list in appendix section, or analysis of test items, e.g., factor analysis); (3) primarily designed to measure problematic gaming or gaming disorder, and not internet addiction or other condition (NB: an exception was made for widely used internet addiction tests that refer specifically to gaming activities, e.g., the Young Internet Addiction Test [YIAT; Young, 1998] and Compulsive Internet Use Scale [CIUS; Meerkerk, Van Den Eijnden, Vermulst, & Garretsen, 2009), and for the Screener for Substance and Behavioural Addictions [SSBA; Schluter, Hodgins, Wolfe, & Wild, 2018] as a behavioral addiction tool); (4) the test was self-report (i.e., not completed by an external rater, e.g., parent); and, (5) the test was original and not a composite of two or more existing measures and/or adapted DSM or other diagnostic criteria (e.g., alcohol-use disorder criteria adapted to gaming). A database search was conducted on 4 April 2019 by the first author (DLK). The *Google Scholar* and academic databases, including *PsychINFO, PsychArticles, PubMed, ScienceDirect, Scopus,* and *Web of Science,* were searched using the following keywords and protocol: ((measure or tool or test or validation or psychometric or screening or diagnostic or item or instrument) AND gaming). Figure 1 presents a PRISMA summary of the database search that yielded a total of 5,828 results (including duplicate results). The titles and abstracts of all results generated by each database were screened for relevance using the above inclusion criteria, which led to the identification of 32 tools, including tests with multiple item formats (see Table 1 for the complete list).

2.2. Identification of empirical literature underlying tools

The second phase of the search protocol involved identifying all empirical studies that have employed at least one of the identified 32 tools. This search protocol involved a procedural examination of the *Google Scholar* citation records for all identified tools (as of April 2019). *Google Scholar* was used because it is rapidly updated to include new results across multiple academic databases. Table 3 provides a summary of the citation records for each tool; citation counts were highly variable and ranged between no citations and 5,413 citations. The total citation count for all tools combined was 12,996 (NB: papers citing a paper that presented multiple versions of a tool were counted only once). These records were examined to identify empirical studies published in English. There were no restrictions on inclusion of studies based on study type (e.g., intervention, epidemiological, neurobiological), publication date, or any potential methodological shortcomings. However, articles that were not peer-reviewed (e.g., dissertation material, conference proceedings) were excluded. Supplementary material 1 presents a numbered list of 128 references, with each number corresponding to the evidence cited in superscript format in Tables 5 and 6.

2.3. Tools evaluation framework

This review aimed to provide a comprehensive descriptive summary and critical evaluation of the conceptual and psychometric properties, and practical considerations, of all 32 identified tools. This evaluation was conducted in stages under the guidance and collaboration of members of the research team, which was composed of 14 experts (i.e., psychiatrists, research professors, psychometrician, clinical psychologists) in the field of GD and behavioral addictions. This review was informed by the structure and protocols of the previous systematic review of GD tools conducted by King et al. (2013). The 2013 review was based on 18 tools (N=63 studies) published up to 2012, and therefore preceded the DSM-5 and ICD-11 recognition of GD (NB: the DSM-5 has only recognized IGD as a provisional disorder, or condition in need of further study). The present review sought to address the question of whether available tools were consistent with, and capable of, assessing IGD/GD as described in the DSM-5 and ICD-11 classifications. The 2013 framework was guided by JARS reporting standards (APA Publications and Communications Board Working Group, 2008), as well as Cicchetti's (1994) and Groth-Marnat's (2009) criteria and guidelines for evaluating psychological tests.

The present review aimed to incorporate the 2013 review's framework components, as well as extend the previous review by including more advanced psychometric information. All tools and studies evaluated in this review (i.e., the master spreadsheets underlying all analyses in this review, which were created by DLK) were checked by at least 3 members of the research team (i.e., JB, NC, and PHD). Any discrepancies or errors in ratings or data entry were resolved by consultation among authors. At every stage of the review process, all members of the research team were provided with relevant updates and documentation outlining the procedures and results. All members were invited to contribute any feedback or other observations on this material. All feedback and suggestions were responded to as team messages for transparency. As explained in Section *2.3.4,* there was a roundtable discussion attended in person by 14 team members at the 6th International Conference on Behavioral Addictions in Yokohama, Japan, which enabled discussion of the review's content (distributed electronically before the meeting) and reach consensus decisions on how to present this information in tabular format for a final report.

2.3.1. Review framework I: Overview

First, all 32 tools were summarized according to the following basic characteristics: (1) *tool abbreviation* (i.e., tool acronym or short-hand name; some tools were disambiguated by using new terms, e.g., 'Petry IGD' and 'Lemmens IGD' to minimise potential confusion with other scales with IGD as their namesake); (2) *author and date* (derived from original publication source); (3) *tool components* (i.e., constructs reportedly measured by the tool); (4) *number of items* (i.e., all tool items, including those which may not be included in scoring considerations (e.g., the AICA-Sgaming and C-VAT2.0 include additional items for clinical judgement but are not scored); (5) *response format* (i.e., response options for tool items, e.g., yes/no); (6) *cut-off score* (i.e., cut-off for 'problematic' or 'addicted' status on the test, based on original source reference, if reported); (7) *age* (i.e., participant with the lowest age who completed the tool in the original study); (8) *country of origin* (i.e., country of research team's institution, with priority given to first author); and, (9) *language versions* (i.e., known language versions of the tool, based on the published evidence base only).

2.3.2. Review framework II: DSM/ICD coverage

The second step involved evaluation of each tool's coverage of the DSM-5 and ICD-11 criteria for IGD/GD. Variation in coverage of criteria or guidelines is important because it may, for example, affect prevalence rates in epidemiological studies or affect diagnostic specificity in clinical practice. At the same time, it is acknowledged that total coverage of DSM/ICD criteria is not essential for the purpose of screening, where having fewer items is considered more feasible for the purpose of obtaining insights into, or estimates of, potential harms. Each test's description of its components (i.e., symptoms) were compared to each of the criteria in their respective DSM-5 and ICD-11 categories. The DSM-5 criteria included 9 symptoms (see American Psychiatric Association, 2013); the ICD-11 guidelines included 3 criteria (i.e., 1: impaired control, 2: increasing priority given to gaming; 3: continuation of gaming despite harm; see WHO, 2019) and then functional impairment which was delineated into 5 main areas (i.e., personal [psychological/physical well-being], social, education, work, and financial). The Table 2 legend provides further information on each type of impairment. The rationale for delineating impairment types was to identify tools that may provide contextually useful information on gaming-related consequences, such as greater applicability to certain investigations (e.g., studies of social correlates of gaming) or special populations (e.g., school-related consequences of excessive gaming).

2.3.3. Review framework III: Quantifying the evidence base

The third step of the evaluation involved a detailed summary of the empirical evidence base for all 32 tools. The evidence was summarized according to: (1) *Google Scholar citation count* (i.e., all citations, irrespective of publication type); (2) *Number of empirical studies* (i.e., studies that employed the tool, excluding any non-quantitative studies); (3) *Validation studies* (i.e., studies that involved tool validation based on nationally representative or clinical samples, delineating studies according to those conducted by: (i) the tool's original author, and (ii) independent research teams); (4) *Intervention studies* (i.e., studies involving any type of intervention, e.g., psychotherapy, prevention); (5) *Clinical sample* (i.e., participants with diagnosed GD or probable GD using a

structured clinical interview by a psychiatrist, clinical psychologist, or qualified registrar; participants seeking treatment for GD; and both aforementioned scenarios); (6) *Longitudinal studies* (i.e., studies with repeated observations of the same sample, with no restriction on intervals between observations); (7) *Prevalence studies* (i.e., studies yielding a GD prevalence rate based on a nationally representative sample or subsample [e.g., adolescents] of the population), and; (8) *Neurobiological studies* (i.e., studies that employ imaging technologies, such as functional magnetic resonance imaging (fMRI), and/or neurocognitive measures. In addition, a descriptive summary of the samples' size and age composition, and recruitment approaches, associated with each tool was generated. The review framework delineated: (1) *Total N* (i.e., the total number of participants who have been administered the tool across all studies); (2) *Participant age* (i.e., adolescent [<18 years] vs. adult age); and (3) *recruitment strategy* (i.e., convenience vs non-convenience sampling). This information was then used to calculate the relative proportion of age groups and convenience sampling for each tool.

2.3.4. Review framework IV: Psychometric properties of tools

The fourth step of the review summarized the research evidence on each test's psychometric properties. The broad aim was to determine the nature, quantity, and overall consistency of research support for each tool's validity and reliability. Applying the framework published by King et al. (2013), and extending this framework to consider new areas of test refinement and clinical utility, this undertaking considered the following areas for evaluation: (1) *Dimensionality* (i.e., type of statistical analysis, e.g., exploratory factor analysis; results, e.g., 1 factor solution); (2) *Reliability* (i.e., internal consistency, test-retest coefficients); (3) *Refinement* (i.e., results of Rasch analysis, item response theory [IRT], measurement invariance); (4) *Validity* (i.e., convergent, criterion); (5) *Relationship to impairment* (i.e., association with recognized measure of functional impairment); and, (6) *Clinical use* (i.e., known utility within a clinical interview; use as an outcome measure for a structured intervention).

Information on each of the above psychometric areas was extracted by systematic review of each of the 328 articles for compilation into Excel spreadsheets. This process was assisted by the advanced search function in *Acrobat Reader DC* to identify relevant keywords. For example, the keyword 'factor' was used to identify all factor analyses reported across the 328 studies. Identification and registration of many of these areas involved a simple transposition of a numerical value (e.g., Cronbach's alpha, or a bivariate correlation from a table). For test areas that involved interpretation of statistical analysis (i.e., Rasch, IRT, measurement invariance), the lead author was assisted by an experienced senior psychometrician (NC) who confirmed or clarified the reported results and interpretation (e.g., consistency of reporting with statistical results, such as correct handling of root mean square of approximation [RMSEA] values). All values and interpretation were compiled into a master spreadsheet for review and consultation by co-authors.

2.3.4. Synthesis of review findings

Given the volume of information yielded for each tool, a final summary table was created to provide a parsimonious overview and comparison of the 32 tools. The primary purpose of this table was to provide readers with a quick reference guide for each tool, which may guide decision-making regarding the utility of each tool for particular uses. The framework for this table was developed by a roundtable discussion involving 11 authors (DLK, SC, JB, KM, MNP, HJR, JS, VS, ZD, MB and SH)

and subsequently discussed and approved by all authors. The Table 7 legend provides a complete explanation of the components and scoring information. The scoring rubric was generated for each area to summarize the quantity and/or consistency of research evidence, where applicable. For example, a score of 0 referred to the absence of research evidence; a score of 1 indicated that only 1 study was available, and; a score of 2 indicated that 2 or more studies were available. For reliability indices, a study reporting a value of .70 or above was considered generally sufficient for inclusion (Cronbach, 1951; Lance, Butts, & Michels, 2006; Schmitt, 1996), acknowledging the caveats of this cut-off, e.g., that longer scales tend to have higher alphas (Cicchetti, 1994; Groth-Marnat, 2009). For criterion and convergent validity, an association of .3 or above was the cut-off (i.e., moderate; Cohen, 1992). Scoring was not weighted and thus does not necessarily reflect the overall quality (i.e., validity, reliability, utility) of each tool. Some tools with relatively more highlighted areas may have other specific deficiencies that make them less suitable than other tools. Some tools with fewer highlighted areas may simply reflect that the tool is more recent and therefore has less supporting evidence.

3. Results

3.1. Overview of tools

Table 1 presents a summary of all 32 tools, including original references and countries of origin, tool components, items, scoring information, and language versions. The tools have been presented in ascending order of publication date, with the non-gaming-specific tools (n=6) positioned separately at the end of the list. This order is used consistently for all tables except for Table 4, which is re-ordered to match the distribution of total N data in cells (i.e., descending frequencies) for ease of comprehension.

Inspection of the list of tools and author information shows that some researchers (and/or their respective research teams) have created or contributed to the development of more than one tool. For example, Mark Griffiths (UK) is a named contributor on 5 tools (including 1 short version); Jeroen Lemmens (the Netherlands) is named on 5 tools (including 2 tools with extended versions); and Tony van Rooij (the Netherlands) is named on 2 tools. The most common country of origin was The Netherlands (n=7), followed by South Korea and Germany (n=4 each), and then the United States, United Kingdom, and Hungary (n=3 each). Most tools used continuous response categories, with the most common type being a 5-point scale (n=16). Only 9 tools employ "Yes/No" responses. Most tools (n=22) reported a cut-off score, and most tools (n=29) were reportedly suitable for respondents under the age of 18 years; however, the minimum age varied (n=15 specified 12-13 years of age). Most tools (n=29) were available in non-English languages. Overall, 29 different languages were represented.

3.2. Coverage of DSM-5 and ICD-11 criteria

Table 2 presents a summary of all tools' coverage of the DSM-5 and ICD-11 criteria. This evaluation referred to the 9 criteria for DSM-5 IGD (i.e., preoccupation, withdrawal, tolerance, unsuccessful attempts to stop/limit, loss of interests due to gaming, continued use despite harm, deception, escape, and harm) and the 3 guidelines for ICD-11 GD (6C51) (i.e., impaired control, increasing priority to gaming, and continued use despite harm) and functional impairment (see Section 2.2). This evaluation showed that, overall, there was inconsistent symptom coverage across the 32 tools. With regard to the DSM-5 criteria, the most consistent criterion (n=31 tools; 97%) was

"9: *Has jeopardised or lost a significant relationship, job, or educational or career opportunity due to gaming.*" Other common criteria were symptoms 1 (*preoccupation*), 2 (*withdrawal*), 3 (*unsuccessful attempts*), and 8 (*escape*), with at least 27 tools including each symptom. Only 9 tools included symptom 6 (*continued use despite harm*). In total, only 8 tools provided coverage of all 9 DSM-5 criteria (i.e., Petry IGD, IGDS9-SF, PIE-9, IGDT-10, CVAT2.0, IGUESS, and DIA).

With regard to ICD-11 GD guidelines, all 32 tools including at least one item for "1: *impaired control over gaming.*" Most tools (n=28) measured "2: *Increasing priority given to gaming to the extent that gaming takes precedence over other life interests and daily activities.*" As noted for the DSM-5 criteria, only 9 tools included the ICD-11 criterion "3: *Continuation or escalation of gaming despite the occurrence of negative consequences.*" Coverage of specific types of functional impairment was inconsistent. Consistent with King et al.'s (2013) findings, the most common type of impairment referred to negative social consequences (n=30), followed by negative personal (n=20) and occupational (n=18) consequences. Only 5 tools referred to negative financial consequences of gaming (e.g., debt, overspending on game content; see Brooks & Clark, 2019; Zendle & Cairns, 2018, 2019). In total, only 3 tools provided coverage of all 3 ICD-11 guidelines and 4 domains of functional impairment (i.e., Petry IGD, PIE-9, and IGDT-10), and an additional 4 tools covered all 3 ICD-11 guidelines and 3 domains of functional impairment (i.e., IGDS9-SF, SCI-IGD, CVAT2.0, and DIA).

Overall, there were 8 tools that provided total coverage of both the DSM-5 and ICD-11 criteria (i.e., Petry IGD, IGDS9-SF, PIE-9, IGDT-10, SCI-IGD, CVAT2.0, IGUESS, and DIA). Of this list, only 3 tools referred to 4 types of functional impairment (i.e., Petry IGD, PIE-9, and IGDT-10).

3.3. The evidence base underlying GD tools

Table 3 summarizes the empirical research literature (n=320 studies, excluding duplicates) that have employed at least one of the 32 GD tools (NB: Supplementary material 1 presents a list of 328 references, which includes 8 duplicated references (i.e., refs 16-93, 19-95, 39-83, 48-177, 82-84, 88-129, 195-212, 196-213; this duplication occurred during coding phases to disambiguate studies with more extensive information on multiple GD tools and/or study parts (i.e., Study 1, Study 2) to ensure that these data were cross-checked properly prior to entering into master spreadsheets.

This evaluation provides an overview of a substantial proportion of the empirical GD literature (NB: it was beyond the scope of this review to estimate the overall size of the total empirical GD literature; i.e., including other studies that employ an unnamed or uncited instrument). Table 3 indicates that the GD field has at least 30 longitudinal studies, 71 prevalence studies (i.e., studies that employ nationally representative samples), and 14 intervention studies. Overall, the most frequently used tools, irrespective of study type, were the YIAT (n=62), GAS-7 (n=45), and IGDS9-SF (n=24). The most common tool used in prevalence studies was the GAS-7 (n=18), which was also the most common tool for longitudinal studies (n=12). The YIAT has been used in 39 intervention studies and reports on treatment-seekers at baseline, with the majority (n=37) conducted in China and South Korea.

The YIAT and YDQ are the most cited tools, which may be due to their prescience (i.e., these tools are presented in papers by the late Kimberly Young, a pioneer of the field of internet addiction; Brand & Potenza, 2019) and their earlier publication relative to other tools in the review. The Petry IGD tool is highly cited (n=447) relative to its use in studies (n=16), which may be due to being cited for its reference to international "consensus" on gaming disorder (i.e., many of its citations include

commentary and debate papers on the topic of consensus on DSM-5 IGD). Similarly, the A-EQ is highly cited (n=798) relative to its use in studies (n=22), which may be attributed to the paper's wider discussion of the importance of distinguishing high (but non-problematic) engagement from addiction (Billieux, Flayelle; Rumpf, & Stein, 2019; Charlton & Danforth, 2007). Only two GD-specific tools (i.e., the AICA-Sgaming and VAT) have been used in both a national prevalence study and an intervention study. Overall, the YIAT, YDQ, and GAS-7 have been the most widely used and cited GD tools; however, none of these tools provide total coverage of the DSM-5 or ICD-11 criteria.

3.4. Samples and recruitment strategies

Table 4 presents a summary of the age composition of, and recruitment strategies for, samples in empirical studies of most GD tools (n=24). For this table, tools with only 1 study (n=8) were excluded due to insufficient cases to yield meaningful proportions. Overall, the total N for all identified studies was 462,249 participants, of whom 56% were aged 18 years or older. The GAS-7 has been administered to more participants (n=94,389) than any other tool, including the greatest number of participants recruited using non-convenience sampling (n=55,518). Studies employing the YIAT and YDQ also reported relatively high figures for total N (n=49,500 and n=30,916, respectively) and for participants recruited by non-convenience sampling (n=31,592 and n=29,810, respectively). The AICA-Sgaming was noteworthy for its relatively large evidence base (n=18, including 10 studies with clinical samples), fourth highest total N (36,306), and the highest percentage of non-convenience sampling (85.7%) among the overall most frequently used tools (i.e., ranked 4th after the GAS-7, YIAT, and CIUS-14). Some tools (i.e., IGDS9-SF, GAS-21, and Petry IGD) have been administered to a relatively large number of participants, but these studies are predominantly based (i.e., >95% of total N) on convenience samples (e.g., online self-selected, non-representative samples). Only the IGUESS tool has total coverage of the DSM-5 and ICD-11 criteria along with an evidence base composed of samples obtained by non-convenience sampling; all other DSM/ICD-compatible tools are based on evidence with 95% convenience sampling.

As a supplementary analysis, all sample figures were examined according to the DSM-5 criteria covered by the tools they had been administered in their respective studies (i.e., based on Table 2 results). The objective of this analysis was to determine the extent to which study participants had been able to respond (e.g., affirmatively, negatively) to specific items referring to GD symptoms. The total N in each study corresponding to each tool was assigned to each of the relevant DSM-5 symptoms covered by each tool. For example, the GAS-7 had been administered to 94,389 participants, and therefore the value 94,389 was assigned to its measurement of symptoms of preoccupation, withdrawal, tolerance, etc. This process yielded an estimated total number of participants with *the possibility of responding* to any given GD symptom (e.g., assuming the item on the survey was completed).

This evaluation showed that, overall, items referring to preoccupation, withdrawal, unsuccessful attempts, and harm had been administered to >96% of participants across the 320 studies. However, items measuring continued use despite harm (i.e., the DSM-5 and ICD-11 criterion) had been administered to only 11.9% of participants. Similarly, items measuring tolerance (67.6%), loss of interests (58.5%), and deception (48.3%) had been administered to relatively fewer participants across studies. When selecting only those studies with non-convenience samples, these figures were relatively similar across these symptoms (i.e., less than 4% difference), except for the item on

continued use despite harm, which reduced from 11.9% to 1.9%. Only 5053 participants out of the 268,081 participants in non-convenience samples had been administered a tool with an item referring to continued use despite harm, indicating that this symptom was underrepresented in the literature.

3.5. Psychometric properties of tools

Table 5 presents the first of two summaries of the psychometric properties of the 32 tools based on their 320 empirical studies. The table summarizes the research evidence on test dimensionality, reliability, and test refinement. Most GD tools (n=27) have been examined using factor analytic techniques (or principal component analysis), but there are inconsistencies across tools in terms of factor analytic (FA) approach. Only 9 tools (i.e., POGU, POGQ, IGD-20, Petry IGD, Lemmens IGD-9 and IGD-27, YDQ, and CIUS-214) have been examined by exploratory followed by confirmatory factor analysis (i.e., EFA and CFA conducted in independent subsamples) within the same study. The majority (n=21) of tools have been subjected to CFA only, with most studies providing the *a priori* reasoning that GD is a unidimensional construct. The majority of FA studies (n=47) reported that the GD construct is unidimensional (i.e., commonly referring to the GD construct composed of nine DSM-5 criteria, with the caveat that most tools actually assess fewer than 9 criteria; see Table 2). There was minimal empirical support for other dimensional structures (e.g., 2-factor solutions; n=3 studies). Some tools demonstrated mixed and/or weak support for their factor structure (e.g., VAT, IGD-20, sIATgaming, YIAT-20, and CIUS-14).

Internal consistency was generally high across the 31 tools (i.e., this was not reported for SCI-IGD). With the exception of the BAM-VG, GAIA, VASC, CIUS-5, CIUS-8 and SSBA tools which had only 1 study reporting on internal consistency, each tool had at least two independent studies reporting Cronbach's alpha values of at least 0.80. Considerably fewer tools (n=7; 9 studies) had examined test-retest reliability; in 8 of these 9 studies, the observed values from a 14-day or 30-day retest were satisfactory. The BAM-VG had a .73 test-retest reliability over a period of 90 days. There were 8 tools (i.e., GAS-7, POGQ, POG-SF, VAT, IGDS9-SF, IGDT-10, CIUS-14, CIUS-8) that had been evaluated by test-refinement analyses (e.g., IRT, Rasch analysis). Generally, these analyses provided support for the model fit and measurement invariance of each respective test. However, the IGDS9-SF reported mixed results for measurement invariance across cultural groups; it bears noting that these data were based largely (>95%) on convenience sampling, which may have affected analyses.

Table 6 summarizes the research evidence on test validity, relationship to impairment, and clinical use of each tool. Convergent validity has often been operationalized in the GD literature as the bivariate association between gaming behavior (i.e., hours per week spent gaming) and total score on a GD tool. The research team discussed and agreed that habitual gaming for long periods (e.g., 6 to 8 hours per day, or longer) was typical in the context of GD, and acknowledged that this behavioral pattern may fluctuate and that the condition was often episodic, but did not consider this association to be generally defining. Gaming may occur for some individuals as a regular and relatively frequent activity without reported associated major negative consequences, as described in previous studies (Király, Tóth, Urbán, Demetrovics, & Maraz, 2017; Triberti et al., 2018). King et al.'s (2013) review

reported the GD-gaming time association for convergent validity; therefore, it is reproduced with caution in this review. Overall, there were varied results on the association between gaming time and total GD tool score, with reported values largely ranging from .2 and .4 (i.e., small to moderate effect; Cohen, 1992).

Criterion validity was evaluated by examining the association between scores on each GD tool and other similar or closely related tools (e.g., measures of gaming-related craving or maladaptive gaming-related cognitions). Most tools (n=28) have been examined in relation to other GD tools, particularly the YIAT (n=8 tool comparisons). The GAS-21, VAT, and the IGDS9-SF have reported the most consistently high correlations with other GD tools. Aside from the YIAT's consistent convergent validity results across 11 studies, the IGDS9-SF was noteworthy for having 5 studies employing 4 different GD tools (i.e., GAS-7, IGD-20, Lemmens IGD-9, and YIAT) that reported associations exceeding r=.70.

Only 6 tools (n=9 studies) have been examined in conjunction with standardized measures of functional impairment or quality of life. The PIE-9 was the only tool that has been evaluated using the recommended standard disability/impairment assessment (i.e., the World Health Organization-Disability Assessment Schedule [WHO-DAS]). The IGD-20 has been examined in conjunction with the DSM global clinician rating scale (i.e., General Assessment of Functioning [GAF]). Other tools (i.e., the A-EQ, GAS-7, Lemmens IGD-9, and YIAT) have been evaluated using standardized quality-of-life measures.

Eleven tools have been employed in studies involving a clinical interview. Notably, the AICA-Sgaming (n=11) and the YIAT (n=31) have been used most frequently. The AICA-Sgaming and Petry IGD are the two most commonly used gaming-specific tools within studies using clinical interviews. In relation to GD tools measuring treatment outcomes, only the YIAT (n=9 studies) and YDQ (n=2 studies) have been used in more than one published study. Four other GD tools (i.e., the POGU, AICA-Sgaming, VAT, and IGD-20) have been employed in one study only.

3.6 Quick reference guide to GD tools

Table 7 presents a synthesis of the main areas of evaluation of the 32 tools. This table was designed to provide a 'quick guide' for researchers and clinicians, identifying tools with specific supporting evidence or use in particular study types. In this way, the table aims to provide a concise overview of the relative strengths and weaknesses for each tool, with the caveats that this table presents: (1) an unweighted representation of data underlying each of the criteria, meaning that the table does not differentiate between tools that meet certain score thresholds (e.g., 2 studies) and those that greatly exceed this basic threshold (e.g., not all tools that score 2 on 'prevalence data' should be considered equivalent); (2) total scores in the rightmost column should not necessarily be considered an overall indicator of tool quality (i.e., higher scores indicate the presence of research evidence in more areas, not higher tool quality); and, (3) higher scores in some areas may be undermined by weaknesses in other areas (e.g., tools with poorly sampled studies are less likely to be valid).

Notwithstanding these limitations, the GD tools that fulfilled the most scoring criteria were the GAS-7, IGDS9-SF, IGDT-10, YDQ and Lemmens IGD-9. The GAS-7 is a much older gaming-specific tool and precedes the DSM-5. Although the GAS-7 is still used frequently in research,

particularly in prevalence studies of European young people, this tool has not yet been used in intervention studies. Only the IGDS9-SF and IGDT-10 have DSM-5 and ICD-11 coverage, and only the YDQ has been used in a study involving clinical interviewing or an intervention. The IGDS9-SF and IGDT-10 had similar profiles in this evaluation, including basic length and scoring, and a comparably sized evidence base (i.e., in terms of total N) predominantly based on convenience samples. Overall, there was a mixed picture of the evidence on GD tools, with several tools with relatively higher evidential support in distinct areas (e.g., GAS-7, Lemmens IGD-9, AICA-Sgaming, and IGDT-10), but there was no markedly superior tool with distinct practical and/or psychometric advantages.

4. Discussion

The present review aimed to systematically evaluate all available instruments for GD. This work was intended to extend a previous major review of GD tools by King et al. (2013), which closely preceded the inclusion of GD as a condition for further study in the DSM-5. The 2013 review highlighted conceptual inconsistencies across GD tools, as well as gaps in empirical evidence underlying each measures' psychometric properties. Seven years have passed since the preparation of King et al.'s (2013) report, and with the official inclusion of GD as a diagnostic category in the ICD-11 in May 2019, it was considered timely to re-evaluate the state-of-the-art in GD screening and assessment. Overall, this evaluation has found that the GD field has greatly expanded in overall size and its array of GD-specific instrumentation, particularly since 2013, with at least 2 new tools, on average, published in each subsequent year. The field has also continued to employ several internet-use-specific tools (e.g., the YIAT, created in 1998) to screen for gaming-related problems, particularly in East Asia. Overall, no single tool emerged from this evaluation as the clearly optimal choice. However, there were some relatively stronger tools (i.e., the AICA-Sgaming, GAS-7, IGDT-10, IGDS9-SF, and Lemmens IGD-9), identified on the basis of conceptual and/or practical considerations and greater volume of evidential support for their psychometric properties.

The GD field is growing rapidly on a global level. This systematic review has identified 320 empirical studies that have employed a combined total of 32 GD tools, with these studies conducted primarily throughout Europe and East Asia. However, there still appears to be some uncertainty or lack of agreement among GD researchers concerning optimal approaches to screening and

assessment, as indicated by the continuing creation of new tools that vary in scope and content. This review has identified inconsistencies in symptom coverage across 32 tools. Screening tools do not necessarily have to measure all criteria or guidelines for any given condition, including GD, in order to be effective. The objective is usually to capture the essential elements of behaviors in a brief format. This review shows that most GD tools tend to converge on the importance of screening for impaired control over gaming and gaming behavior that jeopardizes a significant relationship, or school or work opportunity. This means, however, that there are some criteria or guidelines for GD that tend to be excluded from tools. Notably, the criterion referring to continued use despite harm (which is included in the DSM-5 and ICD-11 criteria) has appeared in only 9 out of 32 tools. This review found that an estimated 88.1% of participants across 320 studies have not been administered a survey item that captures this particular symptom. The 8 tools that provide total coverage of both the DSM-5 and ICD-11 criteria (i.e., Petry IGD, IGDS9-SF, PIE-9, IGDT-10, SCI-IGD, CVAT2.0, IGUESS, and DIA) have been used in a combined total of only 5 nationally representative prevalence studies, or 7% of the prevalence study literature. These observations raise the issue of adequate representation of GD symptomatology in research, and whether these observed gaps in measurement should be factored into current estimates of incidence and prevalence.

Sixteen GD-specific tools have been created since the recognition of IGD in the DSM-5 in 2013. However, many research teams have instead opted to use other tools that precede the DSM-5. Notably, the YIAT and YDQ measures have been used in numerous studies conducted within China and South Korea, particularly those involving clinical interviews to determine eligibility for neuroimaging evaluation and/or interventions. Similarly, research teams in Europe have often employed the GAS-7 for large-scale prevalence and cohort studies. The GD field therefore appears to be shaped by two main types of researchers: (1) those who continue to use older (i.e., pre-DSM-5) tools (i.e., the GAS-7, YIAT, and YDQ) despite the availability of new tools and guidelines for GD; and, (2) those who develop and attempt to validate their own tools which are often conceptually and practically similar, i.e., a tool of between 10 and 20 items that measures a unidimensional addiction construct derived from DSM-5 criteria. Further, regarding (2), the majority of new tools tend to be

psychometrically evaluated exclusively by the researchers who created them. The IGDT-10 and AICA-Sgaming are the only post-DSM-5 tools that have been employed in a study designed to validate psychometric properties by an independent research team. This suggests that there is an isolationist quality to the field's collective efforts to study GD. Many teams appear to be operating in research silos rather than working collaboratively to develop a unified evidence base around a smaller, more manageable subset of measures. Reaching a consensus about the use of specific psychometrically validated screening tools in studies worldwide would not only optimize prevalence estimates, but may also be helpful for studies addressing psychological mechanisms underlying GD and for testing hypotheses on these processes as suggested in theoretical models (e.g., Brand et al., 2019; Dong & Potenza, 2014; Wei et al., 2017). For comparison, the field of gambling disorder, which like GD is a disorder due to addictive behaviors recognized by the ICD-11 (albeit with a more extensive academic history than GD), has coalesced around 8 main tools used between 2000 and 2015[1] (see Calado & Griffiths, 2015).

4.1 GD tools with greater evidential support

This review identified five GD tools that have a relatively greater volume of evidential support for basic psychometric properties compared to other tools. These tools are the GAS-7, IGDS9-SF, IGDT-10, YDQ, and Lemmens IGD-9. However, no one tool appears clearly superior because many of the tools have similar strengths (e.g., the GAS-7 and Lemmens IGD-9 had a comparable evaluation profile), and each tool's limitations were offset by other positive attributes. Some tools are untested in some contexts (e.g., treatment) or have not yet been evaluated psychometrically in some ways (e.g., item response theory). It may be anticipated that future research will address some of these gaps. On the other hand, given that research teams have tended to specialize in particular research areas (e.g., epidemiological, neuroimaging, treatment) and teams have often favored their own tools, some tools appear unlikely to be evaluated in certain types of research. The AICA-Sgaming, for example, has

---

[1] These 8 tools are the South Oaks Gambling Screen (SOGS), Problem Gambling Severity Index (PGSI), DSM-IV criteria for pathological gambling, Diagnostic Interview Schedule (DIS) for pathological gambling, Diagnostic Interview for Gambling Severity (DIGS), National Opinion Research Center DSM Screen for Gambling Problems (NODS), Gamblers Anonymous Twenty Questions (GA20), and the Lie/Bet scale.

been used in at least 10 studies of treatment-seekers, whereas 17 other tools have not been used in any research involving clinical samples. This situation may eventually lead to divergent streams of evidence for specific tools, such that some tools become the standard for some study types but not others, thereby complicating future tasks of synthesizing evidence across a broad literature.

The GAS-7 was the most frequently positively rated tool (N.B., not to be conflated with psychometric superiority) due to its multiple positive features and large evidence base. Its research base included numerous datasets from prevalence studies throughout Europe, good criterion validity and reliability (internal, test-retest), and satisfactory performance on test refinement analyses (e.g., measurement invariance). However, the GAS-7 has not been used clinically, and therefore its utility as an outcome measure in treatment, or sensitivity to treatment-related changes, has not been investigated. In addition, the GAS-7 has incomplete coverage of DSM-5 and ICD-11 criteria, which may not be essential for screening purposes but remains noteworthy as the field turns its attention to locating measurement approaches consistent with ICD-11 classification. This may include research initiatives that involve developing new screening tools for behaviors that appear to have overlapping features with gaming (e.g., online social media use) by adapting items from existing GD tools[2].

Of the 5 tools with the broadest empirical support as identified in Table 7, the IGDS9-SF and IGDT-10 were the only tools that provide total coverage of the DSM-5 and ICD-11 criteria. The IGDS9-SF and IGDT-10 had similar profiles in this evaluation, including number of items and scoring approach, and a comparably sized evidence base (i.e., in terms of total N). Both tools were limited by their study samples; most studies (i.e., 23 out of 24) that employed the IGDS9-SF involved convenience samples, and the IGDT-10 has exclusively been used in convenience samples (i.e., all 7 studies). The IGDT-10 is noteworthy, however, for its numerous language translations and use in relatively more countries, notably China and Japan (N.B., only 2 of the 32 tools have been translated into Japanese). Therefore, the IGDT-10 seems well positioned to bridge the research divide between Western and Eastern countries that is common in the GD literature. With regard to Eastern research, the YIAT and YDQ tools were most commonly used in Chinese studies (i.e., those published in

---

[2] See recent papers on 'scope creep' (Haslam, 2016) and confirmatory bias in connection to the study of behavioral addictions (Billieux et al., 2015).

English), which stands in contrast to most research teams across Western countries that now rarely employ these two tools.

4.2 Implications for debates on GD evidence

The present review provides insights that should contribute to ongoing debate on the general quality and attributes of the GD literature. For example, some critics have asserted that the GD evidence base is flawed due to its reliance on convenience samples. A recent 37-author debate paper by van Rooij et al. (2018), entitled "*A weak scientific basis for gaming disorder*" argued against the inclusion of GD in ICD-11 on the basis that scientific standards had not been met. Among other issues raised, the authors argued that sampling approaches were inadequate because studies recruited "healthy high-school/college students or non-representative online samples recruited from Internet gaming forums" (p.3). The authors also referred to the example of a Singaporean dataset, published in 2011 (Gentile, 2011) and subsequently used in multiple publications without cross-attribution of the data, to support their contention that the literature suffered from "poor methodological choices that undermine our confidence in the findings" (van Rooij et al., 2018, p.4). While it is necessary that authors cite specific examples for academic arguments, it is important that such examples do not form the basis for unwarranted generalizations that may then be misconstrued as scientific consensus. The present review's findings did not support this particular criticism regarding sampling. Based on 320 GD studies of which the majority report independent datasets, there was actually a *slight majority* of participants (N=268,081 or 58% of all participants) recruited using non-convenience sampling methods (e.g., nationally representative studies, stratified sampling by age, region, urbanicity, and treatment-seekers). It is inaccurate to conclude, therefore, that the GD evidence as a whole is fundamentally flawed or "weak" as a consequence of its recruitment strategies and sample sizes. Indeed, the present review may provide a useful resource for other matters of debate concerning the size and quality of GD evidence.

4.3 Future research directions

This review suggests some potential future research avenues to improve GD assessment. As recommended by King et al. (2013), there is a continued need for high-quality epidemiological and

intervention studies, including within these studies with a focus on sensitivity/specificity estimates. Studies of gaming behavior should include consistent measures of comorbidity (e.g., to address questions regarding the presence of other mental disorders, such as depression, anxiety, attention-deficit/hyperactivity or other factors such as past trauma that may affect risk of GD). Similarly, future work should consider not only addictive aspects, but also other relevant perspectives and concepts related to understanding repetitive behaviors, notably impulsivity and compulsivity, in further developing and refining instruments and diagnostic criteria. Non-problematic gaming habits, as well as diagnostically subthreshold entities such as 'hazardous gaming' (QE22) in the ICD-11, deserve closer attention (Potenza, 2018). The focus of most studies of GD has been on the harmful consequences of gaming, without taking into account the potential benefits of gaming activities for some individuals. GD symptoms and negative consequences of gaming should be weighed against reported benefits of gaming, particularly at lower levels of problematic gaming (e.g., meeting between 1 and 4 DSM-5 criteria). This may determine whether some individuals classified within 'low risk' categories might in fact report that gaming has a net benefit on their quality of life and psychological wellbeing or whether this sub-diagnostic level is associated with more mental health concerns, as is typically the case in gambling disorder (Desai & Potenza, 2018).

Another avenue for future research is the use of player data in combination with GD tools and related measures. The field has often relied on self-report approaches to validate tools, which has unavoidable limitations (e.g., biased recall, denial/defensiveness, lack of insight). Conventional survey and interview approaches may be supplemented by player data to provide an objective historical account of gaming behavior; i.e., to describe or corroborate patterns of behavior that may otherwise be difficult to recall. Such data may be acquired by using an app-like or similar monitoring device or software. Considerations regarding how to work together with groups from the gaming industry warrant transparent discussion in order to help ensure scientific integrity in academic/industry collaborations (Griffiths & Pontes, 2019; King & Delfabbro, 2019). Another area for future research concerns the evolving technological nature of modern online video games, particularly the monetization of in-game content (e.g., in-game purchasing, microtransactions, and 'loot boxes'; see King et al., 2019; Zendle & Cairns, 2018). Problematic gaming that involves interactions with monetized content may be more financially involved and share features in common

with gambling disorder (e.g., spending more than one can afford, borrowing or stealing money) (King & Delfabbro, 2018; King et al., 2019). GD tools may need to reflect some of these structural elements in gaming activities, such as additional questions to examine different behaviors and consequences related to different types of games and modes of access (e.g., smartphones, virtual reality) (King, Koster, & Billieux, 2019). The measurement of more in-depth player and gaming information (e.g., game types) is beyond the scope of screening approaches, but is suited to a semi-structured diagnostic interview for GD (i.e., akin to the Structured Clinical Interview for DSM or Mini International Neuropsychiatric Interview), which could be developed and used internationally.

4.4. Limitations of the review

The present review has several limitations that should be acknowledged. First, this review was based on English-language studies only, which excluded a significant proportion (i.e., potentially, the majority) of the East Asian literature (i.e., studies in Chinese, Korean, and Japanese, in particular). This review also did not include data from studies published in German, Dutch, and French. Although researchers working in these countries often publish their work in English journals, the potential omission of relevant data is likely to have affected the evaluation of tools originating from these regions, e.g., the AICA-Sgaming, CSAS, GAS-7, and sIAT-gaming. Similarly, in some countries, there is a substantial grey literature (e.g., government-led health surveys that include standard GD questions), which was not included in this review. This review also did not include studies that employed "adapted" DSM 5 criteria, which would not have affected the main evaluation of tools, but should be taken into account when considering this review's observations of the broader GD literature (e.g., total number of prevalence studies). The review framework aimed to be more comprehensive than any previous review but there were still some gaps. This review did not consider, for example, the sensitivity/specificity of tools, because: (1) this information was very rarely reported; and, (2) the external standard was not always clear in relevant studies. Other areas that were not evaluated were predictive and divergent validity, due to inconsistencies in reporting that made it difficult to extract these data. This review takes into consideration the current conceptualizations of GD and IGD, which may be revised as more evidence accumulates. Finally, the review framework itself was limited by the

fact that many of the criteria were inter-related and often affected by other considerations (e.g., sampling method, sample size).

4.5 Conclusions

The inclusion of GD in the ICD-11 was a significant milestone for the field. The GD diagnosis is likely to stimulate new research investigations on a global level, in important areas of epidemiology, neurobiology, treatment, prevention and public health. The present review aimed to inform the next era of research by providing a comprehensive evaluation of all available English-language GD tools, including a critical appraisal of their associated empirical evidence. The framework employed in this review may be useful for scale evaluation in other areas. Overall, this evidence was found to be mixed, with no clearly optimal tool among 32 tools used across Western and Eastern countries. While the research base has grown rapidly and largely improved its methodologies, the field is hindered by the overproduction of conceptually similar tools which have divided research efforts and created uncertainty among researchers. Despite the abundance of new instrumentation, some tools have relatively greater evidential support for their psychometric properties, including the GAS-7, IGDS9-SF, IGDT-10, and Lemmens IGD-9. Given that most new tools were developed following the inclusion of the provisional DSM-5 criteria, it seems likely that researchers will again "rush to market" to develop new tools that purportedly measure the new ICD-11 GD classification. For the field to prosper and attain greater legitimacy in the field of addiction studies, a more unified approach to measurement is important. Isolated research that creates a multiplicity of tools generates an incohesive and less convincing evidence base. The development of a gold standard tool, following past examples of screening for use of addictive substances (e.g., Saunders et al., 1993; ASSIST; WHO ASSIST Working Group, 2002), would be invaluable for steering this nascent field toward achieving valid identification of gaming-related harms, and developing more effective intervention strategies for those in need.

Role of Funding Sources

Contributors

DLK designed the review and conducted literature searches and provided summaries of the identified content, with substantial assistance from JB, NC, and PHD. All authors contributed to the design and interpretation of the analysis. Authors DLK, SC, JB, KM, MNP, HJR, JS, VS, ZD, MB and SH developed the summary table (Table 7). DLK wrote the first draft of the manuscript, and all authors contributed to and approved the final manuscript.

Conflict of Interest

Dr Chamberlain consults for Promentis and Ieso Digital Health, and receives a stipend for editorial work at Comprehensive Psychiatry & Neuroscience and Biobehavioral Reviews. The other authors report no potential conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES[3]

Allison, S. E., Von Wahlde, ?, Shockley, T., & Gabbard, G. O. (2006). The development of the self in the era of the internet and role-playing fantasy games. *American Journal of Psychiatry, 163*, 381-385.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.) (DSM-5)*. Washington, DC: Author.

APA Publications and Communications Board Working Group on Journal Article Reporting Standards. (2008). Reporting standards for research in psychology: What do we need them? What might they be? *American Psychologist, 63,* 839-851.

Baggio, S., Dupuis, M., Studer, J., Spilka, S., Daeppen, J. B., Simon, O., ... & Gmel, G. (2016). Reframing video gaming and internet use addiction: Empirical cross-national comparison of heavy use over time and addiction scales among young users. *Addiction, 111,* 513-522.

---

[3] See Supplementary Material for a complete list of references for the 32 GD tools and 320 empirical studies.

Beranuy, M., Carbonell, X., & Griffiths, M. D. (2013). A qualitative analysis of online gaming addicts in treatment. *International Journal of Mental Health and Addiction, 11,* 149-161.

Billieux, J., Flayelle, M., Rumpf, H.-J., & Stein, D.J. (2019). High involvement versus pathological involvement in video games: A crucial distinction for ensuring the validity and utility of gaming disorder. *Current Addiction Reports, 6*, 323-330

Billieux, J., Schimmenti, A., Khazaal, Y., Maurage, P., & Heeren, A. (2015). Are we overpathologizing everyday life? A tenable blueprint for behavioral addiction research. *Journal of Behavioral Addictions*, *4*(3), 119-123.

Brand, M. & Potenza, M. N. (2019). In memory of Dr. Kimberly S. Young: The story of a pioneer. *Journal of Behavioral Addictions*, *8*(1), 1-3.

Brand, M., Wegmann, E., Stark, R., Müller, A., Wölfling, K., Robbins, T. W., & Potenza, M. N. (2019). The Interaction of Person-Affect-Cognition-Execution (I-PACE) model for addictive behaviors: Update, generalization to addictive behaviors beyond Internet-use disorders, and specification of the process character of addictive behaviors. *Neuroscience and Biobehavioral Reviews, 104,* 1-10.

Brooks, G. A. & Clark, L. (2019). Associations between loot box use, problematic gaming and gambling, and gambling-related cognitions. *Addictive Behaviors, 96,* 26-34.

Calado, F. & Griffiths, M. D. (2016). Problem gambling worldwide: An update and systematic review of empirical research (2000–2015). *Journal of Behavioral Addictions*, *5*, 592-613.

Charlton, J.P, & Danforth, I.D.W (2007). Distinguishing addiction and high engagement in the context of online game playing. *Computers in Human Behaviors, 23*, 1531–48

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardised assessment instruments in psychology. *Psychological Assessment, 6,* 284-290.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika. 16,* 297-334.

Desai, R. A. & Potenza, M. N. (2008). Gender differences in the associations between past-year gambling problems and psychiatric disorders. *Social Psychiatry and Psychiatric Epidemiology, 43,* 173-183.

Dong, G. & Potenza, M. N. (2014). A cognitive-behavioral model of Internet gaming disorder: Theoretical underpinnings and clinical implications. *Journal of Psychiatric Research, 58,* 7-11.

Dong, G., Wang, L., Du, X., & Potenza, M. N. (2017). Gaming increases craving to gaming-related stimuli in individuals with Internet gaming disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 2,* 404-412.

Feng, W., Ramo, D., Chan, S., & Bourgeois, J. (2017). Internet gaming disorder: Trends in prevalence 1998–2016. *Addictive Behaviors, 75,* 17-24.

Fisher, S. (1994). Identifying video game addiction in children and adolescents. *Addictive Behaviors, 19,* 545-553.

Gentile, D. A., Choo, H., Liau, A., Sim, T., Li, D., Fung, D., & Khoo, A. (2011). Pathological video game use among youths: A two-year longitudinal study. *Pediatrics*, *127*, e319-e329.

Granic, I., Lobel, A., & Engels, R. C. (2014). The benefits of playing video games. *American Psychologist, 69,* 66-78.

Griffiths, M. D. (2010). The role of context in online gaming excess and addiction: Some case study evidence. *International Journal of Mental Health and Addiction, 8*, 119-125.

Griffiths, M. D. & Hunt, N. (1998). Dependence on computer games by adolescents. *Psychological Reports, 82*, 475-480.

Griffiths, M. D. & Pontes, H. M. (2019). The future of gaming disorder research and player protection: what role should the video gaming industry and researchers play? *International Journal of Mental Health and Addiction.* DOI: 10.1007/s11469-019-00110-4

Griffiths, M., King, D. L., & Demetrovics, Z. (2014). DSM-5 internet gaming disorder needs a unified approach to assessment. *Neuropsychiatry, 4,* 1-4.

Groth-Marnat, G. (2009). *Handbook of Psychological Assessment (5th Edition).* New York: John Wiley & Sons.

Groves, C., Gentile, D., Tapscott, R., & Lynch, P. (2015). Testing the predictive validity and construct of pathological video game use. *Behavioral Sciences, 5,* 602-625.

Haagsma, M. C., Pieterse, M. E., & Peters, O. (2012). The prevalence of problematic video gamers in the Netherlands. *Cyberpsychology, Behavior, and Social Networking, 15,* 162-168.

Haslam, N. (2016). Concept creep: Psychology's expanding concepts of harm and pathology. *Psychological Inquiry*, *27*, 1-17.

Higuchi, S., Nakayama, H., Mihara, S., Maezono, M., Kitayuguchi, T., & Hashimoto, T. (2017). Inclusion of gaming disorder criteria in ICD-11: A clinical perspective in favor: Commentary on: Scholars' open debate paper on the World Health Organization ICD-11 Gaming Disorder proposal (Aarseth et al.). *Journal of Behavioral Addictions, 6,* 293-295.

Jap, T., Tiatri, S., Jaya, E. S., & Suteja, M. S. (2013). The development of Indonesian online game addiction questionnaire. *PloS One, 8,* e61098.

Kaptsis, D., King, D. L., Delfabbro, P. H., & Gradisar, M. (2016). Withdrawal symptoms in Internet gaming disorder: A systematic review. *Clinical Psychology Review, 43,* 58-66.

Khazaal, Y., Chatton, A., Rothen, S., Achab, S., Thorens, G., Zullino, D., & Gmel, G. (2016). Psychometric properties of the 7-item game addiction scale among French and German speaking adults. *BMC Psychiatry, 16*, 132.

Király, O., Tóth, D., Urbán, R., Demetrovics, Z., & Maraz, A. (2017). Intense video gaming is not essentially problematic. *Psychology of Addictive Behaviors*, *31*, 807-817.

King, D. L. & Delfabbro, P. H. (2014). The cognitive psychology of Internet gaming disorder. *Clinical Psychology Review, 34,* 298-308.

King, D. L. & Delfabbro, P. H. (2018). Predatory monetization features in video games (e.g., 'loot boxes') and Internet gaming disorder. *Addiction, 113,* 1967-1969.

King, D. L. & Gaming Industry Response Consortium. (2018). Comment on the global gaming industry's statement on ICD-11 gaming disorder: A corporate strategy to disregard harm and deflect social responsibility? *Addiction, 113*, 2145-2146.

King, D. L. & Delfabbro, P. H. (2019). *Internet gaming disorder: Theory, assessment, treatment, and prevention*. Cambridge, MA: Elsevier Academic Press.

King, D. L. & Delfabbro, P. H. (2019). Video game monetization (e.g., 'loot boxes'): A blueprint for practical social responsibility measures. *International Journal of Mental Health and Addiction, 17,* 166-179.

King, D. L., Koster, E., & Billieux, J. (2019). Study what makes games addictive. *Nature, 573,* 346.

King, D. L., Delfabbro, P. H., Zwaans, T., & Kaptsis, D. (2013). Clinical features and axis I comorbidity of Australian adolescent pathological Internet and video-game users. *Australian and New Zealand Journal of Psychiatry, 47,* 1058-1067.

King, D. L., Haagsma, M. C., Delfabbro, P. H., Gradisar, M., & Griffiths, M. D. (2013). Toward a consensus definition of pathological video gaming: A systematic review of psychometric assessment tools. *Clinical Psychology Review, 33,* 331-342.

King, D. L., Delfabbro, P. H., Gainsbury, S. M., Dreier, M., Greer, N., & Billieux, J. (2019). Unfair play? Video games as exploitative monetized services: An examination of game patents from a consumer protection perspective. *Computers in Human Behavior, 101,* 131-143.

King, D. L., Delfabbro, P. H., Deleuze, J., Perales, J. C., Király, O., Krossbakken, E. & Billieux, J. (2019). Maladaptive player-game relationships in problematic gaming and gaming disorder: A systematic review. *Clinical Psychology Review, 73,* 101777.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cut-off criteria: What did they really say? *Organizational Research Methods*, *9*, 202-220.

Lemenager, T., Gwodz, A., Richter, A., Reinhard, I., Kaemmerer, N., Sell, M., & Mann, K. (2013). Self-concept deficits in massively multiplayer online role-playing games addiction. *European Addiction Research, 19,* 227-234.

Lortie, C. L. & Guitton, M. J. (2013). Internet addiction assessment tools: Dimensional structure and methodological status. *Addiction, 108*, 1207-1216.

Mak, K. K., Lai, C. M., Watanabe, H., Kim, D. I., Bahar, N., Ramos, M., ... & Cheng, C. (2014). Epidemiology of internet behaviors and addiction among adolescents in six Asian countries. *Cyberpsychology, Behavior, and Social Networking, 17,* 720-728.

Marino, C. & Spada, M. M. (2017). Dysfunctional cognitions in online gaming and internet gaming disorder: A narrative review and new classification. *Current Addiction Reports, 4,* 308-316.

Meerkerk, G. J., Van Den Eijnden, R. J., Vermulst, A. A., & Garretsen, H. F. (2009). The compulsive internet use scale (CIUS): Some psychometric properties. *Cyberpsychology & Behavior, 12*, 1-6.

Meng, Y., Deng, W., Wang, H., Guo, W., & Li, T. (2015). The prefrontal dysfunction in individuals with Internet gaming disorder: A meta- analysis of functional magnetic resonance imaging studies. *Addiction Biology, 20,* 799-808.

Mentzoni, R. A., Brunborg, G. S., Molde, H., Myrseth, H., Skouverøe, K. J. M., Hetland, J., & Pallesen, S. (2011). Problematic video game use: Estimated prevalence and associations with mental and physical health. *Cyberpsychology, Behavior, and Social Networking, 14,* 591-596.

Mihara, S. & Higuchi, S. (2017). Cross- sectional and longitudinal epidemiological studies of Internet gaming disorder: A systematic review of the literature. *Psychiatry and Clinical Neurosciences, 71,* 425-444.

Moher, D., Tsertsvadze, A., Tricco, A., Eccles, M., Grimshaw J., Sampson, M., & Barrowman N. (2008). When and how to update systematic reviews. *Cochrane Database of Systematic Reviews*, Issue 1. Art. No.: MR000023.

Müller, K. W., Janikian, M., Dreier, M., Wölfling, K., Beutel, M. E., Tzavara, C., ... & Tsitsika, A. (2015). Regular gaming behavior and Internet gaming disorder in European adolescents: Results from a cross-national representative survey of prevalence, predictors, and psychopathological correlates. *European Child & Adolescent Psychiatry, 24*, 565-574.

Peng, W. & Liu, M. (2010). Online gaming dependency: A preliminary study in China. *Cyberpsychology, Behavior, and Social Networking, 13*, 329-333.

Petry, N. M. (2011). Commentary on Van Rooij et al. (2011): 'Gaming addiction'– a psychiatric disorder or not? *Addiction*, *106*, 213-214.

Petry, N. M., Rehbein, F., Gentile, D. A., Lemmens, J. S., Rumpf, H. J., Mößle, T., ... & Auriacombe, M. (2014). An international consensus for assessing internet gaming disorder using the new DSM-5 approach. *Addiction, 109,* 1399-1406.

Potenza, M. N. (2018). Do gaming disorder and hazardous gaming belong in the ICD-11? Considerations regarding the death of a hospitalized patient that was reported to have occurred while a care provider was gaming. Journal of Behavioral Addictions, 7, 206-207.

Porter, G., Starcevic, V., Berle, D., & Penech, P. (2010). Recognising problem video game use. *Australian and New Zealand Journal of Psychiatry, 44,* 120-128.

Przybylski, A. K., Weinstein, N., & Murayama, K. (2016). Internet gaming disorder: Investigating the clinical relevance of a new phenomenon. *American Journal of Psychiatry, 174*, 230-236.

Rehbein, F. & Baier, D. (2013). Family-, media-, and school-related risk factors of video game addiction: A 5-year longitudinal study. *Journal of Media Psychology-Theories Methods and Applications, 25,* 118-128.

Rehbein, F., Kliem, S., Baier, D., Mößle, T., & Petry, N. M. (2015). Prevalence of Internet gaming disorder in German adolescents: Diagnostic contribution of the nine DSM- 5 criteria in a state- wide representative sample. *Addiction, 110,* 842-851.

Rumpf, H. J., Achab, S., Billieux, J., Bowden-Jones, H., Carragher, N., Demetrovics, Z., ... & Saunders, J. B. (2018). Including gaming disorder in the ICD-11: The need to do so from a clinical and public health perspective: Commentary on: A weak scientific basis for gaming disorder: Let us err on the side of caution (van Rooij et al., 2018). *Journal of Behavioral Addictions, 7,* 556-561.

Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption- II. *Addiction*, *88*, 791-804.

Saunders, J. B., Hao, W., Long, J., King, D. L., Mann, K., Fauth-Bühler, M., ... & Chan, E. (2017). Gaming disorder: Its delineation as an important condition for diagnosis, management, and prevention. *Journal of Behavioral Addictions*, *6*, 271-279.

Scharkow, M., Festl, R., & Quandt, T. (2014). Longitudinal patterns of problematic computer game use among adolescents and adults—A 2- year panel study. *Addiction, 109*, 1910-1917.

Schluter, M. G., Hodgins, D. C., Wolfe, J., & Wild, T. C. (2018). Can one simple questionnaire assess substance- related and behavioural addiction problems? Results of a proposed new screener for community epidemiology. *Addiction, 113*, 1528-1537.

Starcevic, V. (2013). Video-gaming disorder and behavioural addictions. *Australian & New Zealand Journal of Psychiatry, 47,* 285-286.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*, 350-353.

Seok, S. & DaCosta, B. (2012). The world's most intense online gaming culture: Addiction and high-engagement prevalence rates among South Korean adolescents and young adults. *Computers in Human Behavior, 28*, 2143-2151.

Sim, T., Gentile, D. A., Bricolo, F., Serpollini, G., & Gulamoydeen, F. (2012). A conceptual review of research on the pathological use of computers, video games, and the Internet. *International Journal of Mental Health and Addiction, 10,* 748-769.

Triberti, S., Milani, L., Villani, D., Grumi, S., Peracchia, S., Curcio, G., & Riva, G. (2018). What matters is when you play: Investigating the relationship between online video games addiction and time spent playing over specific day phases. *Addictive Behaviors Reports*, *8*, 185-188.

van Rooij, A. J., Schoenmakers, T. M., Vermulst, A. A., Van Den Eijnden, R. J., & Van De Mheen, D. (2011). Online video game addiction: Identification of addicted adolescent gamers. *Addiction, 106,* 205-212.

van Rooij, A. J., Ferguson, C. J., Colder Carras, M., Kardefelt-Winther, D., Shi, J., Aarseth, E., ... & Deleuze, J. (2018). A weak scientific basis for gaming disorder: Let us err on the side of caution. *Journal of Behavioral Addictions*, *7*, 1-9.

Wang, C. W., Chan, C. L., Mak, K. K., Ho, S. Y., Wong, P. W., & Ho, R. T. (2014). Prevalence and correlates of video and internet gaming addiction among Hong Kong adolescents: A pilot study. *The Scientific World Journal, 2014,* Article ID 874648

Wegmann, E., Oberst, U., Stodt, B., & Brand, M. (2017). Online-specific fear of missing out and Internet-use expectancies contribute to symptoms of Internet-communication disorder. *Addictive Behaviors Reports, 5*, 33-42.

Wei, L., Zhang, S., Turel, O., Bechara, A., & He, Q. (2017). A tripartite neurocognitive model of Internet Gaming Disorder. *Frontiers in Psychiatry, 8*(285).

World Health Organization (WHO). (2019). *6C51 Gaming disorder*. Retrieved online: https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/1448597234

WHO ASSIST Working Group. (2002). The alcohol, smoking and substance involvement screening test (ASSIST): development, reliability and feasibility. *Addiction, 97*, 1183-1194.

Xin, M., Xing, J., Pengfei, W., Houru, L., Mengcheng, W., & Hong, Z. (2018). Online activities, prevalence of Internet addiction and risk factors related to family and school among adolescents in China. *Addictive Behaviors Reports, 7,* 14-18.

Yao, Y. W., Liu, L., Ma, S. S., Shi, X. H., Zhou, N., Zhang, J. T., & Potenza, M. N. (2017). Functional and structural neural alterations in Internet gaming disorder: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews, 83,* 313-324.

Yee, N. (2006). Motivations for play in online games. *CyberPsychology & Behavior, 9*, 772-775.

Zendle, D. & Cairns, P. (2018). Video game loot boxes are linked to problem gambling: Results of a large-scale survey. *PloS One, 13, e0206767.*

Zendle, D. & Cairns, P. (2019). Correction: Video game loot boxes are linked to problem gambling: Results of a large-scale survey. *PloS One, 14,* e0214167.

Table 1 An overview of all reviewed tools (n=32) for problematic gaming and gaming disorder

Tool	Author	Components (N: Name)	Items	Response format	Cut-off	Age	Country of origin	Languages
PVP Scale	Salguero & Moran (2002)	7: Preoccupation; tolerance; loss of control; withdrawal; escape; deception; disregard	9	Yes/No	NR	13	Spain	ENG; FR; NOR; SPA
A-EQ	Charlton & Danforth (2007)	2: Addiction; engagement	29	7-point	Unclear	-	United Kingdom	ENG; FR
GAS-7	Lemmens et al. (2009)	7: Salience; tolerance; mood modification; relapse; withdrawal;	7	5-point	$4/7^1$	12	Netherlands	ENG; CH; DE; FA; FI; FR; NOR;

		conflict; problems						TR
GAS-21	Lemmens et al. (2009)	7: Salience; tolerance; mood modification; relapse; withdrawal; conflict; problems	21	5-point	Unclear	12	Netherlands	ENG; FR; NOR; SRB; TR
POGU	Kim & Kim (2010)	5: Euphoria; health problem; conflict; failure of self-control; preference for online relationship	20	NR	NR	10	South Korea	ENG; KOR
AICA-Sgaming	Wölfling et al. (2010)	8: Loss of control; tolerance; withdrawal; continued use; loss of interests; emotion regulation; jeopardising; craving	15	4-point	13.5	13	Germany	CZ; ENG; DE; ISL; GRK; LTU; NL; POL; ROU; SPA
POGQ	Demetrovics et al. (2012)	6: Preoccupation; overuse; immersion; social isolation; interpersonal conflicts; withdrawal	18	5-point	66/90	12	Hungary	ENG; FA; FR; FIN; HUN; ITL; KOR; MAL; NOR; SLO; SPA
VAT	van Rooij et al. (2012)	5: Loss of control; preoccupation; withdrawal; conflict, coping	14	5-point	NR	13	Netherlands	ENG; NL; PT
POGQ-SF	Pápay et al. (2013)	6: Preoccupation; overuse; immersion; social isolation; interpersonal conflicts; withdrawal	12	5-point	32/60	12	Hungary	CZ; ENG; FIN; FI; FR; HUN; ITL; KOR; MAL; NOR; SLO; SPA
sIAT-gaming	Pawlikowski et al. (2013)	2: Loss of control/time management; craving/social problems	12	5-point	NR	9	Germany	ENG; DE
IGD-20	Pontes et al. (2014)	6: Salience; mood modification; tolerance; withdrawal; conflict; relapse	20	5-point	71/100	16	United Kingdom	AR: ENG; KOR; PT; SPA;
GAIA	Wong & Hodgins (2014)	6: Loss of control and consequences; engagement; withdrawal (agitated/mournful); coping; shame	26	5-point	NR	18	Canada	ENG
Petry IGD	Petry et al.	9: Preoccupation; tolerance; withdrawal;	9	Yes/No	5/9	10	United	ENG; CH; DE; FR;

	(2014)	reduce/stop; loss of interests; continued use; deception; escape; conflict					States	ITL; JP; KOR; NL; PT; SPA; TR
IGDS9–SF	Pontes & Griffiths (2015)	9: Preoccupation; tolerance; withdrawal; reduce/stop; loss of interests; continued use; deception; escape; conflict	9	5-point	25/45	10	United Kingdom	AL; ENG; CH; FA; ITL; POL; PT; SI; TR;
Lemmens IGD-9	Lemmens et al. (2015)	9: Preoccupation; tolerance; withdrawal; persistence; escape; problems; deception; displacement; conflict	9	Yes/No	5/9	13	Netherlands	AR; ENG; CRO; DE; NL; TR;
Lemmens IGD-27	Lemmens et al. (2015)	9: Preoccupation; tolerance; withdrawal; persistence; escape; problems; deception; displacement; conflict	27	Yes/No 4-point	Unclear	13	Netherlands	ENG; NL; TR
GAIT	Vadlin et al. (2015)	9: Preoccupation; tolerance; losing track of time; craving; withdrawal; relapse; chasing losses; loss of interests; conflict	15	5-point	Unclear	13	Sweden	ENG; SE
CSAS	Rehbein et al. (2015)	9: Preoccupation; withdrawal; tolerance; unsuccessful; continued use; loss of interests; escape; deception; conflict	18	4-point	5/9	13	Germany	ENG; DE
PIE-9	Pearcy et al. (2016)	9: Preoccupation; withdrawal; tolerance; unsuccessful attempts; loss of interests; continued use; deception; escape; harm	9	5-point	5/9	16	Australia	ENG
BAM-VG	Sanders & Williams (2016)	3: Impaired control; Problems; other issues	19	Yes/No	3	18	Canada	ENG; FR
IGDT10	Király et al. (2017)	Preoccupation; withdrawal; tolerance; loss of control; giving up activities; continuation; deception; escape; negative consequences	10	3-point	5/9	14	Hungary	CZ; ENG; CH; CRO; FA; FI; FR; HUN; JP; NOR; POL; SPA
SCI-IGD	Koo et al. (2017)	9: Preoccupation; withdrawal; tolerance; unsuccessful attempts; loss of interests; continued	12	Yes/No	5/9	12	South Korea	ENG; KOR

		use; deception; escape; conflict						
VASC	Yılmaz et al. (2017)	4: Self-control; reward/reinforcement; problems; involvement	21	5-point	90/105	9	Turkey	ENG; TR
C-VAT2.0	van Rooij et al. (2017)	9: Preoccupation; withdrawal; tolerance; unsuccessful attempts; loss of interests; continued use; lying; mood modification; problems	11	Yes/No	5/9	13	Netherlands	ENG; NL
IGUESS	Jo et al. (2018)	9: Preoccupation; withdrawal; tolerance; unsuccessful attempts; loss of interests; continued use; deception; coping; conflict	9	4-point	10/18	10	South Korea	ENG; KOR
DIA	Ryu et al. (2019)	9: Salience; withdrawal; tolerance; difficulty in regulating use; loss of interests; persistent use; deception; use to avoid feelings; interference	10	Yes/No	5/10	13	South Korea	ENG; KOR
YIAT	Young (1998)	6: Salience; excessive use; neglecting work; anticipation; lack of control; neglecting social life	20	5-point	70/100	10	United States	ENG; CH; DE; FR; ITL; KOR; PT; TR
YDQ	Young (1998)	7: Preoccupation; tolerance; loss of control; withdrawal; deception; escape; conflict	8	Yes/No	5/8	13	United States	ENG; CH; DE; FR; KOR; ITL; LTU; PT; ROU; SPA; TR
CIUS-14	Meerkerk et al. (2009)	6: Loss of control; preoccupation; withdrawal; conflict; coping; conflict	14	5-point	28/70	14	Netherlands	ENG; CH; DE; FIN; FR; NL
CIUS-5	Besser et al. (2017)	3: Loss of control; conflict; coping	5	5-point	7/25	16	Germany	ENG; CH; DE; FIN; FR NL;
SSBA	Schluter et al. (2018)	4: Overuse; loss of control; coping; continued use	4	6-point	NR	18	Canada	ENG
CIUS-8	Gmel et al. (2019)	6: Loss of control; preoccupation; withdrawal; conflict;	8	5-point	13/40	15	Switzerland	ENG; CH; DE; FIN; FR; NL;

	coping; conflict

NR: Not reported. (y); years. Age: Age, in years, of the youngest respondent to complete the test.

[1]Lemmens et al. specify a cut-off of 4 out of 7 items met. Items are met if respondent indicates at least 3 ('sometimes') out of 5.

NB: The first 26 tools measure gaming-related problems only and are presented in order of publication date. The remaining 6 tools are Internet-related or all-purpose measures; these tools are also presented in order of publication date.

List of tools: A-EQ: Addiction-Engagement Questionnaire; AICA-Sgaming: Assessment of Internet and Computer Addiction Scale-Gaming; BAM-VG: Behavioral Addiction Measure for Video Gaming; CSAS: Video Game Dependency Scale; C-VAT 2.0: Clinical – Video Game Addiction Test 2.0; DIA: Diagnostic Interview for Internet Addiction; GAIT: Game Addiction Identification Test; GAIA: Game Addiction Inventory for Adults; GAS-7: Game Addiction Scale-7 items; GAS-21: Game Addiction Scale-21 items; IGD-20: Internet Gaming Disorder-20 Test; IGDS9-SF: Internet Gaming Disorder Scale-9 Short Form; IGDT-10; Internet Gaming Disorder Test-10 items; IGUESS: Internet Game Use Elicited Symptom Screen; Lemmens IGD-9: Internet Gaming Disorder Scale-9 items; Lemmens IGD-27: Internet Gaming Disorder Scale-27 items; Petry IGD: Petry et al. (2014) consensus statement on IGD criteria; PIE-9: Personal Internet Gaming Disorder Evaluation-9 items; POGU: Problematic Online Game Use; POGQ: Problematic Online Gaming Questionnaire; POGQ-SF: Problematic Online Gaming Questionnaire-Short Form; PVP Scale: Problematic Video game Playing Scale; SCI-IGD: Structured Clinical Interview-Internet Gaming Disorder; sIATgaming: Short Internet Addiction Test-Gaming; VASC: Video Game Addiction Scale for Children; VAT: Video Game Addiction Test; CIUS: Compulsive Internet Use Scale [NB: CIUS-5, CIUS-8 and CIUS-14 are not the published names but have been adopted for this review to minimize confusion between these multiple versions. Actual test names: CIUS-14 is 'CIUS'; CIUS-8 is 'Short form of the CIUS'; CIUS-5 is the 'Short CIUS'. Each number accompanying CIUS in this review refers to the number of items in each version]; YIAT: Young Internet Addiction Test; YDQ: Young Diagnostic Questionnaire.

Languages: AL: Albanian; AR: Arabic; ENG: English; CH: Chinese; CRO: Croatian; CZ: Czech; DE: German; FA: Farsi; FIN; Finnish; FR: French; HUN; Hungarian; ISL: Icelandic; ITL: Italian; FI: Finnish; GRK: Greek; JP: Japanese; KOR: Korean; LTU; Lithuanian; MAL: Malay; PT: Portuguese; NL: Dutch; NOR: Norwegian; POL; Polish; ROU: Romanian; SE: Sweden; SLO: Slovenian; SPA: Spanish; SRB: Serbian; TR: Turkey.

Table 2 Coverage of DSM-5 and ICD-11 gaming disorder criteria across all tools (N=32)

Tool	1. Preoccupation	2. Withdrawal	3. Tolerance	4. Unsuccessful attempts	5. Loss of interests	6. Continued use	7. Deception	8. Escape	9. Jeopardised life	ICD-11	1. Impaired control	2. Increasing priority	3. Continuation	Impairment - Personal	Impairment - Social	Impairment - Education	Impairment - Work	Impairment - Financial	TOTAL: DSM-5[1]	TOTAL: ICD-11[2]
PVP	●	●	●	●	○	○	●	●	●		●	●	○	●	●	●	●	○	**7**	3
A-EQ	●	●	●	●	○	○	○	●	●		●	●	○	●	●	○	●	●	5	3
GAS-7	●	●	●	●	○	○	●	●	●		●	●	○	●	●	●	●	○	7	3
GAS-	●	●	●	●	●	○	●	●	●		●	●	○	●	●	●	●	○	8	3

Test																			DSM-5[1]	ICD-11[2]
POGU	○	○	●	●	○	○	○	○	●	●	●	○	●	○	●	○	○		3	3
AICA-	●	●	●	●	●	○	○	●	●	●	○	○	●	●	●	●	●		7	2
POGQ	●	●	○	●	●	○	○	○	●	●	●	○	●	●	○	○	○		5	3
VAT	●	●	○	●	○	○	○	●	●	●	●	○	●	●	○	○	○		5	3
POGQ	●	●	○	●	●	○	○	○	●	●	●	●	●	●	○	○	○		5	3
sIAT-	●	●	○	●	●	○	●	○	●	●	●	●	●	●	●	○	○		6	3
IGD-	●	●	●	●	●	○	●	●	●	●	●	○	●	●	●	●	○		8	3
GAIA	○	●	●	○	○	○	●	●	●	●	●	○	●	●	○	●	●		5	3
Petry	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	○		<u>9</u>	<u>4</u>
IGDS9	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●	●	○		<u>9</u>	<u>4</u>
Lemm	●	●	●	●	●	○	●	●	●	●	●	○	○	●	○	○	○		8	3
Lemm	●	●	●	●	●	○	●	●	●	●	●	●	●	●	○	○	○		8	3
GAIT	●	●	●	●	●	○	○	○	●	●	●	○	●	●	○	○	○		6	3
CSAS	●	●	●	●	●	○	●	●	●	●	●	○	●	●	●	●	○		7	3
PIE-9	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●	●	○		<u>9</u>	<u>4</u>
BAM-	●	●	●	●	○	○	●	●	●	●	●	○	●	●	●	●	●		7	3
IGDT1	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	○		<u>9</u>	<u>4</u>
SCI-	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●	●	○		<u>9</u>	<u>4</u>
VASC	○	○	○	●	○	○	○	○	●	●	○	○	●	●	○	○	○		4	2
C-	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	○	○		<u>9</u>	<u>4</u>
IGUES	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●	○	○		<u>9</u>	<u>4</u>
DIA	●	●	●	●	●	●	●	●	●	●	●	○	●	●	●	○	●		<u>9</u>	<u>4</u>
YIAT	●	●	○	●	○	○	●	○	●	●	●	○	●	●	●	●	○		6	3
YDQ	●	●	●	●	○	○	●	●	●	●	○	○	●	●	●	○	○		7	2
CIUS-	●	●	○	●	○	○	●	●	●	●	●	○	●	●	●	●	○		6	3
CIUS-8	●	●	○	○	○	○	○	●	●	●	●	○	●	●	●	○	○		4	3
CIUS-5	○	○	○	○	○	○	○	●	●	●	●	○	●	●	●	○	○		2	3
SSBA	○	○	○	○	○	●	○	●	○	●	○	●	○	○	○	○	○		2	1

Note: ● assessed; ○ not assessed. NB: See Table 1 for full test names. Underlined values represent full coverage of respective criteria.

[1]Refers to total DSM-5 criteria covered by test (max. 9).

[2]Refers to total ICD-11 criteria covered by test (max 4); NB: 1 point for any specific impairment items (listed below).

<u>Impairment types</u>:

Personal: Sleep, appetite, well-being due to excessive use, basic hygiene, other health-related issues.

Social: Relationship conflict/interference, neglect, includes relationships with partner, children, other family members, and friends.

Education: Conflict/interference/disruption, neglect, loss of educational [school/university/learning] opportunities, productivity, outcomes.

Work: Conflict/interference/disruption, neglect, loss of occupational [paid/volunteer work] opportunities, productivity, outcomes.

Financial: Problems related to spending too much money on gaming activities.

Table 3 The empirical evidence base for reviewed tools for problematic gaming and gaming disorder

Instrument	Citations	Studies	Validat‡	Interv.	Clinical		SDI &	Long.	Prev.	Neurobiological	
---	---	---	---	---	---	---	---	---	---	Imaging	Cognitive
PVP Scale	494	22	-	-	-	2	-	-	7	-	2
A-EQ	798	17	-	-	-	-	-	1	2	-	-
GAS-7	665	45	6:6	-	-	-	-	12	18	1	3
GAS-21	665	16	-	-	-	-	-	2	1	-	-
POGU	144	2	-	1	-	-	1	-	-	-	-
AICA-	55	18	3;1	1	-	1	10	-	5	3	-
POGQ	155	8	-	-	-	-	-	-	-	-	2
VAT	83	7	1;0	1	-	-	-	-	3	1	-
POGQ-SF	78	6	1;0	-	-	-	-	-	2	-	1
sIAT-	153	8	-	-	-	-	-	-	-	-	-
IGD-20	150	8	-	1	-	-	2	-	-	-	-
GAIA	15	1	-	-	-	-	-	-	-	-	-
Petry IGD	445	13	-	-	4	1	-	-	-	4	1
IGDS9–SF	152	24	1;0	-	2	-	-	2	3	-	-
Lemmens	186	16	1;0	-	-	-	-	1	4	-	-
Lemmens	186	2	1;0	-	-	-	-	-	1	-	-
GAIT	15	4	-	-	-	1	-	1	3	-	-
CSAS	12	2	-	-	-	-	-	-	2	-	-
PIE-9	10	3	-	-	-	-	-	-	-	-	-
BAM-VG	13	2	-	-	-	-	-	1	-	-	-
IGDT-10	68	7	1;1	-	-	-	-	-	1	-	-
SCI-IGD	14	1	1;0	-	1	-	-	-	-	-	-
VASC	7	1	-	-	-	-	-	-	-	-	-
C-VAT2.0	40	1	1;0	-	-	-	1	-	-	-	-
IGUESS	6	4	2;0	-	2	-	-	-	1	-	-
DIA	-	1	-	-	1	-	-	-	-	-	-
YIAT	2.387	62	-	4	23	-	16	5	2	24	3
YDQ	5,413	12	1;1	1	-	-	4	2	7	-	2
CIUS-14	580	12	4;2	-	1	-	-	-	7	-	1
CIUS-8	-	1	1;0	-	1	-	-	-	-	-	-
CIUS-5	5	1	1;0	-	-	-	-	-	1	-	-
SSBA	2	1	-	-	-	-	-	-	1	-	-
TOTAL	**12996**	**328**	**26;11**	**14**	**35**	**5**	**34**	**30**	**71**	**33**	**15**

Interv: Intervention studies. Long: Longitudinal studies. Prev: Prevalence studies. SDI: Structured/semi-structured diagnostic interview. T-S: Treatment-seeking sample. ‡Validation samples, where the first value is the number of studies using nationally representative or clinical samples, and the second value in the total number of independent studies (i.e., studies not conducted by test author). NB: 328 studies due to duplicated entries for papers with multiple tests and studies (study 1, study 2, etc).

Underlined: Most frequent in column.

Table 4 Sample size and recruitment for studies using the reviewed tools (n=24*), ranked by total N

Tool	Studies	Total N	Adol.	Adult	Conv.	Non-Conv.	%Adol	%Non-Conv.
GAS-7	45	**94,389**	22,616	**69,789**	**38,393**	**55,618**	24.0	58.9
YIAT	62	49,509	1973	47,536	16,347	31,592	4.0	63.8
CIUS-14	12	46,235	11,763	26,340	22,334	23,901	25.4	51.7
AICA-Sgaming	18	36,306	28,447	7,859	5,211	31,095	78.3	85.7

YDQ	12	30,916	**29,810**	1,106	1,106	29,810	**96.4**	**96.4**
PVP Scale	22	26,260	7,626	18,634	6,527	19,066	29.0	74.5
CSAS	2	26,171	26,717	0	0	26,171	100	100
IGDS9–SF	24	25,503	8,894	14,609	22,432	1,071	34.9	4.2
IGDT-10	7	21,702	8,883	12,819	21,702	0	40.9	0
Lemmens IGD-9	16	19,865	10,197	9,668	12,910	6,955	51.3	35
POGQ-SF	6	18,366	14,809	3,557	3,557	14,809	80.6	80.6
VAT	7	13,478	13,198	0	1,826	11,372	97.9	84.3
POGQ	8	9585	2,524	7,061	9,585	0	26.3	0
A-EQ	17	8113	3,332	4,781	4,781	3,332	41.1	41.1
GAS-21	16	5807	1,676	4,131	5,500	88	29.9	1.5
IGD-20	8	5454	930	4,524	4,379	1,075	17.1	19.7
Petry IGD	13	4728	861	3,56	4,542	186	18.2	3.9
BAM-VG	2	4448	0	4448	506	3,942	0	88.6
IGUESS	4	3796	3,796	0	0	3,796	100	100
GAIT	4	3745	3,745	0	0	3,745	100	100
Lemmens IGD-27	2	2901	2,444	457	2,444	457	84.3	15.8
sIATgaming	8	2625	0	2,625	2,625	0	0	0
POGU	2	1505	1,442	63	1,505	0	95.8	0
PIE-9	3	842	0	842	842	0	0	0
TOTAL	**320**	**462,249**	**205,683**	**244,716**	**189,054**	**268,081**	**44.5**	**58.0**

*Tools not listed due to 1 study only: SCI-IGD, VASC, C-VAT2.0, GAIA, DIA, CIUS-8, CIUS-5, SSBA.

Adol.: N of participants aged <18 years. Adult: N of participants aged 18+ years. Conv.: Convenience sample N. Non-Conv.: Non-convenience sample N. Includes nationally representative studies, cohort studies, randomly selected samples, treatment-seeking populations. NB: Studies with undifferentiated samples of adolescents and adults are included in 'Total N' but not listed in other columns. Underlined, bold: Largest N in column. For % columns, the highest value with 10+ studies is bolded and underlined.

Table 5 Psychometric properties of GD tools I: Dimensionality, reliability, and scale refinement

Instrument	Dimensionality		Reliability		Refinement	
	Method	Factor structure	IC‡	Test-retest	Rasch; IRT	Measurement invariance
PVP Scale	PCA[1,7,11], EFA[7], CFA[2]	1-factor[1,2,7,11]	.59[2], .62[21], .66[2], .69[1,15], .74[18,20], .75[9], .78[6,11], .80[13], .84[13]	-	-	-
A-EQ	PCA & EFA[22], CFA[28]	2-factor[22,28]	.69[32,33], .75[28], .77[30], .79[37], .86[23,38], .87[27], .91[24]	-	-	-
GAS-7	EFA[50,72,73] CFA[39,45,48,50,51,72]	1-factor[39,48,50,51,71,72]	.66[53], .72[67], .79[68,71], .80[50], .81[39,55], .83[43], .84[65,70], .85[47,52], .86[39,57], .88[42], .89[48], .90[57], .93[64], .94[64]	.82 (2 weeks)[82]; .83 (2 weeks)[48]	Rasch: partial credit model: satisfactory[48]; IRT using GRM: all items with high discrimination parameters, some items with high DIF[51]	Sex, Age groups: Configural[45]; Sex, Usage: Configural, Metric, Scalar[48]; Language groups (French, German): Configural[72]
GAS-21	CFA[83,88,97]	1-factor, 7 second-order factors: supported[83,97] not supported (1-factor only)[88]	.92[83,87,88], .94[83,84,94,95,96], .96[97]	.76 (30 days)[88]; .84 (2 weeks)[84]	-	-
POGU (Kim)	EFA & CFA[120], CFA[121,123,125,126]	6-factor[120,121,123,125,126]	.84[122], .91[121,127], .92[125], .93[123,126]	-	-	-
AICA-Sgaming	PCA[102,113]	1-factor[102,113]	.70[117], .79[112], .82[107]; .84[114,119], .89[102,113], .92[109]	-	-	-
POGQ	EFA & CFA[120], CFA[121,123,125,126]	6-factor[120,121,123,125,126]	.84[122], .91[121,127], .92[125], .93[123,126]	-	-	Online vs Offline groups: Configural, Metric, Scalar[125]
VAT	EFA[132,133], CFA[128]	1-factor[128], 3-factor[133], Inadequate	.81[133], .92[132], .93[128,134]	.76 (30 days)[132]	-	Sex, Ethnicity, Grade

		model[132]				groups: Configural, Metric[128]
POGQ-SF	CFA[135]	6-factor[135]	.90[137], .91[135], .92[140], .93[136,139]	-	-	Sex groups: Configural, Metric, Scalar[135]
sIATgaming	PCA & CFA[141], CFA[143,145]	1-factor[145]; 2-factor[141,143]	.70[142], .81[147], .84[144], .85([143,145]), .88[148], .90[141], .91[146]	-	-	-
IGD-20	EFA & CFA[151], CFA[149,152,155]	1-factor[151]; 6-factor[149,152,155]	.85[155], .87[152,156], .88[149], .92[150,151,154]	-	-	-
GAIA	EFA[157]	6-factor[157]	.94[157]	-	-	-
Petry IGD	EFA & CFA[164,170]	1-factor[164,170]	.69[160], .70[170], .72[162], .77[166], .78[162,169], .82[1?4], .86[163]	-	-	-
IGDS9-SF	EFA & CFA[171,190] CFA[173,179,182-189, 194]	1-factor[171,179,182-189,190,194]	.81[185], .8?[1?3,187], .84[1?6,19?], .8?[181,?5,186,191], .?8[171,172,180,183,184,19?], .?9[184,190], .90[172,174,183,188,189], .91[182,183,184], .92[174,175,176], .93[179], .96[178,185,194]	.87 (2 weeks)[189]	2PLM IRT findings: High discrimination parameters for all items[182]; Items 6,7,8 have poor fit[187], Rasch: partial credit model: Item 4 had high DIF[189]	Aust, US, UK: Mixed support[183]; US, India, UK; Mixed support[184]; Albania, US, UK, Italy: Mixed support[185]; Limited support[186]; Gender, Gaming time: Mixed support[189]; Gender, Age: Configural, Metric, Scalar[194]
Lemmens IGD-9	EFA & CFA[196], CFA[195,198]	1-factor[195,196,198]	.67[210], .73[208], .74[201,206], .76[208], .77[201,206], .82[197,200], .83[195,197], .85[202], .86[198], .93[196,205], .95[199]	-	-	-

Lemmens IGD -27	EFA & CFA[196], CFA[195]	1-factor[196]; 1-factor, 9 second-order factors[195]	.95[196], .94[195]	-	-	-
GAIT	EFA[217]	1-factor[217]	.89[214], .90[214], .91[215,216]	-	-	-
CSAS	-	-	.92[219], .94[218]	-	-	-
PIE-9	EFA, CFA[220]	1-factor[220]	.86[220], .89[220]	.77 (2 weeks)[220]	-	-
BAM-VG	PCA[223]	2-factor[223]	.87[223]	.73 (20 days)[223]	-	-
IGDT-10	EFA[227], CFA[225,228,230]	1-factor[225,227,228,230]	.62[230], .68[225], .73[229], .75[230], .76[227], .85[226], .87[228]	-	2PL-IRT: Acceptable[225]	Language, Gender groups: Configural, Scalar invariance[230]
SCI-IGD	-	-	-	.41 to .91 (4 weeks)[232]	-	-
VASC	EFA, CFA[233]	4-factor[233]	.89[233]	-	-	-
IGUESS	-	-	.85[236], .86[238], .94[235]	-	-	-
DIA	-	-	.72[239]	-	-	-
YIAT-20	PCA[297], CFA[279]	1-factor: poor fit[279], 3-factor, 4-factor[297]	.86[263], .88[284], .89[282,301], .90[269,299], .93[289], .94[242,283,298,299], .96[266,285,294], .97[263]	-	-	-
YDQ	EFA & CFA[311]	1-factor[311]	.62 & .66 (KD-20)[311], 67[303], .70[313], .76[306], .81[306], .83[310], .86[306], .95[312]	-	-	-
CIUS-14	EFA[318], EFA & CFA[315,317]CFA[314,316,323]	1-factor with poor fit[314,315,317,323], 1-factor[316,318]	.83[315], .86[315], .87[315], .89(314), .91[317,319], .93[323], .95[318]	-	-	Sex, Age, Education, Internet use groups: Configural[316]

CIUS-8	CFA[326]	1-factor[326]	.89[326]	-	IRT-graduated response model: High discrimination parameters[326]	Sex, Region, Age groups: Configural, Metric, Scalar for Region and Age (not Sex)[326]
CIUS-5	-	-	.77[327]	-	-	-
SSBA	PCA[328]	1-factor[328]	.87 -.95[328]	-	-	-

Superscript numbers refer to references; see Supplementary file 1 for complete reference list. C-VAT2.0 excluded due to lack of relevant results.

‡Cronbach's alpha value. Aust; Australia. CFA: Confirmatory factor analysis. EFA: Exploratory factor analysis. IC: Internal consistency. KD-20: Kuder-Robinson;

PCA: Principal components analysis. IRT: Item response theory. UK: United Kingdom. US: United States.

Table 6 Psychometric properties of GD tools II: Criterion validity, impairment, and clinical use

| Instrument | Convergent validity | Criterion validity | Impairment/QOL | Clinical use | |
	Gaming behavior	Other GD tests	Type: result	Independent evaluation	Treatment changes
PVP Scale	.22[21], .23[8], .24[21], .28[16,18], .30[7,19], .48[2,9], .54[14], .61[6], .63[11], .64[1,15]	SOD: .47[1], .52[?], YDQ: .27[12], VIAT: .?2[?], .53[6], GAS-21: .70[16], .74[19], PIE-9: .43[220]	-	Psychiatric interview including K-CIDI & PVP[4]	-
A-EQ	.20[28], .21[33], .29[30]	Petry IGD: .56[28]	WHOQOL-BREF: All 4 scales sign. lower for IGD: small to medium effect[35], large effect[23]	-	-
GAS-7	.22[50], .39[46], .40[50], .43[73,75], .47[44,82], .48[73], .51[75], .55[39],	IGDS9-SF: .40[48,177],	PedsQL: -.15 (SEM, B)[48]; HBSC: Mixed	-	-

	.58[39], .67[57], .68[57]	sIAT: .51[53], GAIT: .83[217], YDQ: .27[77], YIAT: .47[46], .62[64], VAT: .74[128], Petry IGD: .76[164], PIE-9: .57[220]	results[52]		
GAS-21	.19[97], .32[93], .42[94,95], .45[84], .55[83], .58[83]	Online Cognition Scale: .73[87], PVP: .70[94], .74[95], VAT: .88[88], YIAT: .45[88], .63[97]	-	-	-
POGU (Kim)	-	Craving: .48[100]		-	Reduced POGU in CBI vs Controls[100]
AICA-Sgaming	Addict>Control [104,106,109,113,116]	Craving: Addict> Control[10]		Psychological consultation, including diagnostic interview and scales[101-106,109,110, 115,116,119]	Sign. post-treatment reduction in AICA-Sgaming score (large effect size)[118]
POGQ	.40[121], .45[125], .46[123,126]	PIUQ-6: .68[121], IGDT-10: .69[230], .77[225]	-	-	-
VAT	Addict>Control[131], .25[128], .37[128], .40[132]	GAS-7: .74[128], .88[129], CIUS-14: .61[128]	-	-	School-based prevention: Decreased VAT & screen time[13]

POGQ-SF	.35[139]	PIU-6: .59[139]	-	-	-
sIATgaming	.51[146]	-	-	-	-
IGD-20	.27[155], .42[152], .49[151], .58[154], .77[149]	DSM-5 IGD criteria: .82[149], IGADS: .80[155], IGDS9-SF: .82[171], KIAS: .59[155], PIE-9: .64[220], BAM-VG: .44[223]	GAF: 'Serious' for IGD group[156]	Psychological consultation and scales[153,156]	Reduced IGD-20 for PIPACTIC group[156]
GAIA	-	YIAT: .80[157], SSBA: .79[328]	-	-	-
Petry IGD	.24[170], .33[163], .38[166], .42[164], .59[169]	A-EQ: .56[28], CIUS-14: .59[160], GAS-7: .76[164], IGCS: .51[169]	-	Psychiatric interview including MINI and other scales[159,161, 165, 168]	-
IGDS9-SF	.32[192], .33[188], .34[171], .36[186], .44[192], .47[179], .52[179], .55[190]	GAS-7: .40[48,17], .8_[194], IGD-20: .82[171], Lemmens IGD-9: .77[190], TIAS: .57[173], YIAT: .82[178], .83[194]	-	-	-
Lemmens IGD-9	.22[210], .23[195], .24[202], .29[204], .33[199], .54[204], .56[196], .66[208]	Parent version of IGDS: .78[198], IGDS9-SF: .77[190], YDQ:	KIDSCREEN-10: -0.42[200], IGD vs non-IGD: PROMIS Global Health Scale: Small effect size, Neuro-	-	-

		.39[196]	QOL Sleep Disturbances: Small effect size[211]		
Lemmens IGD-27	.23[195], .56[196]	-	-	-	-
GAIT	.48[217]	GAS-7: .83[217]	-	-	-
PIE-9		GAS-7: .57[220], IGD-20: .64[220], PVP: .43[220]	IGD vs non-IGD: WHO-DAS: Large effect size[220]		
BAM-VG	-	IGD-20: .44[223]	-	-	-
IGDT-10	.03 (non-sig)[227], .30[230], .45[228]	PIUQ-6: .53[231], POGQ: .69[230], .77[225]	-		-
SCI-IGD	-	IGD > non-IGD group: K-Scale[232]	-	Psychiatric interview and scales[232]	-
C-VAT 2.0	-	-	-	Treatment professional consultation[234]	-
IGUESS	IGD > non-IGD[236]	YIAT: .20[235], K-Scale[235]	-	Psychiatric interview and scales, with blinded scoring of tests[235,236]	-
DIA	-	K-Scale: .42[239], YIAT: .39[239]	-	Trained clinicians under psychologist supervision[239]	
YIAT	.10[289], .21[284], .31[298], .51[299], .57[285], .58[283]	GAS-7: .47[46], .62[64], GAS-27: .45[88], .63[97], PVP: .27[12],	IGD vs non-IGD group: WHOQOL-BREF: Sign diff[253]	Psychiatric consultation, including structured interview with scales[240,243-246,248, 251-255, 257-260,265, 267-]	Non-sig results[240], Reduced YIAT scores for groups: Drug<Placebo/Control[251,252,253,291], Drug+CBT<CBT[257], 2xDrug vs Control[258], Single drug group[292], CBT=VRT[293]

		GAIA: .80[157], IGDS9-SF:.82[178], .83[194], IGUESS: .90[235], DIA: .39[239], CIUS-14: .62[315]			269,273,276,277,280, 287,288, 290,291,293,294, Unspecified consultation[249], 261,263,292
YDQ	IGD > non-IGD[307], .34[311]	PVP: .27[12], Lemmens IGD-9: .39[196]	-	Psychiatric consultation, including structured interview with scales[30,313], Unspecified psychiatric consultation[304]	Reduced YDQ for treatment group (no control)[305], Reduced YDQ in two treatment groups[313]
CIUS-14	IGD > non-IGD[320,324], .28[315], .33[314], .35[315], .42[314], .47[315]	VAT: .61[128], YIAT: .62[315], .66[315], 75[315], Petry IGD: .59[160]	-		-
CIUS-5	-	-	-	Structured clinical interview for selected sample[327]	-
SSBA	-	GAIA: .79[328]	-	-	-

Superscript numbers refer to references; see Supplementary file 1 for complete reference list. CSAS, CIUS-8, and VASC excluded due to lack of relevant results.

CBI: Craving Behavior Intervention. CBT: Cognitive Behavior Therapy. GAF: Global Assessment of Functioning. Gaming behavior: Hours per week spent playing games. HBSC: Health Behavior in School-Aged Children Symptom Checklist; IGADS: Internet Game Addiction Diagnostic Scale (Korean); IGCS: Internet Gaming Cognition Scale. K-CIDI: Korean version of the Composite International Diagnostic Interview. KIAS: Korean Internet Addiction Scale. SOD: Severity of Dependence Test. PedsQL: Pediatric Quality of Life Inventory Short Form. PIPACTIC: Individualized Psychotherapeutic Program for the Addiction to Information and Communication Technologies. PIUQ-6: Problematic Internet Use Questionnaire-6 items. SEM: Structural Equation Modeling. TIAS: Turkish Internet Addiction Scale. VRT: Virtual Reality Therapy. WHODAS: World Health Organization Disability Assessment Schedule. WHOQOL-BREF: World Health Organization Quality of Life - Brief Instrument.

Table 7 Quick reference guide to all reviewed tools (n=32) for problematic gaming and gaming disorder

Tool	1. DSM-5 coverage	2. ICD-11 coverage	3. Cut-off score	4. Validation sample	5. Prevalence data	6. Longitudinal studies	7. Dimensionality	8. Internal consistency	9. Test-retest reliability	10. Criterion validity	11. Test refinement	12. Impairment	13. Structured interview	TOTAL SCORE (/23)
PVP	○	○	○	○	●	○	○	●	○	●	○	○	○	6
A-EQ	○	○	○	○	●	◐	○	●	◐	○	○	○	○	6
GAS-7	○	○	●	●	●	●	●	●	●	●	●	○	○	**17**
GAS-21	○	○	●	○	◐	●	◐	●	●	○	○	◐	○	11
POGU	○	○	○	○	○	○	◐	◐	○	○	○	○	◐	3
AICA-	○	○	●	●	●	○	○	○	○	○	○	○	●	9
POGQ	○	○	●	○	○	○	●	○	◐	○	○	○	○	6
VAT	○	○	○	◐	○	○	○	●	○	●	●	○	○	8
POGQ-	○	○	◐	◐	○	○	○	◐	○	◐	○	○	○	7
sIAT-	○	○	○	○	○	○	◐	○	◐	○	○	○	○	4
IGD-20	○	○	●	○	○	○	◐	●	○	●	○	○	●	7
GAIA	○	○	○	○	○	○	○	◐	○	◐	○	○	◐	2
Petry	●	●	●	○	○	○	●	●	○	●	○	○	○	9
IGDS9–	●	●	●	◐	●	◐	●	●	◐	●	●	○	○	**16**
Lemme	○	○	●	●	◐	◐	◐	●	●	●	○	○	○	**11**
Lemme	○	○	●	◐	◐	◐	◐	●	○	○	○	○	○	7
GAIT	○	○	●	◐	●	◐	○	●	◐	○	○	○	○	9
CSAS	○	○	●	○	●	◐	○	●	○	○	○	○	○	7
PIE-9	●	●	●	○	◐	○	◐	●	◐	●	○	◐	○	**10**
BAM-	○	○	●	○	◐	○	○	◐	◐	◐	○	○	○	4
IGDT10	●	●	●	○	◐	◐	◐	◐	◐	●	◐	○	○	**13**
SCI-IGD	●	●	●	○	○	○	○	○	◐	○	○	○	○	5
VASC	○	○	●	◐	○	○	○	○	○	○	○	○	○	3
C-	●	●	◐	○	○	○	○	○	○	○	○	○	◐	5
IGUESS	●	●	●	◐	◐	◐	●	○	●	○	○	○	○	10
DIA	●	●	◐	○	○	○	○	◐	○	●	○	○	○	6
YIAT	○	○	●	●	○	●	●	●	○	○	○	○	●	9
YDQ	○	○	◐○	●	○	●	◐	●	○	○	○	○	●	**12**
CIUS-14	○	○	●	●	●	◐	●	○	●	○	○	○	○	**12**
CIUS-8	○	○	●	◐	◐	◐	◐	●	○	◐	○	○	○	7
CIUS-5	○	○	●	◐	○	○	○	◐	○	○	○	○	◐	4
SSBA	○	○	○	○	◐	◐	○	◐	○	◐	○	○	○	4

Full score is 2 points, except for the criteria 1, 2 and 3, where full score is 1 point. ● Full score. ◐ Half-score. ○ No score. Underlined values denote the tests with the highest scores. See Table 1 for full test names.

Scoring criteria:

Criterion 1: Coverage of DSM-5 Internet gaming disorder criteria [0: Does not provide coverage of all DSM-5 IGD criteria. 1: Coverage of all DSM-5 IGD criteria.]; **_Criterion 2:_** Coverage of ICD-11 Gaming disorder criteria [0: Does not provide coverage

of all ICD-11 criteria. 1: Coverage of all ICD-11 GD criteria.]; *__Criterion 3:__* Cut-off score is present [0: No cut-off or unclear cut-off score. 1: Has a cut-off score.]; *__Criterion 4:__* Validation sample quality, indicated by psychometric properties evaluated in a nationally representative or clinical sample [0: No available data; 1: At least one study; 2: at least 1 study, and another study undertaken by an independent team]; *__Criterion 5:__* Prevalence data available [0: No available data. 1: Has 1 nationally representative prevalence study. 2: Has 2+ nationally representative studies.]; *__Criterion 6:__* Longitudinal data available [0: No available data. 1: Has 1 longitudinal study. 2: Has 2+ longitudinal studies.]; *__Criterion 7:__* Dimensionality assessed by exploratory factor analysis and confirmatory factor analysis in two independent samples [0: No available data. 1: 1 study. 2: 2+ studies]; *__Criterion 8:__* Internal consistency, as assessed by Cronbach's alpha [0: No available data. 1: 1 study reporting alpha > .70. 2: 2+ studies reporting alpha > .70]; *__Criterion 9:__* Test-retest reliability [0: No available data. 1: 1 study with alpha > .70. 2: 2+ studies with alpha > .70]; *__Criterion 10:__* Criterion validity, as assessed by association with other GD/IGD tests [0: No available data. 1: 1 study reporting association reporting .3 or higher. 2: 2+ studies reporting association .3 or higher.]; *__Criterion 11:__* Rasch or Item Response Theory [0: No available data. 1: At least 1 study that reports strong support for items/model fit using Rasch or IRT analysis. 2: At least 1 study that gives strong support for items/model fit using Rasch or IRT analysis, and 1 study undertaken by independent research team.]; *__Criterion 12:__* Test is significantly related to validated measure of functional impairment [0: No available data. 1: 1 study with correlation of at least .3. 2: 2+ studies with correlation of at least .3]; *__Criterion 13:__* Test used in conjunction with structured/semi-structured interview involving a sample of treatment-seeking gamers or problem gamers with evidenced functional impairment [0: No available data. 1: 1 study; 2+ studies.]

Figure 1. Search results and tool selection in accordance with PRISMA guidelines

Highlights

- Numerous tools for gaming disorder (GD) have been developed in recent years.
- We evaluated 32 GD tools and their evidence base from 320 empirical studies.
- Several instruments had greater evidential support than others.
- No single tool emerged as the clearly optimal choice.
- A standard international tool would be invaluable to advance the GD field.

Author Biography

Daniel L. King, PhD, MPsych (Clin), is a Senior Research Fellow and registered clinical psychologist in the College of Education, Psychology, and Social Work at Flinders University. He is the author (with Paul Delfabbro) of *Internet Gaming Disorder: Theory, Assessment, Treatment & Prevention*, and has authored more than 100 peer-reviewed papers on the topic of digital technology-based problems, including gaming and simulated gambling.

Dr Samuel Chamberlain is Wellcome Trust Clinical Fellow and Honorary Consultant Psychiatrist at the University of Cambridge. His research and clinical work focuses on the neurobiology and treatment of impulsive and compulsive disorders.

Natacha Carragher, PhD, ATCL, is a Consultant with the World Health Organization, Geneva Switzerland, and a conjoint Senior Lecturer at the University of New South Wales, Sydney Australia. She has published 70 peer-reviewed papers relating to the classification and assessment of mental health, substance use and addictive behaviours.

Joël Billieux, PhD, is Associate Professor of Clinical Psychology, Psychopathology, and Psychological Assessment at the University of Lausanne, Switzerland. His main area of research regards the psychological factors (cognitive, affective, motivational, interpersonal) involved in the etiology of addictive behaviors, with a particular focus on self-regulation-related processes. Another focus of his research is the conceptualization and diagnosis of behavioral addictions.

Dan J. Stein, FRCPC, PhD, is Professor and Chair of the Dept of Psychiatry at the University of Cape Town, and Director of the South African Medical Research Council Unit on Risk & Resilience in Mental Disorders. He is interested in the compulsive-impulsive spectrum of disorders, as well as in work at the intersection of global mental health and neuroscience.

Kai W. Müller, Dr., is a Senior Research Fellow and clinician at the Outpatient Clinic for Behavioral Addictions at the University Medical Center Mainz, Germany. His main research focus is on non-substance-related addiction disorders with a special focus on underlying mechanisms of gaming disorder and gambling disorder

Marc N. Potenza, MD, PhD, is an Addiction Psychiatrist and a Professor of Psychiatry at the Yale School of Medicine where he also holds secondary appointments as a Professor of Neuroscience and a Professor at the Child Study Center. He has authored over 600 manuscripts with a focus on addictions and mental health.

Hans-Jürgen Rumpf, PhD, is an Assistant Professor at the Department of Psychiatry and Psychotherapy, University of Lübeck, Germany. He has authored more than 300 peer-reviewed papers on substance use and mental disorders as well as disorders due to behavioral addictions with a focus on epidemiology, assessment, comorbidity and brief interventions.

John B. Saunders MD, FRACP, FRCP is a professor and consultant physician in internal medicine and addiction medicine, with appointments at the Universities of Queensland and

Sydney, Australia. He has a large consulting practice in all types of addictive disorders. He has 360 peer-reviewed publications and his main areas of research are the AUDIT questionnaire, brief interventions, treatment of alcohol, opioid and psychostimulant dependence, and medical education.

Vladan Starcevic, MD, PhD, FRANZCP, is a psychiatrist and associate professor at the University of Sydney, Faculty of Medicine and Health, Sydney Medical School and Nepean Clinical School. He has had a long-standing professional and research interest in anxiety and related disorders, obsessive-compulsive disorders and behavioural addictions. He has authored or co-authored more than 300 peer-reviewed journal articles and book chapters and several books.

Zsolt Demetrovics, PhD, is professor of psychology at the ELTE Eötvös Loránd University, Budapest, Hungary where he serves as director of the Institute of Psychology and head of the Department of Clinical Psychology & Addiction. He has published numerous peer-reviewed papers and book chapters on the epidemiology, assessment and psychological correlates of substance use behavior and behavioral addictions. He is funding Editor-in-Chief of the Journal of Behavioral Addictions.

Matthias Brand, PhD, is a Professor in the Department of Computer Science and Applied Cognitive Science, Faculty of Engineering, University of Duisburg-Essen, Campus Duisburg. He is also director of the Erwin L. Hahn Institute for Magnetic Resonance Imaging, Essen, and head of the Center for Behavioral Addiction Research (CeBAR), Duisburg. He has published numerous peer-reviewed papers and book chapters on the assessment, psychological and neurobiological correlates of addictive behaviors.

Hae Kook Lee, MD, PhD, is Professor of Psychiatry in the College of Medicine, The Catholic University of Korea, Seoul, South Korea. He specializes in Alcohol-related disorders, dementia, Smoking Cessation & Moderate Drinking Clinic, depression, bipolar disorder, schizophrenia, and panic disorder

Marcantonio M. Spada is Professor of Addictive Behaviours and Mental Health at London South Bank University where is Head of the Division of Psychology and Deputy Lead of the Centre for Addictive Behaviours Research. He is the author over 140 peer-reviewed journal articles, several book chapters and three books. Professor Spada also serves as Editor-in-Chief of Addictive Behaviors and as Founding Editor-in-Chief of Addictive Behaviors Reports.

Katajun Lindenberg, PhD, is a Professor for Developmental Psychology and a clinical psychologist at the Institute for Psychology, University of Education, Heidelberg, Germany. Her research focuses on prevention and psychotherapy treatment for mental disorders in children and adolescents, especially Internet-related disorders.

Anise M. S. Wu, PhD, is a Professor at the Department of Psychology, Faculty of Social Sciences, University of Macau. Her research interests are in clinical and health psychology, with a key focus on addictive behaviors, and she has made more than 100 publications in refereed journals.

Tagrid Lemenager, PhD, is a researcher in the Department of Addictive Behavior and Addiction Medicine, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

Ståle Pallesen, PhD, MPsych (Clin), is professor at the Faculty of Psychology, University of Bergen, Norway. He is leader of the Norwegian Competence Center for Gambling and Gaming Research. He has authored more than 300 peer-reviewed papers. His main research interests are sleep and addictions.

Sophia Achab, MD, MSc, PhD, PD, is an executive psychiatrist and lecturer at University of Geneva. She is in charge of WHO Collaborating Center in training and research in Mental Health addiction section and has authored 55 peer-reviewed papers on the topic of addictive behaviours and 9 book chapters on addiction medicine topic.

Susumu Higuchi, MD, PhD, is the Director of the National Hospital Organization Kurihama Medical and Addiction Center in Yokosuka, Japan. He has authored over 350 manuscripts with a focus on addictions and mental health.

Mike Kyrios holds the position of Vice-President and Executive Dean of the College of Education, Psychology and Social Work at Flinders University. His wide-ranging research focuses on: obsessive-compulsive spectrum disorders (OCD, hoarding disorder, body dysmorphic disorder, trichotillomania), behavioural addictions (compulsive buying, problem gambling), depression, anxiety disorders, chronic medical illness, the self in psychological

disorders, development, evaluation and dissemination of evidence-based psychological treatments, including e-therapies.

Naomi A Fineberg, MA, MBBS, MRCPsych, is Professor of Psychiatry at the University of Hertfordshire and Consultant Psychiatrist at Hertfordshire Partnership University NHS Foundation Trust, where she leads the Highly Specialised Obsessive Compulsive Disorders Service. She chairs the H2020 COST Action into advancing the investigation of Problematic internet Usage.

Paul H. Delfabbro, PhD, is a Professor in Psychology in the School of Psychology at the University of Adelaide. He has published extensively in several areas, including the psychology of gambling, child protection and child welfare and has been a regular advisor to State and Federal Government bodies. He has over 300 publications in these areas including over 200 national and international refereed journal articles.