Toward a consensus definition of pathological video-gaming: A systematic review of psychometric assessment tools

Daniel L. King\textsuperscript{a,c,*}
Maria C. Haagsma\textsuperscript{b}
Paul H. Delfabbro\textsuperscript{a}
Michael Gradisar\textsuperscript{c}
Mark D. Griffiths\textsuperscript{d}

\textsuperscript{a}School of Psychology, The University of Adelaide
\textsuperscript{b}Department of Psychology, Health, and Technology, University of Twente
\textsuperscript{c}School of Psychology, Flinders University
\textsuperscript{d}International Gaming Research Unit, Psychology Division, Nottingham Trent University

*Corresponding author. School of Psychology, Level 4, Hughes Building, The University of Adelaide, Adelaide, SA 5005, Australia. Tel: +61 8 83133740; fax: +61 8 8303 3770

Financial Disclosure
This study received financial support from a 2011 Young Researcher Grant funded by the European Association for the Study of Gambling.

Conflict of Interest
The authors declared no conflicts of interest.
ABSTRACT

Pathological video-gaming, or its proposed DSM-V classification of “Internet Use Disorder”, is of increasing interest to scholars and practitioners in allied health disciplines. This systematic review was designed to evaluate the standards in pathological video-gaming instrumentation, according to Cicchetti (1994) and Groth-Marnat’s (2009) criteria and guidelines for sound psychometric assessment. A total of 63 quantitative studies, including eighteen instruments and representing 58,415 participants, were evaluated. Results indicated that reviewed instrumentation may be broadly characterized as inconsistent. Strengths of available measures include: (i) short length and ease of scoring, (ii) excellent internal consistency and convergent validity, and (iii) potentially adequate data for development of standardized norms for adolescent populations. However, key limitations included: (a) inconsistent coverage of core addiction indicators, (b) varying cut-off scores to indicate clinical status, (c) a lack of a temporal dimension, (d) untested or inconsistent dimensionality, and (e) inadequate data on predictive validity and inter-rater reliability. An emerging consensus suggests that pathological video-gaming is commonly defined by (1) withdrawal, (2) loss of control, and (3) conflict. It is concluded that a unified approach to assessment of pathological video-gaming is needed. A synthesis of extant research efforts by meta-analysis may be difficult in the context of several divergent approaches to assessment.

*Keywords*: pathological video-gaming; assessment; clinical diagnosis; systematic review; Internet use disorder; video-gaming addiction; DSM-V
INTRODUCTION

Video-gaming is an increasingly prevalent activity worldwide and has attracted increasing research attention in psychology and psychiatry (Sim et al., 2012; Weinstein & Lejoyeux, 2010). Clinicians and researchers in allied mental health disciplines have proposed that, under certain conditions, video-gaming may become psychologically, socially, and/or physically detrimental to the user (e.g., Kuss & Griffiths, 2012a; Salguero & Moran, 2002). The question of whether a pattern of problematic video-gaming behaviors over a sustained period of time may constitute a psychological disorder is the topic of ongoing debate (Blaszczynski, 2006; King, Delfabbro, & Zajac, 2011; Wood, 2008). On May 1, 2012, the DSM-V Task Force and Work Groups proposed that Internet Use Disorder, which primarily refers to maladaptive video-gaming (or “Internet Gaming”) behavior, should be included in Section III of the DSM-V as the subject of further empirical inquiry. This announcement marked the first occasion of video-gaming being formally recognised as a disorder, albeit tentatively, in clinical nomenclature.

The proposed Internet Use Disorder classification contains nine criteria: (1) preoccupation with Internet gaming; (2) withdrawal symptoms when Internet is taken away; (3) tolerance: the need to spend increasing amounts of time engaged in Internet gaming, (4) unsuccessful attempts to control Internet gaming use; (5) continued excessive Internet use despite knowledge of negative psychosocial problems; (6) loss of interests, previous hobbies, entertainment as a result of, and with the exception of Internet gaming use; (7) use of the Internet gaming to escape or relieve a dysphoric mood; (8) has deceived family members, therapists, or others regarding the amount of Internet gaming; and (9) has jeopardized or lost a significant relationship, job, or educational or career opportunity because of Internet gaming use. Recent
commentaries by King and Delfabbro (2012) and Starcevic (2012) have highlighted that 7 of the 9 criteria specifically refer to “Internet gaming”, whereas the remaining criteria refer to Internet use more generally. Consequently, although this proposed set of criteria was intended to provide greater clarity to the clinical formulation of Internet-related disorders, the diagnostic category may promote further confusion with its conflation of video-gaming and Internet use for other purposes. For the purpose of this review, the term “Internet Use Disorder” (when used) refers to Internet gaming specifically (i.e., pathological video-gaming), rather than general Internet use behaviors.

Two conceptual definitions of pathological video-gaming preceded the Internet Use Disorder classification. These definitions were based on the underlying components of the DSM-IV-TR classifications for substance dependence and impulse control disorder (see Albrecht, Kirschner, & Grüßer, 2007; Sim et al., 2012; Table 1 and 2 also present a list of diagnostic features of each classification). However, the specific constituents of these two diagnostic categories that should be prioritized in conceptualizing pathological video-gaming have been debated. For example, Blaszczynski (2006) has argued that impaired control and harmful consequences should be considered fundamentally important criteria for defining pathological video-gaming. Similarly, Charlton and Danforth (2007) have argued that some features of addictive video-gaming, including cognitive salience, tolerance, and euphoria, may in fact represent peripheral criteria of addiction which may be inappropriate diagnostic features given their overlap with high but otherwise normal engagement with video-gaming activities. Further attempts to classify problematic video-gaming have referred to the amount of time spent in the activity. For example, “excessive” video gaming has been defined by some as more than five hours of play.
per day (Messias et al., 2011), and “dependent” video-gaming as more than 10 hours per week (Huang, 2006). However, as Griffiths (2010) has shown using case studies, the time spent engaged in gaming is not necessarily an indicator of problematic play and that context is critical when using time as criterion for addictive gaming.

An alternative model has proposed six features or components of addictive behavior (Griffiths, 2005). The criteria include: (1) salience, when video-gaming has become the most important activity in a person’s life, dominating thoughts (preoccupation and cognitive distortions), (2) mood modification, which refers to changes in a person’s mood state that occur as a result of playing video-games; (3) tolerance, which refers to the process whereby increasing amounts of video-game play are required to achieve the former mood-modifying effects, (4) withdrawal, which refers to aversive mood states and/or physical effects that occur when video-game play is suddenly discontinued or reduced, (5) relapse, which refers to the tendency for the player to revert back to earlier patterns of video-game play, and for even the most extreme patterns, typical of the height of excessive video-game play, to be restored quickly after periods of abstinence or moderation, and (6) harm, which refers to the negative consequences of excessive video game play, including personal psychological distress as well as conflicts with other people (family members and friends) and/or other activities (job, school, social life, hobbies and interests).

There has been increasing sophistication in relation to issues concerning assessment and measurement of pathological video-gaming. In the last few years, instruments have been developed that have more robust psychometric properties in terms of reliability and validity. However, there are still some concerns as many of the most widely used screening instruments were adapted from adult screens and much of the video game literature has examined children and adolescents. King, Delfabbro,
Griffiths, and Gradisar (2011) assert that, to enable future advances in the development and testing of interventions for video game-related problems, there must be some consensus among clinicians and researchers as to the precise classification of these problems.

Available evidence suggests that, internationally, a significant number of individuals with video-gaming-related problems have received some form of treatment from a mental health or medical service provider (Baer, Bogusz, & Green, 2011; Han et al., 2009). This is particularly evident in South East Asia (e.g., China, Taiwan and South Korea), where the estimated prevalence of technology-related problems among adolescents ranges from 0.3% to over 10% (King et al., 2012). Several clinical trials of psychological and pharmacological treatment have already been conducted; however, meaningful comparison of treatment outcomes has been difficult without standard assessment protocols (King, Delfabbro, Griffiths, et al., 2011). Therefore, there exists a need for consensus on measurement of pathological video-gaming for consistent assessment treatment outcomes of current and future intervention studies.

The qualities of pathological video-gaming assessment tools that should be prioritized may vary in public health research studies as compared to clinical practice. For epidemiological purposes, Koronczai et al. (2011) claimed that the most appropriate measures for assessing pathological internet use (including online gaming) should meet the following six requirements. The instrument should have: (i) brevity (to facilitate incorporation into time-limited surveys); (ii) comprehensiveness (to examine as many – if not all – aspects of pathological video-gaming as possible); (iii) reliability and validity for different age groups (e.g., adolescents and adults); (iv) reliability and validity for any method of data collection (e.g., paper-and-pencil,
online, face-to-face); (v) cross-cultural reliability and validity; and (vi) been validated on clinical samples. They also recommended that an ideal assessment instrument should also serve as a basis for defining cut-off scores for dependence.

Clinicians are often faced with the critical task of selecting the most appropriate available psychometric tool for an assessment of various disorders in childhood, adolescence, or adulthood. This task is challenging in the field of pathological video-gaming, given the numerous clinical formulations and assessment tools that have emerged in empirical research and clinical intervention studies (King, Delfabbro, Griffiths, & Gradisar, 2011; 2012). Given these varying definitions and clinical indicators of pathological video-gaming, it is perhaps predictable that a growing number of different assessment approaches and tools have been developed. Some researchers have also attempted to combine theoretical models of pathological video-gaming by developing composite instruments, i.e., measures based on a combination of selected items from multiple instruments (see Chou & Ting, 2003; Smahel, Blinka, & Ledabyl, 2008).

The primary aim of this systematic review was to summarize and critique available research evidence on clinical assessment of pathological video-gaming. This review was designed to determine the strengths and limitations of instruments according to accepted standards in psychological assessment (Cicchetti, 1994; Groth-Marnat, 2009). Although several recent reviews of the pathological video-gaming literature have highlighted limitations in regard to clinical conceptualization (Sim et al., 2012), etiology and risk factors (Kuss & Griffiths, 2012a), and quality of intervention studies (King, Delfabbro, Griffiths, et al., 2011), this review is unique due to its focus on psychometric tools. No previous reviews have focussed specifically on the quality of diagnostic instruments. Therefore, it was intended that
this review may assist researchers and practitioners in the selection and use of assessment tools for pathological video-gaming, and guide future research endeavours toward refining measurement of the disorder. In addressing issues of measurement of pathological video-gaming, this review also intended to describe and critique the emerging consensus on the classification (i.e., diagnostic features) of the disorder.

METHOD

Study selection

A computer database search of Academic Search Premier, PubMed, PsychINFO, ScienceDirect, and Web of Science was conducted, using the following search terms and logic: (patholog* OR problem* OR addict* OR compulsive OR dependen*) AND (video OR computer) gam*. All searches were limited to full text papers published from 2000 to 2012 because studies conducted in this era of “online gaming”\(^1\) were reasoned to be most relevant to the DSM-V category of Internet Use Disorder. These database search parameters yielded a total of 4,120 hits, which included the following results in each database: Academic Search Premier (967 results), PubMed (235 results), PsychINFO (957 results), ScienceDirect (1,677 results), and Web of Science (284 results). The reference lists of systematic reviews of pathological video-gaming were also examined (i.e., Ferguson, Coulson, & Barnett, 2011; Kuss & Griffiths, 2012a, 2012b; Sim et al., 2012), as well as the references of the included studies.

\(^1\) From 1999, video gaming had expanded significantly into the online medium where games could be played as part of a gaming community, with the notable emergence of Massively Multiplayer Online Role Playing Games (MMORPGs) (e.g., Everquest [1999], Ultima Online [1997], and Asheron’s Call [1999]) (Griffiths, Kuss, & King, 2012).
Studies were selected on the basis of employing a psychometric instrument to assess pathological video-gaming. The purpose of this study was to conduct an exhaustive review of the standards in assessment of all published quantitative research on pathological gaming. Therefore, studies were not necessarily excluded on the basis of methodological shortcomings, such as low sample size. However, a study was excluded if: (i) the name of the instrument employed was not identified, (ii) a composite measure of pathological video-gaming was employed, (iii) case note material or anecdotal evidence concerning pathological video game use was presented only, or (iv) it was not published in English or Dutch. Studies that employed instruments to assess Internet addiction were included if at least 50% of the participants’ Internet use involved video-gaming. This approach was taken to prevent the exclusion of studies that focussed primarily on pathological video-gaming, but that also assessed other forms of Internet use. Table 1 presents a summary of the 18 instruments, which were employed across 63 studies, identified for selection by this process of review.

2.2 Instrument assessment

A review framework for evaluating instruments within identified studies was developed by adapting standards and guidelines for psychometric assessment. Specifically, the framework was modelled on 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 primary goal of the review framework was to assess theoretical, psychometric, and practical aspects of the instruments. Relevant psychometrics included: components (i.e., underlying theoretical construct),
dimensionality, validity (convergent, predictive, and criterion), reliability (internal consistency, inter-rater), and availability of normative and prevalence data. Practical considerations included: number of items, administration time, reading age level, item sensitivity, time-scale, diagnostic categories, country of origin, and language versions.

All included studies were independently coded by the first two authors so that information regarding relevant characteristics of each instrument could be extracted. The coding method involved each reviewer reading identified studies and coding all information applicable to the review framework. To demonstrate this coding method, for the framework item “Reliability: internal consistency”, the first two authors read the methods and results sections of each paper in search of a Cronbach’s alpha value. A code of “NA” (Not Assessed) was given if authors did not refer to internal consistency at all, or a code of “NR” (Not Reported) was given if the paper referred to internal consistency but did not report a statistical coefficient. In instances where authors could have but did not provide relevant test information, a code of NR was generally employed. Identified characteristics of each study were then discussed and systematically entered into a computer database using Microsoft Excel© 2010. Misunderstandings and/or discrepancies in coding occurred for five papers, and were resolved by consultation and consensus among the first two authors.
RESULTS

Table 1 presents a summary of key characteristics of the 18 pathological videogaming psychometric tools initially identified by review. Although only six instruments had a corresponding evidence base of more than two empirical studies, all 18 instruments were included for analysis of their psychometric properties and practical considerations. Tables 2 and 3 present a summary of these analyses.

[INSERT TABLE 1]

Theoretical orientation

A clinician’s choice of psychometric instrument is guided initially by consideration of available tests’ theoretical orientation (Groth-Marnat, 2009). The question is raised: Do the test items correspond to the theoretical description of the construct? A total of 16 behavioral addiction indicators were identified to assess the theoretical description of each reviewed instrument. As noted in the Introduction, six frequently described (i.e., “core”) components of behavioral addiction include: salience, euphoria, withdrawal, tolerance, relapse, and conflict. Conflict was conceptualised along six domains: work/school, household duties, relationships, sleep, finances, and illegal behaviors. Additional indicators based on the DSM-IV-TR classifications of substance dependence and pathological gambling (i.e., those not already assessed by the components model) were added to this list of indicators. These included: impaired control, escape (i.e., using the activity to escape from problems or relieving unwanted mood states), dependency on others for a financial
bailout, deception (i.e., secrecy and/or lies about use), and continued use despite negative consequences.

Table 2 presents a summary of the 18 instruments’ capacity to assess common indicators of behavioral addiction. The proposed DSM-V criteria for Internet Use Disorder were included in this analysis for the purpose of comparison. Two general observations of the instrument base were made initially: (1) no two instruments were alike with respect to their yielded profile of diagnostic features, and (2) interpersonal conflict was the only addiction indicator that was assessed across all 18 instruments.

Further inspection revealed that dimensions of conflict or harm arising from excessive video-gaming were not assessed consistently across the instruments. With the exception of the adapted DSM-IV-TR pathological gambling criteria and the KIAS, commitment of illegal acts to finance video-gaming or a need of a financial bailout by others were not assessed by instruments. Similarly, only two instruments (i.e., adapted DSM-IV-TR pathological gambling criteria and OAST) examined negative financial consequences associated with video-gaming. Negative sleep consequences associated with video-gaming were assessed by four instruments only (i.e., CIUS, GAS, VAT, and YIAT). Continued use of a video-game despite knowledge of adverse consequences, an item that is unique to the DSM-IV-TR substance dependence classification, was assessed only by the PVP Scale.

Total coverage of addiction indicators across instruments was first assessed. Cognitive salience (n=14), withdrawal symptoms (n=17), loss of control (n=14), and conflict associated with interpersonal relationships (n=18), and work and/or school commitments (n=14) were the most common diagnostic features assessed by all instruments. Overall, the adapted DSM-IV-TR pathological gambling criteria demonstrated the greatest coverage of addiction indicators, by assessing 13 out of 16
The two instruments that provided the most clinical information (in terms of raw data) about pathological video-gaming symptoms were the Game Addiction Scale (GAS) and the Young Internet Addiction Test (YIAT), each of which assessed 11 addiction indicators.

Each instrument’s “cross-over” (i.e., the number of overlapping addiction indicators) with the components model of addiction and/or the adapted DSM-IV-TR classifications of pathological gambling and substance dependence was then considered. Notably, the adapted DSM-IV-TR pathological gambling criteria, the Addiction-Engagement Questionnaire (and Engagement-Addiction Questionnaire), and the GAS were the only instruments capable of assessing all six “core” components of Griffiths’ (2005) behavioral addiction model. In contrast, none of the instruments were able to assess all criteria of the adapted DSM-IV-TR classifications of pathological gambling and substance dependence. However, of the four instruments capable of assessing the components model of addiction, those most capable of also assessing the addiction indicators within the DSM-IV-TR classification of pathological gambling were the GAS (with 10 out of 13, indicating 77% cross-over) and the YIAT (9 out of 13; 69%).

Finally, each instrument’s capacity to assess the proposed DSM-V classification of Internet Use Disorder was examined. Only one instrument, the PVP Scale, demonstrated this capacity, although the GAS was capable of assessing 8 of the 9 criteria. The diagnostic features of Internet Use Disorder most commonly assessed by available instruments included relationship conflict \((n=18)\), work or school conflict \((n=14)\), withdrawal \((n=17)\), and loss of control \((n=14)\). Conversely, the Internet Use Disorder criteria that were most frequently not assessed by instruments was
“continued use despite knowledge of problems” ($N=2$), followed by “deception of others about the amount of time spent Internet gaming” ($N=5$).

[INSERT TABLE 2]

Validity

Three types of instrument validity were assessed: convergent, predictive and criterion. Convergent validity was defined as the extent to which scores on each instrument had expected or hypothesised relationships to relevant variables (DeVellis, 1991). Although pathological video-gaming represents a relatively new field of clinical psychology, research evidence has linked the disorder to aggressive behavior (Anderson et al., 2010), attention problems (Swing et al., 2010), Axis I disorders (Mentzoni et al., 2011), poor academic achievement (Smyth, 2007), reduced empathy (Bartholomew et al., 2005), and impaired social functioning (Gentile et al., 2011). These factors, in addition to factors of relevance to addiction (e.g., impulsivity, sensation-seeking), were evaluated. Four instruments (i.e., adapted DSM-IV-TR pathological gambling criteria, GAS, PVP Scale, and YIAT) demonstrated strong convergent validity by having statistically significant associations with seven or more clinical indicators. The POGU and KFN-CSAS-II each had five indicators, although it should be noted that the KFN-CSAS-II has been tested on the largest population ($N>15,000$) of all 18 instruments.

Predictive validity was defined according to an instruments’ ability to predict pathological video-gaming status over time. The dearth of longitudinal studies in the field of pathological video-gaming has been well-documented (King, Delfabbro, Griffiths, et al., 2011). Accordingly, only two of the 18 instruments (i.e., the CUIS
and the adapted DSM-IV-TR pathological gambling criteria) had relevant research data. Of these two instruments, the adapted DSM-IV-TR pathological gambling criteria demonstrated superior predictive validity in a comparable adolescent population and over a longer time-frame as compared to the CIUS. However, caution is advised in drawing firm conclusions based on these limited data.

Criterion validity was defined as the positive association between severity and/or number of pathological video-gaming symptoms and time spent playing video games. Although Gentile (2009) has reported that this relationship may not be simply isomorphic, it has been generally observed that frequent and repetitive video-gaming typically increases risk of problem gaming behavior. Scores on seven reviewed instruments were significantly positively related (at alpha level of <.01) to greater time spent playing video games. Larger bivariate associations (i.e., >.40) were observed among studies employing the GAS or the KFN-CSAS-II, PVP Scale, and the YIAT. The VAT and CUIS demonstrated moderately lower (but still significant) associations. No data were available for 11 instruments.

**Reliability**

Instrument reliability was assessed according to ratings of internal consistency (IC) and inter-rater reliability. According to Cicchetti (1994), an alpha coefficient of .70 to .79 may be considered “fair”; an alpha of .80 to .89 is “good”; and an alpha of .90 or higher is “excellent”. Groth-Marnat (2009) recommended that reliability estimates should be .90 for clinical decision-making and .70 for research purposes. Cronbach’s alpha coefficients for all instruments generally fell into a range of .70 to .90. Overall, the YIAT demonstrated the most consistently high internal consistency, with the highest number of studies (n=7) reporting alpha values of .90 or higher.
Notably, the largest study of pathological video-gaming, conducted by Rehbein et al. (2010) and employing the KFN-CSAS-II, demonstrated a high IC (.92). Although the adapted DSM-IV-TR pathological gambling criteria had high IC (.92) in one study (Lemmens et al., 2009), its internal consistency generally fell within the .70 to .80 range. This may be partly explained by the relatively lower inter-item correlations for items concerning financial conflict and illegal acts, as compared to other test items (Gentile, 2009). The research literature for 7 out of 18 instruments did not provide measures of internal consistency.

Inter-rater reliability was referred to in only one of the reviewed studies (i.e., Bioulac et al., 2008), which employed a parent-report version of the instrument but the authors did not provide accompanying statistical information. Independent ratings by psychiatrists were obtained in three studies (Dong et al., 2011; Han et al., 2007; Pawlikowski & Brand, 2011), although psychiatric evaluation was performed to determine broad clinical status rather than to verify instrument ratings. Test-retest reliability was not reported in any of the 63 reviewed studies. Discriminant validity was assessed in only one study (Gentile, 2009).

In total, 16 studies have examined test dimensionality of 12 instruments using factor analysis techniques. Four instruments (i.e., GAS, PVP Scale, ISS-20 and YIAT) demonstrated a single-factor structure, with that factor termed “addiction”. However, two of these instruments (i.e., GAS and YIAT) have also demonstrated alternative factor structures in other studies (Chang & Law, 2008; Lemmens et al., 2009). Only the PVP Scale (Salguero & Moran, 2002) demonstrated a factor structure that had been confirmed in a follow-up study (Hart et al., 2009) in the absence of contrasting findings. Among those instruments with more than two published studies, the CUIS was the only test of which dimensionality had not been assessed.
Standardization

A central issue of standardization relates to the adequacy of norms. The basis on which individual test scores have meaning relates directly to the similarity between the individual being tested and the sample (Groth-Marnat, 2009). All 18 reviewed instruments lack a user manual reporting standardized norms. Accordingly, some researchers have employed their own statistical approaches to classify study participants as normal or pathological video-gamers. For example, Stetina et al. (2011) reported that scores on the Problematic Internet Use Scale that fell on or beyond the 88th percentile were indicative of “problematic” video-gaming. However, a purely statistical cut-off score may not be an optimal indicator of clinical significance, particularly in small samples or non-normal distributions, given the likelihood of under- or over-representing clinical cases (McCluskey & Lalkhen, 2007).

A total of 58,415 individuals have been assessed by the 18 reviewed pathological video-gaming instruments. The majority of participants have been assessed by the KFN-CSAS-II \( (n=15,168) \), followed by the adapted DSM-IV-TR pathological gambling criteria \( (n=9,995) \), the CIAT \( (n=7,874) \), the PVP Scale \( (n=4,988) \), the CUIS \( (n=3,744) \), the POGQ \( (n=3,415) \), and the GAS \( (n=3,413) \). The availability of “large” population studies of adolescents (i.e., \( N>1,500 \)) using the KFN-CSAS-II \( (n=15,168: \text{Rehbein et al.}, 2010) \), the adapted DSM-IV-TR pathological gambling criteria (i.e., \( n=2,998: \text{Choo et al.}, 2010; n=3,034; \text{Gentile et
al., 2011; \(n=1,945\) and the CUIS (i.e., \(n=1,572\): Van Rooij et al., 2010) suggest potential for development of norms for these measures.

The sampling procedure should also be taken into account in evaluating psychometric data yielded from large survey studies. Non-representative populations may vary significantly from representative populations (e.g., according to gender, age, socioeconomic status, type of gaming activity, and various psychosocial variables including personality, temperament, etc.). A non-representative sample therefore has the potential to bias test item responses, thereby affecting many psychometric attributes of an instrument. One way in which sample representativeness may be compromised is by participant self-selection. Self-selection was a commonly observed sampling method. In 13 of the 63 reviewed studies (i.e., Achab et al., 2011; Billieux et al., 2011; Charlton & Danforth, 2007; Choi et al., 2008; Collins et al., 2012; Hussain & Griffiths, 2009; Khazeel et al., 2008; Kim et al., 2008; King, Delfabbro, & Zajac, 2011; Lafreniere et al., 2009; Meerkerk et al., 2010; Porter et al., 2011; Stetina et al., 2011), participants were recruited by an advertisement posted in either a real-life or online video-gaming context (e.g., Internet café, or online gaming community website). This recruitment method may have excluded some segments of the general population, thereby limiting capacity to develop meaningful norms from obtained data.

**Practical considerations**

A clinician’s choice of instrumentation is often guided by various practical constraints, such as ease of use and administration time. Notably, 15 out of 18 instruments contained 20 or fewer items, and thus could be completed by an assessor or by self-report in less than 10 minutes. Although all instruments had been employed
in research on adolescents, only eight instruments had a specified reading age level. Two instruments were reported as being appropriate for those aged 8 years and over (i.e., adapted DSM-IV-TR pathological gambling criteria and Addiction-Engagement Questionnaire), whereas five measures specified a minimum age of 12 or 13 years (i.e., GAS, OAST, PVP Scale, VAT, and Video Game Dependency Scale). The majority of instruments (n=11) employed continuous rather than dichotomous item-response categories to increase sensitivity. However, only four instruments specified a time-scale for test items. Information on scoring and interpretation for each measure also varied. Three instruments (i.e., adapted DSM-IV-TR pathological gambling criteria, PVP Scale, and YIAT) had inconsistent cut-off scores across studies to indicate clinical status. For example, five different cut-off scores were reported across 11 studies that employed the YIAT. Similarly, two cut-off scores on the adapted DSM-IV-TR pathological gambling criteria have been employed, and three studies did not specify a cut-off score (Lemmens et al., 2006; Li et al., 2011; Topor et al., 2011). Prevalence rates of pathological video-gaming appeared to vary according to instrument used, although it was difficult to control for the possible influence of region (i.e., differences in Eastern versus Western populations). In considering only those instruments with multiple (i.e., >2) studies, it was found that studies that employed the adapted DSM-IV-TR pathological gambling criteria tended to report prevalence rates around twice as high (i.e., 7-8%) as studies that employed other measures (i.e., 1-4%). The majority of instruments were developed in Europe (n=10) or the United States (n=4), and seven instruments were published in English only.
DISCUSSION

The primary aim of this study was to conduct a systematic review of current instrumentation for pathological video-gaming, and to quantify the psychometric properties and practical considerations of these instruments. A secondary aim was in line with the proposed revision to Section III of the DSM-V (i.e., to identify research avenues in service of refining the conceptualization and assessment of pathological video-gaming). Overall, the results indicated that available pathological video-gaming instrumentation may be broadly characterised as inconsistent. Although 63 quantitative research studies, representing over 50,000 participants, have been conducted since the year 2000, multiple inconsistencies in assessment raise significant concerns about the potential for synthesis and meta-analysis of research findings. Strengths of the instruments used to date include their short length, ease of scoring and administration, and excellent internal consistency and convergent validity.

Defining pathological video-gaming: The emerging consensus

This analysis examined the theoretical orientation of pathological video-gaming instruments, including their alignment with several models of behavioral addiction (i.e., the DSM-IV-TR classifications of substance dependence and pathological gambling, the components model, and the proposed DSM-V Internet Use Disorder). Four instruments were capable of assessing the components model. However, none of the reviewed instruments were capable of assessing the DSM-IV-TR classifications. Only one of the 18 instruments was capable of assessing Internet Use Disorder. This finding suggests that, although the conceptual development of instruments has been purportedly based on established addiction criteria, in practice very few instruments
are grounded in this way. Disconcertingly, no two instruments are alike in their theoretical orientation and ability to ‘map out’ diagnostic features of problem video-gaming behaviour.

If there is no clear consensus on which model of addiction applies to videogaming behavior, then the next question is whether there is some agreement on the diagnostic features of most importance in defining pathological video-gaming. By adopting a utilitarian approach to defining the disorder (i.e., consensus decision-making), the majority of instruments assessed three features: (1) withdrawal, (2) loss of control, and (3) conflict associated with interpersonal relationships, work and/or school commitments. Notably, these characteristics align with criteria B, C, and D for the proposed DSM-V Internet Use Disorder. These three features are also present within the DSM-IV-TR classifications and the components model of addictive behaviour, with the exception of loss of control that does not specifically appear in the components model (as it is subsumed within the ‘conflict’ criterion). On this basis, these features may be put forward as the “core” features of pathological video-gaming. All other features may be considered “peripheral” but not essential to indicate a diagnosis of pathological video-gaming.

To make genuine advances in the conceptualization and measurement of pathological video-gaming, researchers and clinicians must collectively determine which of the aforementioned approaches to adopt. In simple terms, the pathological video-gaming field is at a cross-road, and now must decide between multiple classification approaches. Further research efforts that empirically compare and contrast the utility of the Internet Use Disorder classification alongside the DSM-IV-TR classifications may be necessary to guide this process. Similarly, studies employing multiple instruments may be needed to determine relative psychometric
strengths, including diagnostic sensitivity. On the basis of available evidence, this review suggests that: (1) the Problem Videogame Playing (PVP) Scale may provide the best overall measure of Internet Use Disorder, and (2) the adapted DSM-IV-TR pathological gambling criteria, the GAS and the YIAT may provide the most clinical information (notwithstanding the unique drawbacks of each test). Researchers or clinicians wishing to prioritize particular psychometric features of instrumentation (e.g., predictive validity) are likely to find Table 3 an invaluable resource.

Improving measurement of pathological video-gaming

This review has identified multiple practical ways in which pathological video-gaming assessment procedures may be improved. Firstly, given that no single instrument provides broad coverage of all common addiction indicators (or diagnostic features) within DSM-IV-TR classifications of pathological gambling and substance dependence, it is recommended that research studies employ more than one instrument for greater scope of clinical information. Second, many instruments may be improved by the addition of two items: (i) an item that assesses whether the individual personally believes that their video-gaming behavior is problematic (i.e., a validity check), particularly as psychometric evidence has shown good correlations between a single self-report item of problematic Internet use and standardized measures of problematic internet use (Widyanto, Griffiths & Brunsden, 2011), and (ii) an item that asks whether significant others in their life would consider that their video-gaming is problematic (i.e., a reliability check). Third, in the absence of psychiatric evaluation to confirm diagnosis and screen for a host of co-morbidities (e.g., psychopathology, pain, trauma, and substance abuse), studies employing pathological video-gaming instruments may benefit from items to guide differential
diagnosis, such as symptomology indicative of obsessive-compulsive disorder and/or a manic episode.

Each instrument would benefit from the development of a user manual, with clear guidelines on scoring and interpretation. Inclusion of time-scales for each item (i.e., the period of time that the item refers to) as well as time of onset is also needed for most instruments, to improve demarcation of current versus historical symptoms of pathological video-gaming. Assessment of whether video-gaming activity occurs online or offline may also provide useful information about the context of the pathological behavior. Notably, the Internet Use Disorder in its current format does not account for problem video-gaming behavior that occurs in the absence of online play (King & Delfabbro, 2012). This omission may need to be revised to encompass a range of video-gaming behaviors (e.g., offline video-gaming, including video-gaming on a portable device [Griffiths, Kuss & King, 2012]).

The concept of harm seems to be a problematic construct within pathological video-gaming. First, it is not clear which types of harm are most indicative of pathological video-gaming, although this review suggests that relationship conflict may be the most prominent type. It may be speculated that a focus on interpersonal (rather than intra-personal) conflict alone may not identify severe cases where social contact has become so disrupted and/or avoided that gaming no longer produces conflict per se. Second, there is variability in terms of which life domains are assessed for potential harm caused by video-gaming. Most instruments assess conflict with work or education, and interpersonal relationships. Although there is a substantial body of research on the negative impact of electronic media on sleep (e.g., Cain & Gradisar, 2010; Eggermont & Van den Bulck, 2006; Schocat et al., 2010), negative sleep outcomes associated with video-gaming were assessed by only 5 of the 18
instruments. Additionally, most instruments did not assess financial conflict. It may be that video-gaming, even at high levels of engagement, has significantly lower associated financial costs aside from the initial expense of acquiring video-gaming equipment and software. The proposed Internet Use Disorder notably excludes pathological gambling-related items on financial difficulties and the need for a bailout from significant others.

**Future research directions**

Several research avenues may lead to improvement in assessment. First, there is a need for overall improvement in standards in design and reporting in empirical research, in line with JARS reporting standards. A similar observation has been made of the clinical intervention literature in the related field of Internet addiction (King, Delfabbro, Griffiths, et al., 2011). Although approximately one-third of the studies have recruited adolescents via secondary schools, the majority of studies involving adults have recruited online, self-selected samples, often with limited inclusion criteria. Given the threat of non-representative data to the psychometric profile of instruments, future studies should employ random selection methods to recruit from the general population in order to reduce selection bias (Wilkinson, 1999). Such research would also aid in the development of manuals reporting standardized norms for each instrument. Another issue is that prevalence rates of pathological video-gaming appear to vary significantly, and the extent to which instrumentation (versus true regional differences) may account for this variation is not clear. Finally, longitudinal research designs would provide needed data on predictive validity of instruments, in addition to fulfilling the broader aim of improving current knowledge of the course and severity of pathological video-gaming.
There is also a need for assessment of specificity and sensitivity of instruments (Demetrovics, Urbán, Nagygyörgy, et al. 2012). When identifying clinical cases for administration of treatment in an inpatient setting, high specificity and lower sensitivity may be prioritized. A converse approach may be more appropriate in epidemiological research. Currently it is not clear which instruments are better suited to serve these competing objectives. Given the range of addiction indicators employed across instruments, it may be worthwhile for studies to employ item-response theory or Rasch analysis to identify test items most indicative of levels of severity. For instruments with multiple published cut-off scores (e.g., the YIAT), this approach may be particularly helpful. Similarly, addiction test indicators that appear infrequently in assessment (e.g., committing illegal acts to finance video-gaming) may be eliminated from the pool of diagnostic features of pathological video-gaming. Charlton and Danforth (2007) have reported that some criteria of addiction may also indicate “high engagement” (i.e., increase sensitivity), therefore further examination of high engagement behavior may also contribute to the research agenda.

Limitations of the review

The method employed to review each instrument and its corresponding evidence base was based on guidelines and accepted standards in assessment (Cicchetti, 1994; Groth-Marnat, 2009). Information of most relevance to clinicians was prioritized for inclusion. However, the task of reviewing every aspect of all the available instrumentation is unnecessarily onerous and beyond the scope of the paper. This review should be considered as only a starting point for further conceptual development of any proposed pathological video-gaming disorder (including Internet Use Disorder). The key performance indicators employed in this review framework
may not be most comprehensive method of evaluating pathological video-gaming instrumentation. For example, this review did not examine how instruments differ in their assessment of subclinical or borderline cases according to different populations. Practical constraints also prevented the quantification of details for every possible indicator of instrument quality (e.g., the relative magnitude of correlations for convergent validity indicators across instruments, weighted by sample size). However, relevant citations are provided for all instruments to enable interested scholars and/or clinicians to refer to the original article material for the purpose of extending analysis presented here. Finally, it should again be noted that the databases used to identify reviewed studies may not have identified those studies published in South East Asian journals, although this is a common limitation of systematic reviews.

Conclusion

The proposed DSM-V category of Internet Use Disorder has signalled a period of transition for the field of pathological video-gaming. For decades, pathological video-gaming, or video-game “addiction”, has been based conceptually on the DSM-IV-TR criteria for pathological gambling and/or substance dependence. Inconsistency of theoretical orientation has led to the development of multiple instruments, which has caused an unfortunate divide in clinical research efforts. A groundswell in basic and applied research has placed an emphasis on instruments’ adaptability for use in survey studies of adolescents. This review has indicated that several psychometric properties of assessment are in need of further refinement. Standardization is particularly important in the treatment literature on pathological video-gaming (King, Delfabbro, Griffiths, & Gradisar, 2012), as idiosyncratic assessment may create major
difficulties in evaluating the relative and absolute efficacy of treatment. In summary, the field of pathological video-gaming would benefit from discontinuing use of multiple conceptually divergent instruments, and applying a unified approach to assessment. Whether this assessment approach should be based on the emerging, research-driven consensus on the classification of pathological video-gaming, or on the diagnostic features within the proposed DSM-V criteria for Internet Use Disorder, remains an issue for further debate.
REFERENCES

References marked with a cross (†) indicate studies included in the review.


psychometric characteristics. *Cyberpsychology, Behavior, & Social Networking*, in press.


