



Measurement and Conceptualization of Gaming Disorder According to the World Health Organization Framework: the Development of the Gaming Disorder Test

Halley M. Pontes¹  · Bruno Schivinski² · Cornelia Sindermann³ · Mei Li⁴ · Benjamin Becker⁵ · Min Zhou⁶ · Christian Montag^{3,5}

Published online: 3 June 2019

© The Author(s) 2019

Abstract

Previous research on gaming disorder (GD) has highlighted key methodological and conceptual hindrances stemming from the heterogeneity of nomenclature and the use of non-standardized psychometric tools to assess this phenomenon. The recent recognition of GD as an official mental health disorder and behavioral addiction by the World Health Organization (WHO) in the 11th Revision of the International Classification of Diseases (ICD-11) opens up new possibilities to investigate further the psychosocial and mental health implications due to excessive and disordered gaming. However, before further research on GD can be conducted in a reliable way and within a robust cross-cultural context, a valid and reliable standardized psychometric tool to assess the construct as defined by the WHO should be developed. The aim of this study was to develop The Gaming Disorder Test (GDT), a brief four-item measure to assess GD and to further explore its psychometric properties. A sample of 236 Chinese (47% male, mean age 19.22 years, SD = 1.57) and 324 British (49.4% male, mean age 26.74 years, SD = 7.88) gamers was recruited online. Construct validity of the GDT was examined via factorial validity, nomological validity, alongside convergent and discriminant validity. Concurrent validity was also examined using the Internet Gaming Disorder Scale—Short-Form (IGDS9-SF). Finally, reliability indicators involving the Cronbach's alpha and composite reliability coefficients were estimated. Overall, the results indicated that GDT is best conceptualized within a single-factor structure. Additionally, the four items of the GDT are valid, reliable, and proved to be highly suitable for measuring GD within a cross-cultural context.

Keywords Videogame addiction · Gaming disorder · ICD-11 · GDT · Behavioral addictions

✉ Halley M. Pontes
contactme@halleypontes.com

Prior to the publication of the fifth revision of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) by the American Psychiatric Association in May 2013 (American Psychiatric Association 2013), researchers and clinicians were unclear about the core diagnostic criteria of the phenomenon gaming disorder (GD) (Griffiths et al. 2015). The adoption of inconsistent terminologies and non-standardized assessment tools by previous research investigating GD (see King et al. 2013) has led to several debates among scholars as to whether the phenomenon represents a unique clinical entity worth being officially recognized as a behavioral addiction (Griffiths and Pontes 2014; Pontes and Griffiths 2014, 2015c).

More recently, scholars have argued that formal recognition of GD as a disorder may result in potentially negative medical, scientific, public health, societal, and rights-based repercussions that should be acknowledged (Aarseth et al. 2017). These concerns often relate to the challenges in the psychometric and clinical assessment of the phenomenon as accurate identification of pathological and non-pathological behavior and actual illness has long been a problem in psychiatric epidemiology, often leading to false-positive diagnoses with significant economic and societal implications (van Rooij et al. 2018). Thus, for some scholars, caution is warranted as the new proposed disorder arguably lacks the necessary scientific support and sufficient clinical utility to justify its medical recognition. Notwithstanding this, other scholars have argued against this view and in favor of legitimating GD as a bona fide disorder. For example, Griffiths et al. (2017) suggested that although GD is a rare phenomenon due to its low prevalence rates, dismissing its clinical significance and the individual impact that excessive gaming can have on overall health may inevitably lead to a number of detrimental outcomes. There are several reasons supporting formalizing GD as a psychiatric disorder. For example, by recognizing GD as an official mental health disorder, new social policy can help the development of better insurance and treatment providers able to offer specialized and efficacious treatments for patients suffering from GD (Griffiths et al. 2017). Although different perspectives can be taken on the merits of recognizing GD, it is the view of the present authors that the extant empirical and clinical research warrants recognition of GD as a mental health disorder that affects a minority of gamers. Hence, future research should aim to provide answers related to how GD can be better understood regarding its prevention and which risk factors may contribute to shedding light on how it develops in the general population. Consequently, developing an updated psychometric assessment tools based on official diagnostic framework will likely lead to greater consistency at the diagnostic level and provide robust parameters for replication in research.

Furthermore, with the publication of the DSM-5 in May 2013, “Internet Gaming Disorder” (IGD) was included in Section III (“Emerging Measures and Models”) of the DSM-5 as a tentative disorder requiring additional research before possible formal recognition could be achieved in future revisions of the DSM (Petry and O'Brien 2013; Pontes and Griffiths 2014). According to the American Psychiatric Association (2013), the clinical diagnosis of IGD comprises a behavioral pattern encompassing persistent and recurrent use of the Internet to engage in online games, leading to significant impairment or distress over a period of 12 months as indicated by endorsing five (or more) of nine criteria. The proposed diagnostic criteria for IGD in the DSM-5 include the following: (1) preoccupation with games; (2) withdrawal symptoms when gaming is taken away; (3) tolerance, resulting in the need to spend increasing amounts of time engaged in games; (4) unsuccessful attempts to control participation in games; (5) loss of interest in previous hobbies and entertainment as a result of, and with the exception of, games; (6) continued excessive use of games despite knowledge of psychosocial problems; (7) deceiving family members, therapists, or others regarding the

amount of gaming; (8) use of games to escape or relieve negative moods; and (9) jeopardizing or losing a significant relationship, job, or education or career opportunity because of participation in games. When the severity of IGD is greatly increased, it may lead to academic failure, job loss, or marriage failure as the problem behavior tends to displace usual and expected social, work and/or educational, relationship, and family activities (American Psychiatric Association 2013).

Due to the initial recognition of IGD in the DSM-5, the focus of the scholarly debate has shifted from questioning whether IGD merited formal recognition as a mental health disorder to highlighting the need to further understand the core experiences of IGD from a broad biological, psychological, and social perspective (Pontes 2018). As a result, researchers in the gaming studies field have suggested the need of additional empirical evidence to help identify the defining features of IGD, obtain cross-cultural data on reliability and validity of specific diagnostic criteria, determine prevalence rates in representative epidemiological samples in countries around the world, evaluate its natural history, and examine its associated biological features (Petry and O'Brien 2013; Pontes and Griffiths 2014).

The call for unification in the field (Griffiths et al. 2014, 2016) has led to an increase in the research being conducted on IGD (Pontes and Griffiths 2015a), with several psychometric assessment tools being developed under the IGD framework defined by the American Psychiatric Association (e.g., Király et al. 2017; Lemmens et al. 2015; Pontes and Griffiths 2015b) as reported in a recent review (Pontes 2016). According to Kuss and Pontes (2019), since the publication of the nine IGD criteria in May 2013, a total of seven clinical psychometric tools were developed to assess IGD. Although progress has been made toward unifying the assessment of IGD, clinical and cross-cultural evidence for these assessment tools remains limited as only the Internet Gaming Disorder Test (IGD-20 Test; Pontes et al. 2014), the Internet Gaming Disorder Scale–Short Form (IGDS9-SF; Pontes and Griffiths 2015b), and the Ten-Item Internet Gaming Disorder Test (IGDT-10; Király et al. 2017) have been extensively examined cross-culturally. More specifically, the IGD-20 Test has been psychometrically validated and culturally adapted for Spanish, Arabic, and Korean speakers, and the IGDS9-SF has been validated for Slovenian, Portuguese, Italian, and Persian speakers. Similarly, the IGDT-10 has been investigated in Hungarian, Iranian, Norwegian, Czech, Peruvian, French, and English samples, and only the Clinical Assessment Tool (C-VAT 2.0; Van Rooij et al. 2017) has preliminary clinical evidence to support its use in clinical samples.

Arguably, the potential unification in the field and the increased consistency in IGD research has led to IGD being officially recognized as a potential mental health disorder by the World Health Organization in September 2018 after an extensive and iterative review process (World Health Organization 2018a). Additionally, the WHO adopted the term “disorder” instead of “addiction.” The choice in the nomenclature accounts for ongoing discussion among scholars debating if excessive gaming is best characterized as an addictive disorder or something else (King and Gaming Industry Response Consortium 2018). Based on this, “Gaming Disorder” (GD) is now classed as a bona fide disorder due to addictive behaviors in relation to excessive gaming. The WHO’s decision to include GD in the 11th Revision of the International Classification of Diseases (ICD-11) was based on extensive reviews of existing peer-reviewed evidence and reflects a consensus among experts from different disciplines (e.g., psychology, psychiatry, neurosciences) and geographical regions that were involved in several consultation meetings organized by the WHO in the process of developing the ICD-11.

More specifically, GD is defined in the beta draft of the ICD-11, as a pattern of persistent or recurrent online and/or offline gaming behavior manifested by three core diagnostic

criteria: (1) *impaired control over gaming* (e.g., *onset, frequency, intensity, duration, termination, context*); (2) *increasing priority given to gaming to the extent that gaming takes precedence over other life interests and daily activities*; and (3) *continuation or escalation of gaming despite the occurrence of negative consequences* (World Health Organization 2018b). Moreover, GD should only be diagnosed when the behavior pattern is of sufficient severity to result in significant impairments in personal, family, social, educational, occupational, or other important areas of functioning (World Health Organization 2018b). In this context, for a diagnosis to be assigned, the pattern of gaming behavior should be continuous or episodic and recurrent, and the gaming behavior and its associated features should be normally evident over a period of at least 12 months. Nonetheless, the required diagnostic duration may be shortened if all diagnostic requirements are met and symptoms are severe (World Health Organization 2018b). In addition to this, exclusion criteria when diagnosing GD have been specified by the WHO in the ICD-11. These include hazardous gaming (QE22), bipolar type I disorder (6A60), and bipolar type II disorder (6A61). This nuanced diagnostic approach proposed by the WHO is in line with previous findings suggesting that gaming can be an intense activity whereby individuals may end up spending copious amounts of time gaming and yet no significant clinical impairment may be experienced as the gaming behavior falls within the notion of a high engagement behavioral pattern toward gaming due to the lack of endorsement of core diagnostic criteria (i.e., conflict, withdrawal symptoms, relapse, and behavioral salience) (Charlton and Danforth 2007).

Although GD has been recognized by the WHO as an official disorder, the field now faces similar research challenges to those that arose when IGD was included in the DSM-5. More specifically, in order to advance research and further the scientific understanding of GD under the new diagnostic framework established by the WHO, researchers should endeavor to establish standardized tools with adequate psychometric properties to assess the core criteria for GD according to this new framework, particularly in reference to the lack of direct diagnostic comparability between the IGD and the GD criteria. Additionally, cross-cultural clinical and psychometric research will be key to shed light on the relevance of each criterion. This is indeed a crucial step to be considered in GD research in order to provide empirical information related to the diagnostic properties and efficacy of the new criteria as it will impact on clinical practice and research on GD around the world.

In light of the aforementioned rationale, the main aim of the present study was to develop the first-ever standardized psychometric tool for assessing GD using the new diagnostic framework devised by the WHO. Additionally, this study will scrutinize key psychometric properties of these criteria adopting a cross-cultural perspective by investigating a sample of Chinese and British gamers. The main reason for focusing on these two specific population results from the fact that addictive gaming has become a major public health issue in Asian countries such as China (with prevalence rates ranging from 3.5 to 17%) (Long et al. 2018) and an emerging issue in developed Western countries such as the UK (with prevalence rates reported around 14.6%) (Lopez-Fernandez et al. 2014).

To the best of the authors' knowledge, this is the first study ever to be conducted on the effects of addictive gaming by adopting the WHO's diagnostic criteria. Therefore, this study contributes to facilitating international research in this field and improving diagnostic approaches in clinical milieus by developing a standardized psychometric tool to assess GD and ascertaining its suitability to measure the problem behavior in a valid and reliable way.

Method

Participants and Procedures

The present study recruited two samples of gamers in China and in the UK. For the Chinese sample, a call for participants was promoted at a large university in Beijing. Moreover, for the British sample, students from two major universities from the East Midlands and Greater London area were invited to participate in the research project. There was no distinction between undergraduate and postgraduate students across the countries.

Data collection for both samples was conducted using online surveys. More specifically, the British sample was recruited using an online survey hosted on Qualtrics (www.qualtrics.com) while the Chinese sample was recruited via an online survey programmed on the Survey Coder platform (www.ckannen.com). The two recruiting online platforms were hosted independently and despite the difference in language, they were equivalent in terms of the questions included on both surveys among the two samples. Participation was entirely voluntary and no financial compensation was offered to participants. Additionally, all participants were assured of anonymity and confidentiality, and the study was granted approval by the College Research Ethics Committee at Nottingham Trent University. All procedures of the study were executed in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 2005.

Participants were eligible to participate in the study upon responding to the following screening question: “have you played any video game in the past 12 months (yes/no)?” Respondents were not allowed to take the survey in case they have responded “no” to this question. A total of 25 participants (4%) were excluded based on this inclusion criterion, resulting in a total of 597 recruited participants.

Table 1 Chinese and British samples, sociodemographic characteristics and related gaming behaviors

Variables	Sample		
	Overall	Chinese	British
Sample size (<i>n</i>)	560	236	324
Gender (male, %)	271 (48.4)	111 (47)	160 (49.4)
Age (years) (mean, SD)	23.6 (7.09)	19.22 (1.57)	26.74 (7.88)
Relationship status (in a relationship, %)	237 (42.3)	66 (28)	171 (52.8)
Weekly time spent gaming (mean, SD)	12.69 (12.05)	11.56 (13.22)	13.51 (11.03)
Percentage of time spent on weekend (%)	46.15	49.76	43.52
Experienced significant problems due to gaming (yes, %)	36 (6.4)	21 (8.9)	15 (4.6)
Depression severity (mean, SD)	17.05 (5.9)	14.75 (4.82)	18.72 (6.21)
Loneliness severity (mean, SD)	6.51 (2.6)	5.91 (2.36)	6.95 (2.68)
Gaming disorder severity (mean, SD)	6.89 (3.17)	6.17 (2.91)	7.42 (3.25)
Gaming disorder prevalence (<i>n</i> , %)	10 (1.8)	4 (1.7)	6 (1.9)
Item-related descriptive statistics (mean, SD)			
GDT item 1	1.85 (0.94)	1.87 (0.96)	1.84 (0.93)
GDT item 2	1.98 (1.06)	1.56 (0.86)	2.29 (1.08)
GDT item 3	1.69 (1.04)	1.43 (0.85)	1.89 (1.13)
GDT item 4	1.36 (0.76)	1.31 (0.76)	1.40 (0.76)

SD, standard deviation; *GDT*, Gaming Disorder Test

The mean age of the sample was 20 years in China ($SD = 1.57$, range 18–24 years) and 26 years in the UK ($SD = 7.81$, range 18–49 years). Overall, the gender split was relatively even with males representing 52% ($n = 310$) of the total sample (see Table 1 for further details).

Measures

Sociodemographic and Gaming-Related Behaviors Sociodemographic and gaming-related behaviors questions were aligned with previous similar psychometric studies on GD (e.g., Pontes and Griffiths 2015b; Pontes et al. 2014; Schivinski et al. 2018). More specifically, sociodemographic data included participants' gender, age, and relationship status. Gaming-related behaviors were controlled by self-reported average time spent playing videogames in a week and weekends. An additional self-report question was included asking if participants had experienced significant problems in their lives due to gaming as their activity in the past 12 months.

The Gaming Disorder Test

The Gaming Disorder Test (GDT) is a brief standardized assessment tool including four items reflecting the key defining diagnostic features of GD in the ICD-11 (World Health Organization 2018b). The GDT examines gaming activities occurring over a 12-month period since the WHO criteria for GD are based on persistent and recurrent gaming. This most often involves specific online and/or offline games, regardless of the device used to play (e.g., consoles, computers, smartphones). All four items are rated on a 5-point Likert scale: 1 (“never”), 2 (“rarely”), 3 (“sometimes”), 4 (“often”), and 5 (“very often”). Total scores are obtained by summing the gamer's answers and they can range from 4 to 20, with higher scores being indicative of higher degrees of disordered gaming. It is also worth noting that the main purpose of this instrument is not to diagnose GD but to assess its severity and accompanying detrimental effects to the gamer's life. However, for research purposes, it is recommended that answers are given as 4 (“often”) or 5 (“very often”) to any of the four items should be coded as endorsement of a specific GD criterion. By adopting this (≥ 4 or 5) diagnostic approach, researchers will be able to distinguish between potentially disordered and non-disordered gamers.

Prior to developing the GDT items, the research team consulted with a group of experts (i.e., *blinded for review purposes*) working in the field in order to ensure that the scale had adequate content validity. Furthermore, the first three items were devised to map on (i) *impaired control over gaming*, (ii) *increased priority given to gaming*, and (iii) *continuation despite negative consequences*. An additional item reflecting the (iv) *experience of significant problems in life* was added given that at higher levels; GD is of sufficient severity to result in significant impairment in personal, family, social, educational, occupational, or other important areas of functioning (World Health Organization 2018b). The inclusion of the fourth item helps to ensure that the scale is able to effectively capture GD at different levels of severity rather than excessive or hazardous gaming according to the definition proposed by WHO.

For the process of the cross-cultural adaptation of the GDT in China, we adopted a standard procedural method (Beaton et al. 2000). Therefore, two bilingual Chinese translators translated the GDT from its original English version to Mandarin. The Chinese GDT was then back-

translated to English. The translators (see acknowledgement) discussed together with the authors the back and forth translations of the GDT in Mandarin to ensure that the meaning of the Chinese version of the GDT fits the original English items. Finally, we assessed face validity of the GDT by running a pilot testing among 47 videogame players with the English version of the scale (51% male, $\text{mean}_{\text{age}} = 21.4$, $\text{SD}_{\text{age}} = 3.5$). The respondents did not report any significant problems when completing the questionnaire.

Internet Gaming Disorder Scale—Short-Form

The Internet Gaming Disorder Scale—Short-Form (IGDS9-SF) (Pontes and Griffiths 2015b) is a psychometric tool adapted from the nine IGD criteria according to the DSM-5 (American Psychiatric Association 2013). This instrument is used to assess the severity of IGD symptoms and its detrimental effects by examining both online and/or offline gaming activities occurring over a 12-month period. The nine items of the IGDS9-SF are answered using a 5-point Likert scale: 1 (“never”), 2 (“rarely”), 3 (“sometimes”), 4 (“often”), and 5 (“very often”). Total scores can be obtained by summing the gamer’s answers to each item and total scores can range from 9 to 45, with higher scores being indicative of higher degrees of disordered gaming. The IGDS9-SF has been extensively investigated in cross-cultural research and has been found to be consistently valid and reliable to assess IGD symptoms across different countries and samples, including in Chinese- and English-speaking samples (Stavropoulos et al. 2018; Yam et al. 2018).

Three-Item Loneliness Scale

Loneliness was assessed using three items from the UCLA Loneliness Scale, version 3 (Russell 1996). These items were extensively tested in large sample surveys using population-based samples, and they displayed excellent psychometric properties for assessing loneliness concisely across both Chinese- and English-speaking samples (e.g., Hughes et al. 2004; Xu et al. 2018). All three items are rated on a 4-point Likert scale: 1 (“never”), 2 (“rarely”), 3 (“sometimes”), and 4 (“often”). The total of the responses to the three questions was used to determine participants’ levels of loneliness.

The Patient Health Questionnaire

The Patient Health Questionnaire (PHQ-9) (Kroenke et al. 2001) is a nine-item instrument designed to assess symptoms of depression and was developed based on the criteria for major depressive disorder in the DSM-IV (American Psychiatric Association 1994). Participants evaluate their symptoms referencing the last 2 weeks, using a 4-point Likert scale for a duration ranging from 0 (“not at all”), 1 (“several days”), 2 (“more than half the days”), and 3 (“nearly every day”). The PHQ-9 has been used as a screening and diagnostic tool and as an outcome measure in both Chinese- and English-speaking samples and displayed excellent psychometric properties (e.g., Merz et al. 2011; Tsai et al. 2014). Total scores of 5, 10, 15, and 20 can be used to determine “mild,” “moderate,” “moderately severe,” and “severe” depression, respectively (Kroenke et al. 2001).

Data Management and Analytic Strategy

Data management involved undertaking several steps to ensure the applied parametric testing would be conducted appropriately. First, the structure of the missing data was tested with Little's Missing Completely at Random (MCAR) test using the package BaylorEdPsych (R Package for Baylor University Education Psychology Quantitative Course Version 0.5) in R system for statistical computing Version 3.4.2 "Short Summer" (<https://www.r-project.org>). The results of this test yielded a chi-square of 8.36, $df = 13$, $p = 0.81$. The hypothesis of the MCAR is rejected at 0.05 significance level, therefore it can be assumed that the data is missing at random. Furthermore, a total of 32 (5.3%) cases were excluded from the analyses at this stage due to severe missing values (i.e., ≥ 5) on the psychometric tests used in the study.

The distribution of all items across all psychometric tests utilized in the present study was examined to assess univariate normality. As a result, no item of the GDT and the other psychometric tests had absolute values of skewness > 3.0 and kurtosis > 8.0 (Kline 2011). Additionally, the data was screened for univariate outliers by computing a standardized composite sum score using the four GDT items. Participants were considered univariate outliers if their composite score obtained was ± 3.29 standard deviations from the GDT z -scores as this threshold includes 99.9% of the normally distributed z -scores (Field 2013). The same procedure was repeated for all the remainder of the psychometric tests utilized in the study (see Table 2). Moreover, multivariate outliers were assessed by examination of the Mahalanobis distances and the critical value for each case (based on the chi-square distribution values). As a result of this procedure, an additional three participants were excluded from the subsequent analyses. Overall, after applying the data cleaning strategy outlined, a final sample of 560 (93.8%) respondents fully eligible for the final statistical analyses was reached.

Statistical Analyses

Statistical analysis of the data collected included (i) descriptive statistics of the main sample's characteristics, (ii) an in-depth psychometric evaluation of the GDT that included analysis of the construct validity (i.e., *factorial validity* via Confirmatory Factor Analysis [CFA], *nomological validity* by estimating a full structural equation model to test a set of a priori theoretical assumptions about the interplay between these constructs; *convergent* and *discriminant*

Table 2 Reliability, convergent, and discriminant validity of the Gaming Disorder Test (GDT)

Overall	α	CR	AVE	(1)	(2)	(3)
(1) Loneliness	0.84	0.85	0.66	<i>0.81</i>		
(2) Depression	0.90	0.90	0.52	0.75	<i>0.72</i>	
(3) GDT	0.84	0.85	0.59	0.45	0.47	<i>0.77</i>
Chinese						
(1) Loneliness	0.82	0.84	0.63	<i>0.79</i>		
(2) Depression	0.87	0.87	0.46	0.74	<i>0.68</i>	
(3) GDT	0.87	0.87	0.63	0.38	0.47	<i>0.79</i>
British						
(1) Loneliness	0.85	0.86	0.67	<i>0.82</i>		
(2) Depression	0.90	0.90	0.53	0.72	<i>0.73</i>	
(3) GDT	0.84	0.85	0.59	0.45	0.41	<i>0.77</i>

The square root of the Average Variance Extracted for each construct is in italics

validity analysis based on the Average Variance Extracted [AVE] coefficients of each latent variable), and a (iii) reliability analysis using different coefficients and indicators of internal consistency (i.e., Cronbach's alpha and Composite Reliability [CR]) was performed. Finally, (iv) concurrent validity of the GDT was investigated using the IGDS9-SF as both tools assess the same construct.

To assess the quality of the structural equation models tested, several fit indices indicators were used to judge the quality of the goodness of fit (GOF). However, because there is no consensus on the fit indices for evaluating structural equation models (see Bollen and Long 1993; Boomsma 2000; Hoyle and Panter 1995), the following fit indices and thresholds were adopted in the present study: χ^2/df [1;4]; Comparative Fit Index (CFI); and Tucker-Lewis Fit Index (TLI) [0.90;0.95]; Root Mean Square Error of Approximation (RMSEA) [0.05;0.08]; and the Standardized Root Mean Square Residual (SRMR) [0.05;0.08] (Bentler 1990; Bentler and Bonnet 1980; Hooper et al. 2008; Hu and Bentler 1999). These fit indices are often used in similar psychological research and provide the best coefficients to assess the quality of the results obtained in latent variable modeling. Furthermore, Full Information Maximum Likelihood estimation method (FIML) was used in the estimation of the structural equation models with 5000 bootstrap samples to yield robust standardized errors (Berkovits et al. 2000).

All the aforementioned statistical analyses were performed using R system for statistical computing. More specifically, the following packages were utilized: Psych (Procedures for Psychological, Psychometric, and Personality Research Version 1.8.10), and Lavaan (Latent Variable Analysis Version 0.6–3).

Results

Descriptive Statistics

With regard to gender distribution in the overall sample, females represented 51.6% ($n = 289$) of all participants, reflecting a balanced gender distribution. Participants' age ranged between 18 and 49 years, and the mean age observed in the sample was 23 years ($SD = 7.09$ years). Finally, about (42.3%, $n = 327$) of participants reported not being involved in a romantic relationship (42.3%, $n = 327$).

As per gaming-related behaviors, the average time spent playing videogames during the week was 12 h ($SD = 12.05$ h), with about 46% of this time being spent over the weekend alone. Finally, a total of 36 participants (6.4%) reported having experienced significant problems in their lives due to their gaming behavior according to the self-report question developed within the sociodemographic section of the survey. In terms of symptomatology, depression ($mean_{overall} = 17.05$; $SD_{overall} = 5.9$; $min = 8$, $max = 32$), and GD severity ($mean_{overall} = 6.89$; $SD_{overall} = 3.17$; $min = 4$, $max = 20$) were higher in the overall sample in comparison to loneliness ($mean_{overall} = 6.51$; $SD_{overall} = 2.6$; $min = 3$, $max = 12$). Gamers from the British sample exhibited significantly higher levels of depression ($t[558] = 8.16$, $p < 0.001$; $d = 0.69$), loneliness ($t[558] = 4.77$, $p < 0.001$; $d = 0.40$), and severity of GD ($t[560] = 4.67$, $p < 0.001$; $d = 0.40$). Finally, the prevalence rates of GD found among the Chinese and British gamers did not differ significantly ($\chi^2 [1, N = 560] = 0.02$, $p = 0.58$). A complete summary of all participants' main sociodemographic characteristics and severity of symptoms of the constructs assessed is presented in Table 1, including all variables investigated in the present study among both the Chinese and British samples.

Factorial Validity: CFA and Unidimensionality Testing

The factor structure of the GDT was investigated and operationally defined under a single-factor solution due to statistical and theoretical reasons as an optimal balance between scale length with parsimony in measurement should be warranted. At the statistical level, given the brevity of the GDT, having more than one dimension would lead to latent factors containing less than three indicators, which is not aligned with the standard recommendations made against retaining factors with fewer than three indicators (Tabachnick and Fidell 2013; Worthington and Whittaker 2006). At the theoretical level, brief IGD standardized tools have been recently operationalized at the latent level as a unidimensional factor containing the key core indicators and clinical symptoms of the disorder (e.g., Király et al. 2017; Lemmens et al. 2015; Pontes and Griffiths 2015b).

Based on this rationale, the unidimensional factor structure of the GD construct as measured with the four GDT items was examined by performing a CFA with relevant covariates in a Multiple-Indicator Multiple-Causes model (MIMIC model) (see Fig. 1). For the MIMIC model specification, GD was predicted by the following variables: age, gender, and weekly time spent playing videogames. These variables were selected due to their role in predicting GD as several empirical studies reported GD to be associated with younger and male gamers, and greater time spent playing videogames (e.g., Dong et al. 2018; Fuster et al. 2016; Lemmens et al. 2009; Pontes and Griffiths 2015b, 2016; Schivinski et al. 2018).

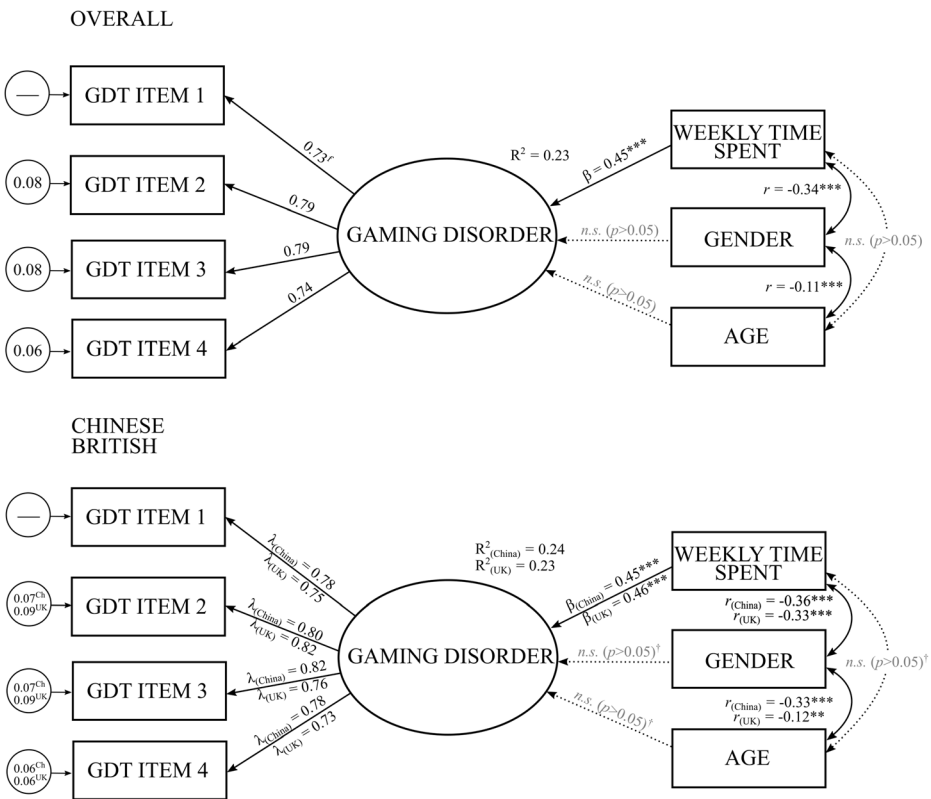
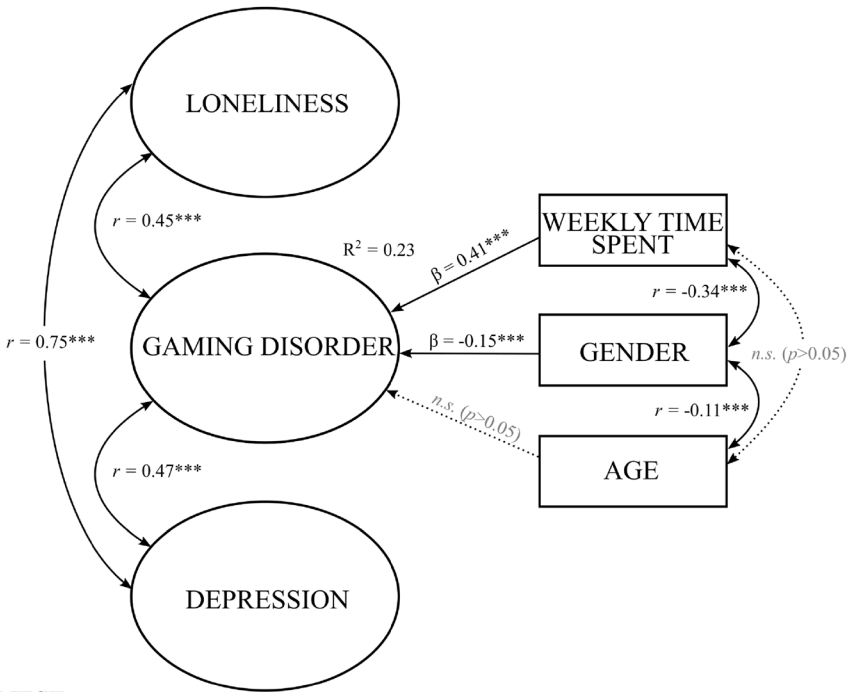


Fig. 1 Unidimensional factor structure of GD construct as measured with four GDT items was examined by performing a CFA with relevant covariates in a Multiple-Indicator Multiple-Causes model

OVERALL



CHINESE
BRITISH

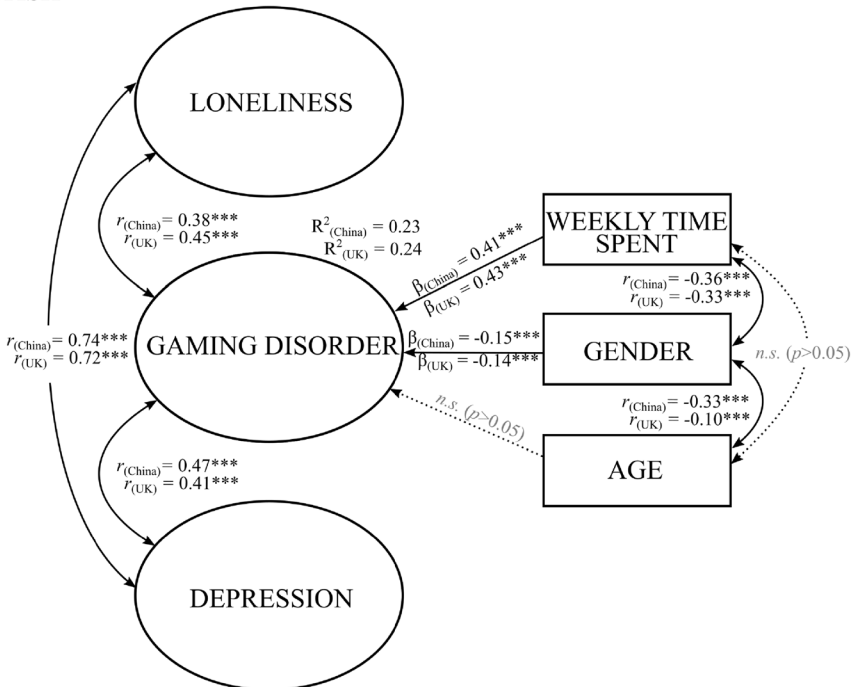


Fig. 2 Total variance explained in GD by weekly time spent, age, and gender

The results of this analysis yielded the following GOF: $FIML_{\text{overall}}\chi^2_{(11)} = 46.22$, $\chi^2/df = 4.20$, $CFI = 0.97$, $TLI = 0.94$, $RMSEA = 0.07$, and $SRMR = 0.03$, further suggesting that the unidimensional factor structure for the GDT construct fits the data well. Overall, similar findings were obtained in the Chinese ($FIML_{\text{Chinese}}\chi^2_{[11]} = 12.31$, $\chi^2/df = 1.12$, $CFI = 0.99$, $TLI = 0.99$, $RMSEA = 0.02$, and $SRMR = 0.02$) and British ($FIML_{\text{uk}}\chi^2_{[11]} = 34.68$; $\chi^2/df = 3.15$, $CFI = 0.93$, $TLI = 0.92$, $RMSEA = 0.08$, and $SRMR = 0.03$) samples. The standardized item loads ranged from $\lambda_{\text{GDT1overall}} = 0.73$ to $\lambda_{\text{GDT3overall}} = 0.79$ and were all above the desired threshold of $\lambda_{ij} \geq 0.50$, $p < 0.001$ (Hair et al. 2010).

Furthermore, the results of this analyses demonstrated that GD was positively influenced by time spent playing videogames during the week ($\beta = 0.45$, $p < 0.001$), while age and gender were not statistically significant ($p = 0.29$ and $p = 0.27$, respectively). A detailed summary of these findings is presented in Fig. 1, including the breakdown of the results according to the overall sample and the two countries.

Construct Validity: Nomological Validation

The assessment of the construct validity of GDT involved identifying a relevant network of key constructs associated with GD to explain the pattern of interrelationships that exists among them (Bryant et al. 2007). This procedure has been extensively discussed by Cronbach and Meehl (1955) who suggested that it is necessary to understand the nature of a construct through the statistical or deterministic laws underlying the network of key constructs, often referred to as the nomological network. The nomological network is considered an essential aspect of construct validity of a given phenomenon and was investigated in the present study by replicating the structural and causal relationships between GD, loneliness, and depression accounting for potential confounding effects. The decision to investigate the interplay between these three latent constructs was informed by empirical findings.

Based on the aforementioned rationale, the structural model testing the nomological network proposed included GD, loneliness, and depression in a mutually reinforcing (bi-directional) association pattern. Additionally, GD has been specified as the outcome of weekly time spent gaming, age, and gender to control for potential confounding effects stemming from key individual differences factors. The results obtained for the overall sample produced an adequate fit to the data ($FIML_{\text{overall}}\chi^2_{[130]} = 480.66$, $\chi^2/df = 3.70$, $CFI = 0.92$, $TLI = 0.91$, $RMSEA = 0.06$, and $SRMR = 0.05$). As expected, GD was positively associated with loneliness ($r = 0.45$, $p < 0.001$) and depression ($r = 0.45$, $p < 0.001$), further corroborating previous findings (see Fig. 2). Finally, the total variance explained in GD by weekly time spent, age, and gender was 23% ($R^2 = 0.23$, $p < 0.05$) (see Fig. 2). Overall, similar findings were obtained in the Chinese ($FIML_{\text{Chinese}}\chi^2_{[129]} = 224.13$, $\chi^2/df = 1.74$, $CFI = 0.94$, $TLI = 0.93$, $RMSEA = 0.05$, and $SRMR = 0.00$) and British ($FIML_{\text{UK}}\chi^2_{[129]} = 313.41$, $\chi^2/df = 2.43$, $CFI = 0.93$, $TLI = 0.92$, $RMSEA = 0.06$, and $SRMR = 0.06$) samples.

Taken together, these results highlight the suitability of the GDT to measure GD symptoms and capture previously reported association patterns between other closely related and implicated mental health constructs.

Convergent Validity, Discriminant Validity, and Reliability Analysis

The literature defines convergent validity as the extent to which items of a psychometric test appear to be indicators of a single underlying construct (Lee et al. 2015). Convergent validity

Table 3 Bootstrapped correlation matrix with 95% bias-corrected and accelerated confidence interval between the GDT and IGDS9-SF scores across all samples

Measure (sample)	GDT	BCa 95% CI	R ²
IGDS9-SF (overall)	0.83**	0.78; 0.86	0.68
IGDS9-SF (China)	0.82**	0.75; 0.88	0.67
IGDS9-SF (UK)	0.82**	0.77; 0.86	0.67

IGDS9-SF, Internet Gaming Disorder Scale—Short-Form; *GDT*, Gaming Disorder Test; *R*², R-squared

Bootstrap results are based on 5,000 bootstrap samples

**Correlation is significant at the 0.01 level

is deemed adequate when the AVE of the latent variable is ≥ 0.50 and composite reliability is ≥ 0.70 and there is no evidence of cross-loadings across the constructs (Fornell and Larcker 1981; Hair et al. 2010). As shown in Table 2, the AVE values for the GDT in the overall, Chinese and British samples were all adequate (0.59, 0.63, and 0.59 respectively), and the composite reliability coefficients were well beyond the desired threshold across the overall, Chinese, and British samples (0.85, 0.87, and 0.85 respectively). Furthermore, no evidence of cross-loadings across the constructs was found.

The notion of discriminant validity refers to the degree to which the measures of distinct constructs differ (Lee et al. 2015) and is demonstrated when the square root of the AVE of each latent variable is higher than the correlations between it and the rest of the constructs (Fornell and Larcker 1981; Hair et al. 2010). The square root of the AVE of each latent variable is in italics in Table 2. The results of this analysis suggest that the value for each latent variable was higher than the correlations between it and the other constructs of the study.

Finally, the internal consistency of the GDT was also assessed using the Cronbach's alpha as another indicator of internal consistency in addition to the CR (see Table 2). More specifically, the Cronbach's alphas of the GDT were all excellent across the overall ($\alpha = 0.84$), Chinese ($\alpha = 0.87$), and the British ($\alpha = 0.84$) samples. Overall, the results of these analyses illustrate that the GDT demonstrated convergent validity, discriminant validity, and reliability across different levels, further supporting its psychometric robustness to measure GD.

Concurrent Validity: GDT and IGDS9-SF

The concurrent validity of the GDT was assessed to examine the degree to which this measure relates to the IGDS9-SF (Pontes and Griffiths 2015b). The main assumption of the concurrent validity analysis is that a psychometric test should show substantial correlation with other measures to which it is theoretically related (Frick et al. 2010). This procedure involves administering a new psychometric tool and an existing well-validated measure of psychopathology to a group of individuals. If a correlation of 0.20 is obtained, then the concurrent validity of the new measure would be questionable. However, a correlation ≥ 0.75 indicates a sufficient degree of concurrent validity of the new measure (Frick et al. 2010).

In the present study, concurrent validity was assessed by examining the degree of association (i.e., bootstrapped correlation with bias-corrected accelerated 95% confidence intervals) between the GDT and IGDS9-SF overall scores across all three subsamples (i.e., overall, Chinese, and British). As expected, participants' level of GD (as assessed by the GDT) was highly associated with the IGDS9-SF scores. More specifically, the obtained correlation coefficients ranging from $r = 0.82$ ($p < 0.001$) to $r = 0.83$ ($p < 0.001$) (see Table 3). Therefore,

it can be concluded that the GDT exhibited satisfactory concurrent validity when considering other closely related measures.

Discussion

The present study sought to develop the first psychometric tool to assess GD based on the newest diagnostic framework developed by the WHO in the ICD-11 (World Health Organization 2018b). In order to achieve this goal, experts in the field developed the GDT, a four-item standardized psychometric tool covering all key GD diagnostic criteria and clinical features. Furthermore, the GDT was investigated in a cross-cultural setting involving gamers from China and the UK. The GDT was scrutinized at the psychometric level using different parameters to support its validity and reliability and whether it reflects the concept of GD. All the psychometric analyses were performed on three subsamples including the overall sample, Chinese sample, and the British sample to ensure consistency of the findings.

Overall, the results obtained supported the new scale's factorial validity and demonstrated that the single-factor solution for GD as measured by the GDT reflects an optimal factor structure for the construct. This finding mirrors similar research in the field that aimed to develop psychometric tools based on the nine IGD criteria as defined in the DSM-5 (American Psychiatric Association 2013). More specifically, researchers have found empirical support to a unidimensional factor structure for IGD across different studies (e.g., Chiu et al. 2018; Király et al. 2017; Pontes and Griffiths 2015b). Although GD was psychometrically operationalized within a unidimensional factor structure, all indicators (i.e., items) of the latent construct reflected the different clinical manifestations of the disorder in order to cover its main clinical features.

In addition to factorial validity, other sources of validity of the GDT were investigated in the present study, such as construct validity (i.e., nomological validity, convergent validity, and discriminant validity) and criterion-related validity (i.e., concurrent validity). Finally, the internal consistency of the GDT was also investigated using different indicators of reliability (i.e., Cronbach's alpha and CR).

With regard to the nomological validity analysis of the GDT, a nomological network of relevant constructs was examined. More specifically, GD, loneliness, and depression were examined within a bi-directional association pattern when controlling for weekly time spent gaming, age, and gender effects. This rationale was informed by previous empirical research suggesting that GD, loneliness, and depression are consistently implicated at the cross-sectional level (e.g., Burleigh et al. 2018; Lee et al. 2018; Lemmens et al. 2015; Myrseth et al. 2017; Pontes 2017). More specifically, these studies reported a positive association between these constructs. Furthermore, a recent large-scale longitudinal study using a representative sample of adolescents in Norway investigating the developmental trajectories of GD and mental health factors such as depression and loneliness found that all three variables were reciprocally associated (Krossbakken et al. 2018). Based on this, the results of this analysis indicated that the nomological network tested was supported and consistently replicated across the three (sub)samples (i.e., overall, Chinese, and British samples), lending further support to the construct validity of the GDT. Additionally, similar findings were obtained regarding the convergent validity and discriminant validity of the GDT as the results suggested that the GDT items reflect a single underlying construct (i.e., convergent validity) and that they also uniquely measure GD when compared against distinct constructs (i.e., discriminant validity). In relation to the examination of the concurrent validity, the GDT scores were consistently and highly associated with the IGDS9-SF

scores obtained by participants across all three (sub)samples. Finally, the internal consistency of the GDT was excellent across all the three (sub)samples. Taken together, the findings concerning the different types of validity and reliability of the GDT were consistent across the three samples, further supporting the population cross-validity of the GDT.

Previous research on IGD assessment reported several inconsistencies undermining the conceptualization and measurement of the construct (Pontes and Griffiths 2014). For instance, most of the existing instruments were found to be inconsistent since no two psychometric tests were alike in their conceptualization and ability to measure key diagnostic features (King et al. 2013). Moreover, the main limitations in existing psychometric tools included (i) inconsistent coverage of core addiction indicators, (ii) varying cut-off scores to indicate clinical status, (iii) lack of a temporal dimension, (iv) untested or inconsistent dimensionality, and (v) inadequate data on predictive validity and inter-rater reliability. It is envisaged that the GDT will uniquely contribute to overcoming previous inconsistencies in the assessment of GD by adopting the latest official theoretical framework for GD as developed by the WHO in the ICD-11. Additionally, this study represents an initial effort to standardize the clinical criteria for GD in a way that it will benefit time-limited research as the GDT is a brief tool that can be easily administered to a large number of gamers in a short period of time.

Previous authors (Koronczai et al. 2011) have suggested that a suitable psychometric tool should meet six key criteria: (i) comprehensiveness (i.e., examining many and possibly all aspects of GD); (ii) brevity, so that it can be used for impulsive individuals and fit time-limited surveys; (iii) reliability and validity for different data collection methods; (iv) reliability and validity across different age groups; (v) cross-cultural reliability and validity; and (vi) validation on clinical samples for determining more precise cut-off points based not only on empirical data. Although the present study was not able to fully cover all the six steps outlined above, it was still able to develop the GDT, which is a (i) comprehensive (i.e., measures all aspects of GD as defined by the latest clinical criteria developed by the WHO); (ii) brief instrument with a minimal amount of indicators able to fully measure the phenomenon while still capable of maintaining a high level of construct and content validity; and presents robust indicators of cross-cultural reliability and validity among two distinct samples. Nevertheless, the present study was not able to entirely fulfill the third, fourth, and sixth criteria outlined above as (i) the findings reported only refer to one type of data collection method (i.e., online survey using convenience sampling); (ii) the age of participants did not include all relevant age-related demographics of all gamers; and the study (iii) lacked a clinical sample to aid the development of robust cut-off points. As such, further research should focus on investigating the psychometric properties of the GDT in more culturally diverse samples, including clinically diagnosed individuals as this will enable further developments of the GDT and refinement to the GD criteria by providing further information about the diagnostic properties and efficiency of the GDT using the GD framework.

Notwithstanding this, the present study provides a useful and much-needed resource for assessing GD in two major nations (i.e., China and the UK). Both the Chinese and English versions of the GDT (see Appendix 1) will help advance empirical and clinical research on GD by facilitating assessment of GD in different countries. In addition to this, the present findings were estimated accounting for the potential effects stemming from excessive gaming and high engagement via increased time spent gaming, therefore a greater degree of certainty and robustness exists regarding how the findings were obtained in relation to the measurement of core GD symptoms as opposed to measurement of peripheral GD symptoms. The GDT is a suitable measure for large-scale and nationwide research projects aiming to establish the

prevalence of GD in a large number of individuals. Thus, future research should aim to investigate additional forms of validity of the GDT using large samples and more sophisticated statistical analyses such as latent profile analysis to ascertain the profile of gamers and further determine the diagnostic properties of the GDT.

In light of previous developments in research regarding the assessment of IGD and the psychometric tools developed to assess the phenomenon, the present work contributes uniquely to the assessment of GD by providing an initial attempt to measure GD strictly using the main criteria outlined by the WHO. Although existing IGD tools are still relevant for measuring disordered gaming based on the IGD framework, fundamental clinical and diagnostic differences will emerge from the IGD-based psychometric assessment tools and the psychometric tools developed under the new WHO framework. For example, the IGD diagnostic framework includes the experience of withdrawal symptoms when gaming is not possible as a core diagnostic criterion. By contrast, the new framework for GD based on the WHO's proposal does not include withdrawal symptoms within its core conceptualization of disordered gaming. Although the present study cannot offer empirical evidence supporting the potential hindrances to diagnostic practices using the two approaches, it does offer an updated, concise, and clinically useful psychometric tool that warrants further testing, which will hopefully shed light on how these differences shape epidemiological rates of GD across countries and how diagnostic approaches may be conducted in clinical settings.

Although this study is likely to contribute to future unification in the assessment of GD (Griffiths et al. 2014), it is indeed not without potential limitations. Firstly, the findings reported were based on convenience samples. Even though the use of convenience sample is common in this field of research, such sampling strategy may be problematic in understanding the full extent and degree of the generalizability of these findings as the present study relied on convenience samples comprising Chinese and British gamers. Consequently, the findings reported may not be necessarily representative of all Chinese and British gamer populations. Secondly, the use of self-report psychometric tests is often criticized as it may be accompanied by possible biases such as social desirability bias and short-term recall bias. Thirdly, the present study was not able to estimate robust cut-off points to aid future diagnostic practices using the newly developed tool. However, this potential limitation arose from the fact that a clinical sample was not recruited, inevitably rendering it impossible to obtain a clinical gold standard for GD.

In conclusion, the present study lends empirical support for the concept of GD as currently defined by the WHO in the ICD-11 (World Health Organization 2018b). The findings obtained also support the viability of further studies investigating GD in greater depth. Moreover, the present findings demonstrated that the GDT can cater to the need for a brief standardized and psychometrically sound tool for assessing GD among Chinese- and English-speaking individuals under the latest diagnostic framework. Therefore, the authors of the present study envisage that the GDT will contribute to facilitating additional research in the field by providing a brief, valid, and reliable psychometric assessment tool to measure the core symptoms and severity of GD.

Funding The position of CM is funded by a Heisenberg grant awarded to him by the German Research Foundation (DFG, MO 2363/3-2).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Appendix

The Gaming Disorder Test (GDT) (Pontes et al. 2019)

Instructions: The questions below are about your gaming activity during the past year (i.e., last 12 months). Here, gaming activity means any gaming-related activity that has been played either from a computer/laptop or from a gaming console or any other kind of device (e.g., mobile phone, tablet) both online and/or offline.

Please indicate how often the following issues occurred on average over the past twelve months until today.

	Never	Rarely	Sometimes	Often	Very Often
1. I have had difficulties controlling my gaming activity.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. I have given increasing priority to gaming over other life interests and daily activities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. I have continued gaming despite the occurrence of negative consequences.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. I have experienced significant problems in life (e.g., personal, family, social, education, occupational) due to the severity of my gaming behavior.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	从不	很少	有时	经常	非常频繁
1. 我在控制自己玩游戏方面有困难。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. 我越来越重视游戏而不是其他生活乐趣和日常活动。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. 尽管出现了不好的后果（比如对人际关系、学习或工作产生了负面影响），我仍然继续玩游戏。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. 游戏的严重度已经导致我在生活中遇到了重大问题（例如对个人、家庭、社交、教育、职业产生了严重影响）。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Scoring information:

Total scores can be obtained by summing up all responses given to all four items of the GDT and can range from a minimum of 4 to a maximum of 20 points, with higher scores being indicative of a higher degree of gaming disorder. In order to differentiate disordered gamers from non-disordered gamers, researchers should check if participants have endorsed all four diagnostic criteria as assessed by each GDT items by taking into account answers as '4: Often' or '5: Very often', which translates as endorsement of the criterion.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Aarseth, E., Bean, A. M., Boonen, H., Colder, C. M., Coulson, M., Das, D., et al. (2017). Scholars' open debate paper on the World Health Organization ICD-11 gaming disorder proposal. *Journal of Behavioral Addictions*, 6(3), 267–270. <https://doi.org/10.1556/2006.5.2016.088>.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>.
- Bentler, P. M., & Bonnet, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606. <https://doi.org/10.1037/0033-2909.88.3.588>.
- Berkovits, I., Hancock, G. R., & Nevitt, J. (2000). Bootstrap resampling approaches for repeated measure designs: Relative robustness to sphericity and normality violations. *Educational and Psychological Measurement*, 60(6), 877–892. <https://doi.org/10.1177/00131640021970961>.
- Bollen, K. A., & Long, J. S. (1993). Introduction. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 1–9). Newbury Park: Sage.
- Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling: a Multidisciplinary Journal*, 7(3), 461–483. https://doi.org/10.1207/S15328007SEM0703_6.
- Bryant, F. B., King, S. P., & Smart, C. M. (2007). Multivariate statistical strategies for construct validation in positive psychology. In A. D. Ong & M. H. M. v. Dulmen (Eds.), *Oxford handbook of methods in positive psychology* (pp. 61–82). Oxford: Oxford University Press.
- Burleigh, T. L., Stavropoulos, V., Liew, L. W. L., Adams, B. L. M., & Griffiths, M. D. (2018). Depression, internet gaming disorder, and the moderating effect of the gamer-avatar relationship: an exploratory longitudinal study. *International Journal of Mental Health and Addiction*, 16, 102–124. <https://doi.org/10.1007/s11469-017-9806-3>.
- Charlton, J. P., & Danforth, I. D. W. (2007). Distinguishing addiction and high engagement in the context of online game playing. *Computers in Human Behavior*, 23(3), 1531–1548. <https://doi.org/10.1016/j.chb.2005.07.002>.
- Chiu, Y. C., Pan, Y. C., & Lin, Y. H. (2018). Chinese adaptation of the ten-item internet gaming disorder test and prevalence estimate of internet gaming disorder among adolescents in Taiwan. *Journal of Behavioral Addictions*, 7(3), 719–726. <https://doi.org/10.1556/2006.7.2018.92>.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. <https://doi.org/10.1037/h0040957>.
- Dong, G., Wang, L., Du, X., & Potenza, M. N. (2018). Gender-related differences in neural responses to gaming cues before and after gaming: implications for gender-specific vulnerabilities to Internet gaming disorder. *Social Cognitive and Affective Neuroscience*, 13(11), 1203–1214. <https://doi.org/10.1093/scan/nsy084>.
- Field, A. (2013). *Discovering statistics using IBM SPSS Statistics (fourth edition)*. London: SAGE Publications Ltd.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.2307/3151312>.
- Frick, P. J., Barry, C. T., & Kamphaus, R. W. (2010). *Clinical assessment of child and adolescent personality and behavior* (3rd ed.). New York: Springer.
- Fuster, H., Carbonell, X., Pontes, H. M., & Griffiths, M. D. (2016). Spanish validation of the Internet Gaming Disorder-20 (IGD-20) Test. *Computers in Human Behavior*, 56, 215–224. <https://doi.org/10.1016/j.chb.2015.11.050>.
- Griffiths, M. D., & Pontes, H. M. (2014). Internet addiction disorder and internet gaming disorder are not the same. *Journal of Addiction Research & Therapy*, 5(4), e124. <https://doi.org/10.4172/2155-6105.1000e124>.
- Griffiths, M. D., King, D. L., & Demetrovics, Z. (2014). DSM-5 internet gaming disorder needs a unified approach to assessment. *Neuropsychiatry*, 4(1), 1–4. <https://doi.org/10.2217/npj.13.82>.
- Griffiths, M. D., Király, O., Pontes, H. M., & Demetrovics, Z. (2015). An overview of problematic gaming. In E. Aboujaoude & V. Starcevic (Eds.), *Mental health in the digital age: grave dangers, great promise* (pp. 27–45). Oxford: Oxford University Press.
- Griffiths, M. D., Van Rooij, A. J., Kardefelt-Winther, D., Starcevic, V., Király, O., Pallesen, S., et al. (2016). Working towards an international consensus on criteria for assessing internet gaming disorder: a critical commentary on Petry et al. (2014). *Addiction*, 111(1), 167–175. <https://doi.org/10.1111/add.13057>.

- Griffiths, M. D., Kuss, D. J., Lopez-Fernandez, O., & Pontes, H. M. (2017). Problematic gaming exists and is an example of disordered gaming: commentary on: Scholars' open debate paper on the World Health Organization ICD-11 Gaming Disorder proposal (Aarseth et al.). *Journal of Behavioral Addictions*, 6(3), 296–301. <https://doi.org/10.1556/2006.6.2017.037>.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis. A global perspective (seventh edition)*. Upper Saddle River: Pearson Prentice Hall.
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53–60.
- Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural equation modeling: concepts, issues, and applications* (pp. 158–176). Thousand Oaks: Sage Publications.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>.
- Hughes, M. E., Waite, L. J., Hawkey, L. C., & Cacioppo, J. T. (2004). A short scale for measuring loneliness in large surveys: results from two population-based studies. *Research on Aging*, 26(6), 655–672. <https://doi.org/10.1177/0164027504268574>.
- 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. <https://doi.org/10.1111/add.14388>.
- 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(3), 331–342. <https://doi.org/10.1016/j.cpr.2013.01.002>.
- Király, O., Slezcka, P., Pontes, H. M., Urbán, R., Griffiths, M. D., & Demetrovics, Z. (2017). Validation of the ten-item Internet Gaming Disorder Test (IGDT-10) and evaluation of the nine DSM-5 Internet Gaming Disorder criteria. *Addictive Behaviors*, 64, 253–260. <https://doi.org/10.1016/j.addbeh.2015.11.005>.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling (third edition)*. New York: The Guilford Press.
- Koronczai, B., Urbán, R., Kökönyei, G., Paksi, B., Papp, K., Kun, B., et al. (2011). Confirmation of the three-factor model of problematic Internet use on off-line adolescent and adult samples. *Cyberpsychology, Behavior and Social Networking*, 14(11), 657–664. <https://doi.org/10.1089/cyber.2010.0345>.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression measure. *Journal of General Internal Medicine*, 16(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>.
- Krossbakken, E., Pallesen, S., Mentzoni, R. A., King, D. L., Molde, H., Finserås, T. R., & Torsheim, T. (2018). A cross-lagged study of developmental trajectories of video game engagement, addiction, and mental health. *Frontiers in Psychology*, 9(2239). <https://doi.org/10.3389/fpsyg.2018.02239>.
- Kuss, D. J., & Pontes, H. M. (2019). *Internet addiction* (Vol. 41). Boston: Hogrefe Publishing Corporation.
- Lee, Z. W. Y., Cheung, C. M. K., & Chan, T. K. H. (2015). Massively multiplayer online game addiction: instrument development and validation. *Information & Management*, 52(4), 413–430. <https://doi.org/10.1016/j.im.2015.01.006>.
- Lee, D., Lee, J., Namkoong, K., & Jung, Y. C. (2018). Subregions of the anterior cingulate cortex form distinct functional connectivity patterns in young males with internet gaming disorder with comorbid depression. *Frontiers in Psychiatry*, 9(380). <https://doi.org/10.3389/fpsyg.2018.00380>.
- Lemmens, J. S., Valkenburg, P. M., & Peter, J. (2009). Development and validation of a game addiction scale for adolescents. *Media Psychology*, 12(1), 77–95. <https://doi.org/10.1080/15213260802669458>.
- Lemmens, J. S., Valkenburg, P. M., & Gentile, D. A. (2015). The Internet Gaming Disorder Scale. *Psychological Assessment*, 27(2), 567–582. <https://doi.org/10.1037/pas0000062>.
- Long, J., Liu, T., Liu, Y., Hao, W., Mauraage, P., & Billieux, J. (2018). Prevalence and correlates of problematic online gaming: a systematic review of the evidence published in Chinese. *Current Addiction Reports*, 5(3), 359–371. <https://doi.org/10.1007/s40429-018-0219-6>.
- Lopez-Fernandez, O., Honrubia-Serrano, M. L., Baguley, T., & Griffiths, M. D. (2014). Pathological video game playing in Spanish and British adolescents: towards the exploration of internet gaming disorder symptomatology. *Computers in Human Behavior*, 41, 304–312. <https://doi.org/10.1016/j.chb.2014.10.011>.

- Merz, E. L., Malcarne, V. L., Roesch, S. C., Riley, N., & Sadler, G. R. (2011). A multigroup confirmatory factor analysis of the Patient Health Questionnaire-9 among English- and Spanish-speaking Latinas. *Cultural Diversity and Ethnic Minority Psychology, 17*(3), 309–316. <https://doi.org/10.1037/a0023883>.
- Myrseth, H., Olsen, O. K., Borud, E. K., & Strand, L. Å. (2017). Predictors of gaming behavior among military peacekeepers – exploring the role of boredom and loneliness in relation to gaming problems. *Journal of Military Studies*. <https://doi.org/10.1515/jms-2017-0001>.
- Petry, N. M., & O'Brien, C. P. (2013). Internet gaming disorder and the DSM-5. *Addiction, 108*(7), 1186–1187. <https://doi.org/10.1111/add.12162>.
- Pontes, H. M. (2016). Current practices in the clinical and psychometric assessment of internet gaming disorder in the era of the DSM-5: a mini review of existing assessment tools. *Mental Health and Addiction Research, 1*(1), 18–19. <https://doi.org/10.15761/MHAR.1000105>.
- Pontes, H. M. (2017). Investigating the differential effects of social networking site addiction and internet gaming disorder on psychological health. *Journal of Behavioral Addictions, 6*(4), 601–610. <https://doi.org/10.1556/2006.6.2017.075>.
- Pontes, H. M. (2018). Making the case for video game addiction: does it exist or not? In C. J. Ferguson (Ed.), *Video game influences on aggression, cognition, and attention* (pp. 41–57). Cham: Springer International Publishing.
- Pontes, H. M., & Griffiths, M. D. (2014). Assessment of internet gaming disorder in clinical research: past and present perspectives. *Clinical Research and Regulatory Affairs, 31*(2–4), 35–48. <https://doi.org/10.3109/10601333.2014.962748>.
- Pontes, H. M., & Griffiths, M. D. (2015a). Internet gaming disorder and its associated cognitions and cognitive-related impairments: a systematic review using PRISMA guidelines. *Revista Argentina de Ciencias del Comportamiento, 7*(3), 102–118.
- Pontes, H. M., & Griffiths, M. D. (2015b). Measuring DSM-5 internet gaming disorder: development and validation of a short psychometric scale. *Computers in Human Behavior, 45*, 137–143. <https://doi.org/10.1016/j.chb.2014.12.006>.
- Pontes, H. M., & Griffiths, M. D. (2015c). New concepts, old known issues: The DSM-5 and internet gaming disorder and its assessment. In J. Bishop (Ed.), *Psychological and social implications surrounding internet and gaming addiction* (pp. 16–30). Hershey: IGI Global.
- Pontes, H. M., & Griffiths, M. D. (2016). Portuguese validation of the internet gaming disorder scale–short-form. *CyberPsychology, Behavior & Social Networking, 19*(4), 288–293. <https://doi.org/10.1089/cyber.2015.0605>.
- Pontes, H. M., Király, O., Demetrovics, Z., & Griffiths, M. D. (2014). The conceptualisation and measurement of DSM-5 internet gaming disorder: the development of the IGD-20 test. *PLoS One, 9*(10), e110137. <https://doi.org/10.1371/journal.pone.0110137>.
- Russell, D. W. (1996). UCLA Loneliness Scale (version 3): reliability, validity, and factor structure. *Journal of Personality Assessment, 66*(1), 20–40. https://doi.org/10.1207/s15327752jpa6601_2.
- Schivinski, B., Brzozowska-Woś, M., Buchanan, E. M., Griffiths, M. D., & Pontes, H. M. (2018). Psychometric assessment of the internet gaming disorder diagnostic criteria: an item response theory study. *Addictive Behaviors Reports, 8*, 176–184. <https://doi.org/10.1016/j.abrep.2018.06.004>.
- Stavropoulos, V., Beard, C., Griffiths, M. D., Buleigh, T., Gomez, R., & Pontes, H. M. (2018). Measurement invariance of the Internet Gaming Disorder Scale–Short-Form (IGDS9-SF) between Australia, the USA, and the UK. *International Journal of Mental Health and Addiction, 16*(2), 377–392. <https://doi.org/10.1007/s11469-017-9786-3>.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics (sixth edition)*. Essex: Pearson Education.
- Tsai, F. J., Huang, Y. H., Liu, H. C., Huang, K. Y., Huang, Y. H., & Liu, S. I. (2014). Patient Health Questionnaire for school-based depression screening among Chinese adolescents. *Pediatrics, 133*(2), e402–e409.
- Van Rooij, A. J., Schoenmakers, T. M., & van de Mheen, D. (2017). Clinical validation of the C-VAT 2.0 assessment tool for gaming disorder: a sensitivity analysis of the proposed DSM-5 criteria and the clinical characteristics of young patients with ‘video game addiction’. *Addictive Behaviors, 64*, 269–274. <https://doi.org/10.1016/j.addbeh.2015.10.018>.
- van Rooij, A. J., Ferguson, C. J., Colder Carras, M., Kardefelt-Winther, D., Shi, J., Aarseth, E., et al. (2018). A weak scientific basis for gaming disorder: let us err on the side of caution. *Journal of Behavioral Addictions, 7*(1), 1–9. <https://doi.org/10.1556/2006.7.2018.19>.
- World Health Organization. (2018a). Gaming disorder. Retrieved from: <http://www.who.int/features/qa/gaming-disorder/en/>.
- World Health Organization. (2018b). ICD-11 beta draft: gaming disorder. Retrieved from <http://apps.who.int/classifications/icd11/browse/f/en#/http%3a%2f%2fid.who.int%2fid%2fentity%2f1448597234>.
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: a content analysis and recommendations for best practices. *The Counseling Psychologist, 34*(6), 806–838. <https://doi.org/10.1177/0011000006288127>.

- Xu, S., Qiu, D., Hahne, J., Zhao, M., & Hu, M. (2018). Psychometric properties of the short-form UCLA Loneliness Scale (ULS-8) among Chinese adolescents. *Medicine*, *97*(38), e12373. <https://doi.org/10.1097/MD.00000000000012373>.
- Yam, C. W., Pakpour, A. H., Griffiths, M. D., Yau, W. Y., Lo, C. L. M., Ng, J. M. T., et al. (2018). Psychometric testing of three Chinese online-related addictive behavior instruments among Hong Kong university students. *Psychiatric Quarterly*, *90*, 117–128. <https://doi.org/10.1007/s11126-018-9610-7>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Halley M. Pontes¹ · Bruno Schivinski² · Cornelia Sindermann³ · Mei Li⁴ · Benjamin Becker⁵ · Min Zhou⁶ · Christian Montag^{3,5}

¹ Division of Psychology, School of Medicine, University of Tasmania, Launceston, Australia

² Department of Management, Birkbeck, University of London, Malet Street, Bloomsbury, London WC1E 7HX, UK

³ Department of Molecular Psychology, Institute of Psychology and Education, Ulm University, 89081 Ulm, Germany

⁴ Student Counseling Center, Beijing University of Civil Engineering and Architecture, Beijing, China

⁵ The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Laboratory for Neuroinformation, University of Electronic Science and Technology of China, Chengdu 611731, China

⁶ Institute of Medical Statistics, Informatics and Epidemiology, University of Cologne, Cologne, Germany