Validation of the Ten-Item Internet Gaming Disorder Test (IGDT-10) and evaluation of the nine DSM-5 Internet gaming disorder criteria

Abstract

Introduction: The inclusion of Internet gaming disorder (IGD) in the DSM-5 (Section 3) has given rise to much scholarly debate regarding the proposed criteria and their operationalization. The present study's aim was threefold: to (i) develop and validate a brief psychometric instrument (Ten-Item Internet Gaming Disorder Test; IGDT-10) to assess IGD using definitions suggested in DSM-5, (ii) contribute to ongoing debate regards the usefulness and validity of each of the nine IGD criteria (using Item Response Theory [IRT]), and (iii) investigate the cut-off threshold suggested in the DSM-5.

Methods: An online gamer sample of 4,887 gamers (age range 14-64 years, mean age 22.2 years [SD=6.4], 92.5% male) was collected through *Facebook* and a gaming-related website with the cooperation of a popular Hungarian gaming magazine. A shopping voucher of approx. 300 Euros was drawn between participants to boost participation (i.e., lottery incentive).

Results: Analysis supported IGDT-10's validity, reliability, and suitability to be used in future research. Findings of the IRT analysis suggest IGD is manifested through a different set of symptoms depending on the level of severity of the disorder. More specifically, "continuation", "preoccupation", "negative consequences" and "escape" were associated with lower severity of IGD, while "tolerance", "loss of control", "giving up other activities" and "deception" criteria were associated with more severe levels. "Preoccupation" and "escape" provided very little information to the estimation IGD severity. Finally, the DSM-5 suggested threshold appeared to be supported by our statistical analyses.

Conclusions: IGDT-10 is a valid and reliably instrument to assess IGD as proposed in the DSM-5. Apparently the nine criteria do not explain IGD in the same way, suggesting that additional studies are needed to assess the characteristics and intricacies of each criterion and how they account to explain IGD.

Keywords

Internet gaming disorder; online game; behavioral addiction; psychometric validation; Item Response Theory; Structural Equation Modeling

1. Introduction

The assessment of problematic gaming (or gaming addiction) has been subject of controversy since it has gained scientific attention (Griffiths, Király, Pontes, & Demetrovics, 2015; King, Haagsma, Delfabbro, Gradisar, & Griffiths, 2013; Kuss & Griffiths, 2012). Over 20 instruments have been developed to assess problematic gaming but few have been properly validated and the lack of consensual definition led to inconsistencies in the assessment criteria (see King et al., 2013; Király, Nagygyörgy, Koronczai, Griffiths, & Demetrovics, 2015; Pontes & Griffiths, 2014). The inclusion of Internet gaming disorder (IGD) in Section 3 of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2013), led to an upsurge of new psychometric tools designed to assess problematic gaming according to the nine DSM-5 IGD criteria (e.g., Lemmens, Valkenburg, & Gentile, 2015; Pontes & Griffiths, 2015; Pontes, Király, Demetrovics, & Griffiths, 2014; Rehbein, Kliem, Baier, Mößle, & Petry, 2015). Additionally, the concept of IGD has stirred debate among scholars in terms of the validity of each nine IGD criteria and how to better operationally define such criteria in light of knowledge on problematic gaming accumulated over three decades (Griffiths, van Rooij, et al., 2015; Király, Griffiths, & Demetrovics, 2015; van Rooij & Prause, 2014).

To date, most of the newly developed IGD assessment measures have two main shortcomings. Firstly, some of the new measures are arguably lengthy, thus not optimal for large-scale surveys using questionnaires that comprise several scales necessary to test complex associations regarding IGD. Secondly, some of the new instruments (including short ones) either do not specifically reflect the nine IGD criteria, or use convoluted wording that might be difficult for gamers to understand and answer. Given this, the first aim of this study is to develop and validate psychometrically, a new brief instrument to assess IGD using definitions suggested in the DSM-5, while also adopting a simple, clear, and more consistent item-wording that adequately reflects the IGD concept.

Furthermore, a few recently published studies (Ko et al., 2014; Lemmens et al., 2015; Rehbein et al., 2015) examining the usefulness and validity of all nine IGD criteria provided interesting and

yet conflicting results. For instance, Rehbein and colleagues (2015) found that the criteria "give up other activities," "tolerance" and "withdrawal" were of key importance for identifying IGD, while Lemmens and colleagues (2015) found "escape" did not add to diagnostic accuracy due to lack of specificity. As a second aim, the present authors contribute to this debate by further investigating the nine IGD criteria as proposed by the DSM-5 and examining how each IGD criterion performs at different severity levels using an Item Response Theory (IRT) approach.

Moreover, little is known empirically about the validity of the proposed cut-off threshold for IGD (i.e., endorsing five or more criteria out of the nine) in the DSM-5 as it was conservatively chosen by the APA, because low thresholds might inflate diagnoses and result in classifying individuals who have not suffered significant clinical impairment (Petry et al., 2014). To the authors' knowledge, only one study (i.e., Ko et al., 2014) has examined the diagnostic validity of the nine IGD criteria and tested the proposed cut-off threshold for IGD. The study concluded that endorsing five or more of the nine IGD criteria was the best cut-off threshold to differentiate gamers with IGD from healthy gamers. A final aim is to contribute to this question by empirically testing the suggested IGD cut-off points with Latent Class Analysis (LCA), sensitivity, and specificity analysis in a large heterogeneous sample of online gamers.

2. Materials and methods

The study was approved by the Institutional Review Board of the Eötvös Loránd University, Budapest, Hungary.

2.1. Participants and procedure

Data collection took place online with the cooperation of a popular Hungarian gaming magazine (i.e., *GameStar*). The magazine's *Facebook* page had approximately 65,000 followers during data collection and an associated online website. A participation call was posted online via the magazine's website and *Facebook* page three times during August to September 2014. Prior to starting the survey, all participants were informed about the goals of the study. They were assured about anonymity and confidentiality, and their informed consent was obtained by ticking into a box if they agreed to continue and participate in the study (14-17 years old children had to tick another box too for parental permission). A shopping voucher of 90000 HUF (approx. 300€) was drawn between participants that fully completed the survey. A total of 7,757 gamers started the

survey. After excluding cases with severe incompleteness or inconsistencies (n=2,870, 37.0%), 4,887 gamers remained, corresponding to a response rate of 63.0%.

2.2. Measures

Ten-Item Internet Gaming Disorder Test (IGDT-10) comprises 10 items and assesses levels of IGD. The instrument was developed theoretically, based on the nine DSM-5 criteria, and also taking into consideration Petry et al.'s (2014) recommendations. During instrument development, the authors sought to operationalize IGD throughout the nine criteria suggested in the DSM-5 in a brief and simple way by adopting a clear wording for each item. The IGD criterion referring to "*jeopardy or losing a significant relationship, job, or educational or career opportunity because of participation in Internet games*" was operationalized via two items given its complexity. Response options for the ten items were "never", "sometimes", and "often" instead of "yes" and "no". The 3-point Likert scale was preferred in order to facilitate the responses given by participants as it makes the decision easier and more realistic. However, during scoring the IGDT-10 items were recoded into a "yes" (1) and "no" (0) format in order to resemble the dichotomous structure of IGD in DSM-5. Since IGD criteria in the DSM-5 suggest that behaviors or problems are frequently repeated or continuously present, only "often" answers were recoded as "yes". Given that question 9 and 10 related to the same criterion, they are combined in the scoring, that is, answering "Often" on either Item 9 or Item 10 (or both items) scores only 1 point (see the Appendix). Consequently, the composite score of IGDT-10 ranges from 0 to 9, higher scores indicating more severe cases of IGD. Cronbach's alpha of the scale was .68.

Problematic Online Gaming Questionnaire (POGQ; Demetrovics et al., 2012). The POGQ is a standardized tool with good validity and reliability indices. It comprises 18 items used to assess problematic online gaming via six factors (i.e., preoccupation, immersion, overuse, social isolation, interpersonal conflicts, and withdrawal). Items were rated on a 5-point Likert scale (from "never" to "almost always/always") with higher scores meaning higher risk of problematic play. Cronbach's alpha for this instrument in the present sample was .90. This instrument was used to ascertain the validity of the IGDT-10.

Brief Symptom Inventory (BSI; Derogatis, 1975) comprises 53 items and assessed psychiatric distress. This instrument assesses nine self-reported clinically relevant psychological symptoms:

somatization, obsession-compulsion, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism. Answers were given on a 5-point Likert scale from "not at all" to "extremely". In the present study, a summarized index (i.e., The Global Severity Index [GSI]) assessed the intensity of general distress where higher scores indicated higher levels of general distress. Cronbach's alpha for this measure in the present sample was .96.

Weekly gaming time was measured as an ordinal variable with six response options: "None"; "Less than 7 hours weekly"; "7-14 hours weekly"; "15-28 hours weekly"; "29-42 hours weekly"; and "More than 42 hours weekly", and it was linearized for the analysis (first and last values were retained with a score of 0 and 42, while intervals were recoded with their mid-points).

2.3. Statistical analysis

2.3.1. Validation of the IGDT-10

To test dimensionality and construct validity of the IGDT-10, confirmatory factor analysis (CFA) was used with robust weighted least square (WLSMV) estimation method in *Mplus* 7.3 (Muthén & Muthén, 1998-2012). To evaluate model fit, multiple indices were used, including the chi-square (χ2) value, the comparative fit index (CFI), the Tucker–Lewis Fit index (TLI), the root mean square error of approximation (RMSEA), its 90% confidence interval (90% CI), and *p* value smaller than .05 for test of close fit. The chi-square test should be non-significant ($p > .05$) for good fitting models. However, this index is sensitive to large sample sizes, therefore goodness of fit was examined using the other indices. Hu and Bentler (1999) suggested an RMSEA value less than 0.06 and CFI and TLI values above 0.95 for good model fit. To further verify the construct validity of the IGDT-10 a pattern of covariates was tested with a fully saturated structural regression model in *Mplus* 7.3, having IGDT-10 and POGQ as endogenous variables, and psychiatric distress and gaming time as exogenous variables. All composite scores (i.e., IGDT-10, POGQ, and psychiatric distress) were entered in the model as latent variables with single indicators in order to control for measurement error (Cole & Preacher, 2014).

2.3.2. Evaluating the nine DSM criteria using Item Response Theory

IRT is a theoretical framework and complementary to Classical Test Theory (CTT). Both approaches share the assumption that measurement involves placing of an individual on the

continuum of the underlying latent variable, in this case the severity of IGD. The IRT has nevertheless some important advantages over the CTT. The IRT models place individuals and items[1] on the same latent variable. Thus, item properties remain invariant with respect to the analyzed sample (e.g., Reeve & Fayers, 2005).

A two-parameter logistic Item Response Theory model (2PL-IRT model) was computed using *Mplus* 7.3. The 2PL-IRT describes measurement performance of each criterion with two parameters. The threshold parameter (*b* parameter) is related to severity of a criterion and its higher value indicates that the criterion is fulfilled only in more severe stages of the disorder. The discrimination parameter (*a* parameter) refers to the accuracy of the criterion in distinguishing between respondents below and above the aforementioned threshold (*b*). Based on the two parameters, an information function can be estimated. It describes how informative (in terms of measurement) criteria on different levels of disorder severity are.

The fit indices for the Item Response Theory (IRT) were equivalent to those obtained in CFA analysis. Additionally, the assumptions of the 2PL-IRT model were verified. The question in regard to unidimensionality was subject to CFA analysis. Monotonicity assumption was verified based on the evaluation of the ratio of respondents endorsing each criterion, nine groups fulfilling different number of criteria. Residual correlation matrix and model indices were analysed to identify associations between the criteria after controlling for the underlying latent factor.

2.3.3. Evaluating IGD's suggested cut-off value

In order to test the cut-off value suggested in the DSM-5 (i.e., endorsement of five or more of nine IGD criteria) a latent class analysis (LCA) was carried out in *Mplus* 7.3 to identify problematic gamers. LCA is a mixture modeling technique used to identify groups of people (categorical latent variables) who give similar responses to specific variables (Collins & Lanza, 2010), in this case the responses given by gamers to the IGDT-10 criteria (dichotomous manifest variables). The group with the highest likelihood of meeting the nine criteria was used as the gold standard to determine the optimal cut-off value that separated this group from the remaining sample. The sensitivity, specificity, positive, negative predictive value, and diagnostic accuracy of each cut-off threshold were calculated and compared to identify the best-fitting value and compare it to the one suggested for IGD in DSM-5.

---

[1]The terms *item* and *criterion* are used interchangeably in the description of statistical analyses.

The LCA analysis was carried out with one to five classes. To determine the number of latent classes, multiple indices were used: measures of parsimony (i.e., Akaike Information Criteria [AIC], Bayesian Information Criteria [BIC], and Sample-size adjusted Bayesian Information Criteria [SSABIC]) with lower values indicating more parsimonious models, the Entropy criterion that determines the accuracy of classifying people into their respective classes (higher values indicating better fit), and the Lo-Mendell-Rubin Adjusted Likelihood Ratio Test (LMR-LRT) that statistically compares the fit of the tested model (e.g., four-class model) to a model with one class less (e.g., three-class model). A significant p value in LMR-LRT indicates the tested model fits better than the previous one (Muthén & Muthén, 1998-2012). Missing data were treated with Full Information Maximum Likelihood (FIML) method in Mplus. Descriptive statistics, prevalence, and criteria endorsement analyses were carried out using SPSS version 22 (IBM Corp., 2013).

3. Results

3.1. Descriptive statistics

After deleting cases with severe incompleteness or inconsistencies, the sample comprised 4,887 online gamers (mean age 22.2 years [SD = 6.4], age range 14-64 years, 92.5% male), most of whom played for more than 15 hours weekly. The mean number of completed years in education was 12.4 years (SD=3.2), and most were single and still studying in education (see Table 1).

Table 1

3.2. Validation of the IGDT-10

A one-factor solution of the nine IGD items was tested with CFA. The model provided a good fit to the data ($\chi 2$=194.4 df=27 p<.0001; CFI=0.971; TLI=0.962; RMSEA=0.036 [0.031-0.040] Cfit>0.90 pclose=1.000; N=4,887). Factor loadings of all nine indicators of the IGDT-10 were above the conventional threshold of .50 (see Table 2). The pattern of covariates of IGDT-10 was tested with a structural regression model (see Figure 1). The two instruments measuring problematic online gaming (i.e., IGDT-10 and POGQ) were strongly correlated (r=.77, p<.001). Additionally, their associations with psychiatric distress and gaming time were comparable, IGDT-10 having slightly higher associations with all three explanatory variables.

Table 2/Figure 1

3.3. IGD prevalence and criteria endorsement

Based on five criteria endorsement suggested in DSM-5 as the cut-off threshold for classifying disordered gamers, IGD prevalence rate in the present sample was 2.9% (n=138). Basic demographic characteristics and weekly gaming time of the participants with and without IGD are outlined in Table 1. The endorsement of each IGD criterion in the total sample and among disordered gamers is in Table 2. "Preoccupation" was the most frequently endorsed criterion (20.1%) among all gamers, while "continuation" was the most commonly endorsed criterion among disordered gamers (92.8%).

3.4. Evaluating the nine DSM criteria using IRT

All assumptions of the 2PL-IRT model were verified. The probability of endorsing each criterion increased with the number of criteria fulfilled, indicating monotonicity. After controlling for the underlying latent factor, no significant correlations between the criteria were found. Fit indices for the IRT model were the same as for the CFA, and therefore acceptable. The Item Characteristic Curves are presented in Table 2 and graphically displayed in Figure 2.

Figure 2

Generally (see Figure 2), a specific criteria endorsement pattern was identified in the responses. The probability (Y-axis) of fulfilling each criterion differed depending on the severity of IGD (X-axis). The criteria "continuation", "preoccupation", "negative consequences" and "escape" were associated with lower IGD severity. "Tolerance", "loss of control", "giving up other activities" and "deception" criteria were associated with more severe levels, and had similar threshold (b) parameters. Highest severity ($b = 2.54$) was noted for "withdrawal" criterion, indicating that withdrawal is associated with most severe IGD symptoms.

Regarding criterion discrimination accuracy (a parameter), some differences were found in performance of various criteria. While most had similar discriminatory power, "preoccupation" and "escape" criteria showed significantly lower values, indicating worse accuracy. Based on threshold and discrimination parameters, Item (criteria) Information Curves were estimated. These indicators (see Figure 3) describe how much information each criterion delivers for estimation of IGD severity. Due to different threshold parameters, each criterion delivered

different amounts of information on various IGD severity stages. In general, criteria with higher discriminatory power delivered more information.

Figure 3

When compared to other criteria, "preoccupation" and "escape" provided the least information. Conversely, "continuation", followed by "negative consequences" and "giving up other activities" provided the most information. The "continuation" criterion provided considerably more information on less severe IGD levels, suggesting it accurately measures the severity of IGD problems already among individuals with low severity of IGD.

3.5. Evaluating the suggested cut-off value

An LCA analysis with two to five latent classes was performed on the the IGDT-10 criteria to test the cut-off value suggested in the DSM-5. Fit indices and test values are presented in Table 3. According to the criteria listed in the 'Statistical analyses' section, the four-class solution was selected. Estimated probabilities of the four latent classes are presented in Figure 4. The fourth class (i.e., 2.6%, N=4,887) represents gamers with the highest estimated probabilities of endorsing all nine IGD criteria, and therefore considered the disordered group. "Continuation" and "negative consequences" criteria had the highest estimated probabilities of being endorsed by these gamers.

Table 3-4/Figure 4

To test whether the cut-off value suggested by the DSM-5 (i.e., five or more criteria) fitted the data empirically, the fourth LCA group was used as the gold standard to determine the optimal cut-off threshold to distinguish gamers of this group from the remaining sample (Table 4). Based on sensitivity, specificity, positive and negative predictive value, and diagnostic accuracy of each cut-off score, the empirical data supported DSM-5's suggestion for IGD cut-off threshold (i.e., endorsement of five or more criteria).

4. Discussion

The primary aim of this study was to develop and validate a new psychometric tool to assess IGD that overcomes the shortcomings present in extant IGD instruments. Furthermore, this study also contributed to ongoing debate regards the usefulness and validity of each of the nine IGD criteria,

9

and investigated the cut-off threshold suggested in DSM-5. As for validity and reliability of the IGDT-10, CFA demonstrated a one-factor structure. This suggests IGD can be assessed using a first-order latent variable reflecting the nine criteria via multiple indicators. The structural regression model testing the pattern of covariates showed high correlation with POGQ, an instrument measuring the same construct, and comparable associations of the two instruments with explanatory variables. Moderate associations with psychiatric distress and weak to moderate associations with gaming time were found in line with the literature (Brunborg, Mentzoni, & Frøyland, 2014; Hsu, Wen, & Wu, 2009; Király, Urbán, et al., 2015; Starcevic, Berle, Porter, & Fenech, 2011; Zanetta Dauriat et al., 2011) and therefore demonstrate construct validity of the instrument. Internal consistency was satisfactory, however, not too high, which may be related to the limited number of binary variables analyzed.

The results of IRT analysis demonstrated the criteria "continuation", "preoccupation", "negative consequences", and "escape" were endorsed more frequently in less severe IGD stages and "tolerance", "loss of control", "giving up other activities", and "deception" were only reported in more severe cases. Based on conditional inference tree analysis, Rehbein et al. (2015) reported that the criterion "giving up other activities" and "tolerance" were key in identifying IGD. Being associated with more severe IGD levels in the present study may help explain why endorsement of these criteria corresponds to a high probability of a positive IGD diagnosis (Rehbein et al., 2015).

In parallel to these results, criteria endorsement rates suggested that criterion "preoccupation" followed by "continuation" were the two most frequently endorsed criteria, whereas "withdrawal" and "giving up other activities" were least endorsed among the whole sample. Furthermore, the endorsement rate pattern was different amongst those meeting five or more (of nine) IGD criteria as "continuation" followed by "negative consequences" were endorsed more often in comparison to the criteria "withdrawal" and "deception". These findings contrast with those of two recent studies (i.e., Lemmens et al., 2015; Rehbein et al., 2015) that found "escape" was the most frequently endorsed criteria, while "negative consequences" (termed as "conflict" in Lemmens et al.'s study, and "risk/lose relationship/opportunities" in Rehbein et al.'s study) was the least endorsed in both the whole sample and among those meeting five or more IGD criteria. It remains unclear whether these differences derive from the different criteria operationalization or the diversity of samples (i.e., different mean age, gender rate, and/or cultural differences).

The present study's results suggest that while the underlying structure for IGD appears to be unidimensional, the phenomenon is primarily manifested via different sets of symptoms depending on the level of disorder severity. Furthermore, the discriminatory (a) and threshold (b) parameters of each criterion provided information about the accuracy of the proposed set of IGD criteria in differentiating between gamers with less and more severe IGD levels. The accuracy of the symptoms differed depending on IGD severity, and was best at more severe levels of the disorder (peaking around $b = 2$), suggesting that the nine IGD symptoms differentiate better among individuals with more severe gaming-related problems. However, "preoccupation" and "escape" criteria had very low discriminatory power, and thus provided little information to the estimation of IGD severity. This parallels the findings of Rehbein et al. (2015) who found that "escape" and "preoccupation" poorly predicted IGD despite being endorsed at high rates.

Additionally, Lemmens et al. (2015) found that "escape" had the lowest specificity in distinguishing between disordered and non-disordered gamers, while Ko et al. (2014) reported that "deception" and "escape" had the lowest diagnostic accuracy to discriminate individuals with IGD from non-problematic individuals. Therefore, replacing these two criteria with new ones that better discriminate (high a parameter) disordered gamers in less severe stages (low b parameter) of IGD might be beneficial. Although it is not the primary goal of DSM-5, implementing such changes in the diagnostic IGD criteria, might also facilitate more reliable assessment among "diagnostic orphans" (i.e., gamers endorsing less than five IGD criteria), which is very valuable from the perspective of research.

IGD prevalence rates were estimated on the basis of endorsement of at least five of nine IGD criteria, and via an LCA by selecting gamers with membership in the class with the highest probability of endorsing all nine criteria. Based on endorsement of five criteria as the cut-off, the IGD prevalence rate in the sample was 2.9%, while, the prevalence based on LCA results was slightly more conservative (2.6%). By adopting a similar strategy to derive IGD prevalence rates in a representative sample of Dutch adolescents, Lemmens et al. (2015) found a 6.6% prevalence rate among gamers, and 5.4% in the whole sample using a threshold endorsement of at least five of nine criteria. However, by using a stricter cut-off (i.e., endorsement of at least six criteria) a 5.2% prevalence rate was reported among gamers and 4% in the whole sample (Lemmens et al., 2015). The lowest IGD prevalence rate reported also by Lemmens et al. (2015) was 4.9% (among gamers) and was derived from the results of an LCA. Additionally, Rehbein et al. (2015) reported

that IGD prevalence was around 1.16% in a representative sample of German adolescents using the cut-off of endorsing five of nine IGD criteria. These prevalence rate discrepancies might also be a consequence of different operationalization of IGD criteria and diversity of samples.

The results from the sensitivity, specificity, positive and negative predictive value, and diagnostic accuracy analysis of each possible cut-off score for IGD appeared to support the DSM-5 suggested threshold (i.e., endorsement of at least five criteria). Drawing on different methodology and study design, Ko et al. (2014) found similar results in a clinical sample of 75 Taiwanese young adults. Future clinical validation studies using samples of Western individuals should be conducted to provide further information on the suitability of IGD cut-off as the findings of this study were entirely derived on the basis of statistical testing.

This study has several limitations that have to be taken in consideration when interpreting or using the results. Although the sample was large and heterogeneous, participants were recruited using non-probability sampling techniques. Additionally, the data was entirely collected using self-report questionnaires, which are known for producing potential biases (e.g., social desirability biases, short-term recall biases, etc.).

5. Conclusion

Generally, the present findings support the validity and reliability of the IGDT-10 to assess IGD using the DSM-5 framework and corroborate the use of the suggested cut-off threshold for classifying individuals with IGD as proposed by DSM-5. The findings of IRT analysis in regard to the criteria endorsement pattern suggested that IGD is manifested differently depending on its severity. For instance, while "continuation" and "negative consequences" were already fulfilled by individuals with less severe IGD, withdrawal symptoms were reported only in most severe cases. Furthermore, as "preoccupation" and "escape" repeatedly yielded low psychometric indices, we suggest their replacement by others (or even exclusion). Our findings pave the way for future studies to assess the characteristics and intricacies of each criterion and how they explain IGD.

Figure captions

Figure 1. Multivariate multiple regression model to test construct validity of IGDT-10. *Note: IGDT-10: Ten-Item Internet Gaming Disorder Test; POGQ: Problematic Online Gaming Questionnaire; values on the simple arrows are standardized regression coefficients; double-headed arrows indicate correlations between the two variables; *** *p<.001*

Figure 2. Criterion response curves for the IGD criteria.

Figure 3. Criterion information curves for the IGD criteria.

Figure 4. Latent class analysis on the nine IGD criteria of the IGDT-10.

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders – Text Revision. 5th edition*. DC, USA: American Psychiatric Association.

Brunborg, G. S., Mentzoni, R. A., & Frøyland, L. R. (2014). Is video gaming, or video game addiction, associated with depression, academic achievement, heavy episodic drinking, or conduct problems? *Journal of Behavioral Addictions, 3*(1), 27-32. doi: 10.1556/JBA.3.2014.002

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods, 19*(2), 300-315. doi: 10.1037/a0033805

Collins, L., & Lanza, S. (2010). *Latent class and latent transition analysis*. Hoboken, NJ: Wiley.

Demetrovics, Z., Urbán, R., Nagygyörgy, K., Farkas, J., Griffiths, M. D., Pápay, O., . . . Oláh, A. (2012). The development of the Problematic Online Gaming Questionnaire (POGQ). *PLoS ONE, 7*(5), e36417. doi: 10.1371/journal.pone.0036417

Derogatis, L. R. (1975). *Brief Symptom Inventory*. Baltimore, MD: Clinical Psychometric Research.

Griffiths, M. D., Király, O., Pontes, H. M., & Demetrovics, Z. (2015). An overview of problematic gaming. In V. Starcevic & E. Aboujaoude (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., . . . Demetrovics, Z. (2015). Working towards an international consensus on criteria for assessing Internet Gaming Disorder: A critical commentary on Petry et al (2014). *Addiction, (in press)*.

Hsu, S. H., Wen, M. H., & Wu, M. C. (2009). Exploring user experiences as predictors of MMORPG addiction. *Computers & Education, 53*(3), 990-999. doi: 10.1016/j.compedu.2009.05.016

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling, 6*(1), 1-55. doi: 10.1080/10705519909540118

IBM Corp. (2013). IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

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. doi: 10.1016/j.cpr.2013.01.002

Király, O., Griffiths, M. D., & Demetrovics, Z. (2015). Internet gaming disorder and the DSM-5: Conceptualization, debates, and controversies. *Current Addiction Reports, 2*, 254-262. doi: 10.1007/s40429-015-0066-7.

Király, O., Nagygyörgy, K., Koronczai, B., Griffiths, M. D., & Demetrovics, Z. (2015). Assessment of Problematic Internet Use and Online Video Gaming. In E. Aboujaoude & V. Starcevic (Eds.), *Mental Health in the Digital Age: Grave Dangers, Great Promise* (pp. 46-68). Oxford: Oxford University Press.

Király, O., Urbán, R., Griffiths, M. D., Ágoston, C., Nagygyörgy, K., Kökönyei, G., & Demetrovics, Z. (2015). Psychiatric symptoms and problematic online gaming: The mediating effect of gaming motivation. *Journal of Medical Internet Research, 17*(4), e88. doi: 10.2196/jmir.3515

Ko, C. H., Yen, J. Y., Chen, S. H., Wang, P. W., Chen, C. C., & Yen, C. F. (2014). Evaluation of the diagnostic criteria of Internet gaming disorder in the DSM-5 among young adults in Taiwan. *Journal of Psychiatric Research, 53*, 103-110. doi: 10.1016/j.jpsychires.2014.02.008

Kuss, D. J., & Griffiths, M. D. (2012). Online gaming addiction in children and adolescents: a review of empirical research. *Journal of Behavioral Addictions, 1*(1), 3-22. doi: 10.1556/JBA.1.2012.1.1

Lemmens, J. S., Valkenburg, P. M., & Gentile, D. A. (2015). The Internet Gaming Disorder Scale. *Psychological Assessment, 27*(2), 567-582. doi: 10.1037/pas0000062

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.

Petry, N. M., Rehbein, F., Gentile, D. A., Lemmens, J. S., Rumpf, H. J., Mossle, T., . . . O'Brien, C. P. (2014). An international consensus for assessing internet gaming disorder using the new DSM-5 approach. *Addiction, 109*(9), 1399-1406. doi: 10.1111/add.12457

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. doi: 10.3109/10601333.2014.962748

Pontes, H. M., & Griffiths, M. D. (2015). Measuring DSM-5 internet gaming disorder: Development and validation of a short psychometric scale. *Computers in Human Behavior, 45*, 137-143. doi: 10.1016/j.chb.2014.12.006

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. doi: 10.1371/journal.pone.0110137

Reeve, B. B., & Fayers, P. (2005). Applying item response theory modelling evaluating questionnaire item and scale properties. In P. Fayers & R. Hays (Eds.), *Assessing Quality of Life in Clinical Trials: Methods and Practice* (pp. 55-74). New York: Oxford University Press.

Rehbein, F., Kliem, S., Baier, D., Mößle, T., & Petry, N. M. (2015). Prevalence of Internet Gaming Disorder in German adolescents: diagnostic contribution of the nine DSM-5 criteria in a statewide representative sample. *Addiction, 110*(5), 842-851. doi: 10.1111/add.12849

Starcevic, V., Berle, D., Porter, G., & Fenech, P. (2011). Problem video game use and dimensions of psychopathology. *International Journal of Mental Health and Addiction, 9*(3), 248-256. doi: 10.1007/s11469-010-9282-5

van Rooij, A. J., & Prause, N. (2014). A critical review of "Internet addiction" criteria with suggestions for the future. *Journal of Behavioral Addictions, 3*(4), 203-213. doi: 10.1556/JBA.3.2014.4.1

Zanetta Dauriat, F., Zermatten, A., Billieux, J., Thorens, G., Bondolfi, G., Zullino, D., & Khazaal, Y. (2011). Motivations to play specifically predict excessive involvement in massively multiplayer online role-playing games: Evidence from an online survey. *European Addiction Research, 17*(4), 185-189. doi: 10.1159/000326070