Performance of the DSM-5-based criteria for Internet addiction: A factor analytical examination of three samples

BETTINA BESSER*, LOTTA LOERBROKS, GALLUS BISCHOF, ANJA BISCHOF and HANS-JÜRGEN RUMPF

Department of Psychiatry and Psychotherapy, University of Luebeck, Luebeck, Germany

(Received: June 13, 2018; revised manuscript received: March 25, 2019; accepted: April 7, 2019)

INTRODUCTION

Because of the large body of studies on video game addiction, the American Psychiatric Association (APA) included Internet Gaming Disorder (IGD) in the research appendix of the fifth edition of the Diagnostic and Statistical Manual for Mental Disorders (DSM-5; APA, 2013; Petry et al., 2014; Petry & O’Brien, 2013) and confined this potential diagnosis to online and offline game players. It is the first attempt to standardize the diagnostic criteria for this relatively new disorder, creating a framework for comparison of research across the board. The World Health Organization (WHO) decided to suggest Gaming Disorder as a condition in the 11th revision of the International Classification of Diseases. This decision is based on clinical evidence and public health needs with the demand of ensuring treatment and preventive measures (Rumpf et al., 2018). Although these approaches are reasonable, given the existing research, other Internet activities besides gaming could also lead to similar addictive behavioral patterns (Kuss & Griffiths, 2011; Pontes, 2017; Rumpf et al., 2014; van Rooij, Schoenmakers, van de Eijnden, & van de Mheen, 2010). In fact, the DSM-5 notes explicitly that the proposed criteria require greater study both in the context of Internet gaming as well as more global Internet use (APA, 2013). This study focuses on the broader concept of Internet addiction (IA) including gaming, social networks site use, compulsive buying, pornography use, or other Internet activities. Therefore, instruments for measurement of IGD have been adjusted using the term “Internet applications” rather than “Internet gaming” to measure IA instead of IGD.

Meanwhile, there have been studies that analyzed the performance of the DSM-5 criteria for IGD (Ko et al., 2014; Rehbein, Kliem, Baier, Mossle, & Petry, 2015) and a number of studies have developed questionnaires to measure IGD according to DSM-5 (Kiraly et al., 2017; Pontes & Griffiths, 2015; Pontes, Kiraly, Demetrovics, & Griffiths, 2014; Pontes, Stavropoulos, & Griffiths, 2017). One of these found a single-factor structure using exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) and another one found six factors (salience, mood modification, tolerance, withdrawal, conflict, and relapse) using CFA (Pontes et al., 2014). All these studies have used questionnaires, one is based on a representative sample of adolescents (Rehbein et al., 2015); the majority has used media solicited or online samples. To sum up, there are quite

* Corresponding author: Bettina Besser; Department of Psychiatry and Psychotherapy, University of Luebeck, Ratzeburger Allee 160, D-23538 Luebeck, Germany; Phone: +49 451 500 98755; Fax: +49 451 500 98754; E-mail: bettina.besser@uksh.de

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium for non-commercial purposes, provided the original author and source are credited, a link to the CC License is provided, and changes – if any – are indicated.

ISSN 2062-5871 © 2019 The Author(s)
a few studies on the performance of the IGD criteria and no data based on clinical interviews.

Aim

This paper briefly summarizes data from three samples based on fully structured clinical interviews. The aim is to analyze the factorial structure of the DSM-5 IGD criteria with a broader focus on Internet activities instead of just Internet gaming and to explore the contribution of each criterion to the explained variance. Response characteristics of the participants to the DSM-5 criteria will be evaluated in form of endorsement rates via multinomial logistic regression analysis.

METHODS

Samples

Sample 1 is based on a large general population sample \( n = 15,023 \) (Meyer et al., 2015; Rumpf et al., 2014) of which participants with elevated levels of problem use of the Internet, as defined by a low- and high-sensitive cut-off of 21 or more points on the Compulsive Internet Use Scale (CIUS; Meerkerk, Van Den Eijnden, Vermulst, & Garretsen, 2009), were reinterviewed face-to-face with a comprehensive assessment \( n = 196 \). Recruitment of the subsample used for the analyses of this study is described in more detail by Zadora et al. (2016). Of the initial general population study, 685 participants scored 21 or more points in the CIUS and 307 of them agreed to attend future studies. In total, 196 participants were reinterviewed. According to a non-response analysis comparing the initial sample of 685 with 196 participants who were reinterviewed, the non-responders were more likely to have at least one parent who was born outside Germany \( (p \leq .001) \) and were more likely to have received less than 10 years of schooling \( (p = .012) \). No differences were found regarding gender, age, unemployment status, and CIUS scores.

Sample 2 comprises job seekers who were screened for excessive Internet use with the CIUS in job agencies \( n = 3,040 \). Screening positive cases defined as having 21 or more points in the CIUS or using the Internet for 4 hr or more per day using a 5-point Likert scale (range: 0–45) were reinterviewed by telephone. Of screening positive cases, 138 telephone interviews could be realized.

Sample 3 was recruited in two vocational schools by systematic screening using the CIUS \( n = 1,209 \). Again, at least 21 points in the CIUS or having 9 or more points in the items suggested by an international consensus group on assessing the DSM-5 criteria (Petry et al., 2014) served as threshold for further assessment in a telephone interview \( (n = 188) \).

A comparison of the three samples revealed differences in age, subjective status of health, and school years attended. The results are shown in Table 1.

Assessment of IA according to DSM-5

The 14-item CIUS was used as a screening instrument covering the following five core criteria: salience, withdrawal, loss of control, conflict, and coping with unpleasant mood. The items are represented by a 5-point Likert-type scale from “never” to “very often.” Cronbach’s \( \alpha \) ranges from .88 to .90 (Meerkerk et al., 2009), suggesting good validity and reliability, and a stable one-factor solution was found across time and different samples.

In all three samples, IA was assessed using a fully structured interview based on the principles of the International Composite Diagnostic Interview (CIDI; Robins, Wing, & Wittchen, 1988; Wittchen, 1994) covering the proposed nine DSM-5 criteria. The term “Internet activities” replaced “gaming” to apply to all probable forms of Internet use. All nine DSM-5 criteria were assessed by a total of 27 questions, covering each criterion by one to four questions. Items are based on the structure and term of CIDI sections for substance-related disorders and gambling disorder adapted to IA. The computer-based interview automatically provides the number of fulfilled DSM-5 criteria. Lifetime as well as past-year symptoms were assessed. As suggested by DSM-5, participants who fulfilled at least five of the nine diagnostic criteria were categorized as fulfilling the diagnosis of IA. For this study, lifetime diagnoses are analyzed. The lifetime diagnosis refers to the same criteria as the past-year diagnosis, except that the reference period in which these occur refers to the previous life instead of the past 12 months. It implies that participants affirmed the inquired criteria at some point in their lives. We used no clustering of criteria in a specific time period except for the past 12 months. An ongoing study analyzed data of a sample of students of vocational schools and found that the diagnostic interview shows excellent reliability for past-year IA (Yule’s \( Y = .84 \)) as well as lifetime diagnosis (Yule’s \( Y = .86 \); Brandt et al., 2018).

Statistical analysis

For all three samples, separate factor analyses were performed. As a structure-detecting method, EFA can be

| Table 1. Comparison of three samples |
|--------------------------------------|
| Sample 1 \( (n = 196) \) | Sample 2 \( (n = 138) \) | Sample 3 \( (n = 188) \) | \( p \) |
|--------------------------------------|
| **Female sex (%)** | 48.2 | 43.8 | 51.4 | .407 |
| **Age [mean (SD)]** | 28.6 (12.9) | 26.4 (9.1) | 19.5 (3.2) | <.001 |
| **Subjective status of health [mean (SD)]** | 3.3 (0.9) | 2.8 (1.1) | 2.6 (0.9) | <.001 |
| **More than 9 years of schooling (%)** | 70.6 | 62.5 | 79.0 | .005 |
| **CIUS sum score [mean (SD)]** | 25.6 (4.4) | 26.4 (9.3) | 26.5 (7.1) | .091 |

**Note. SD**: standard deviation; CIUS: Compulsive Internet Use Scale.
used to examine the underlying construct of the DSM-5 criteria. The inspection of diagnostic criteria using factor analysis techniques has been found to be a valuable approach (Boelen, Spuij, & Lenferink, 2019; Hartwell & Ray, 2018; McSweeney, Koch, Saules, & Jefferson, 2016). The binary data were primarily prepared using an underlying-variable-approach (UVA) by calculating polychor correlations of each item couple. These correlations are considered to be specific correlation coefficients for ordinal data that extinguish every item to exhibit a latent variable whose range of values has been fragmented into the intervals of the ordinal variable. The UVA attempts to estimate correlations between these latent variables via tetracor correlations of the theoretical variables using the maximum likelihood method based on the assumption that the bivariate item couples are normally distributed. Factor analysis was performed on the basis of the resulting correlation matrix via principle axis method. Furthermore, endorsement rates of the criteria were analyzed descriptively for each sample and compared via multinomial logistic regression analysis.

Ethics

The study procedures were carried out in accordance with the Declaration of Helsinki. The studies were approved by the ethics committee of the University of Luebeck. All subjects were informed about the study and all provided informed consent. Parental consent was sought for those aged less than 18 years.

RESULTS

Kaiser–Meyer–Olkin (KMO) Test and sphericity

KMO Test of sampling adequacy shows a value of 0.65 in Sample 1, 0.56 in the Sample 2, and 0.68 in Sample 3. The Bartlett’s test of sphericity shows a significant result in all three samples (p ≥ .001).

Factor structure

Separately for all three samples, EFA was performed. Based on decisions on grounds of the eigenvalues, a single factor explains the data best in Samples 1 and 2. The single factor of Sample 1 has an eigenvalue of 3.50 and explains 38.86% of the total variance. The single factor of Sample 2 has an eigenvalue of 4.66 and explains 51.81% of the total variance. With eigenvalues of 1.19 and 1.03, second factors of both samples explain little more variance than they contribute and should be neglected in favor of the single-factor solution.

Contrary to expectations, the eigenvalues indicate a two-factor solution for the data of Sample 3. The first factor has an eigenvalue of 3.16 and explains 35.13% of the total variance. With an eigenvalue of 1.36, the second factor explains considerably more variance than it contributes. Therefore, a two-factor solution should be considered, which would explain 50.19% of the total variance. However, factor loadings of the two-factor solution show that only one item (escape from a negative mood) is associated with the second factor. In favor of better interpretability, a single-factor solution was chosen for Sample 3 after all. Eigenvalues of the factors are shown in Table 2. Factor loadings, communalities, and explained variance of all samples relating to a single-factor solution are displayed in Table 3.

Endorsement rates

Response characteristics of the participants to the DSM-5 criteria for IGD have been evaluated in all three samples and are shown in form of endorsement rates (Figure 1). The samples were compared to each other in relation to fulfilling the different criteria via multinomial logistic regression analysis. Sex and age as possible confounders were added as covariables. Sample 3 is significantly younger than Samples 1 and 2 (p < .001). In addition, regardless of age and sex as possible confounding variables, the analysis indicates that the endorsement rate of criterion no. 5 (loss of interests) is significantly lower in the Sample 3 in comparison to other samples.

With endorsement rates of 76.5%, 77.5%, and even 85.5%, the DSM-5 criterion no. 8 (using the Internet to escape or release a negative mood) scores highest in all three samples.

Table 2. Eigenvalues of the factor analysis (principal axis method) for Samples 1, 2, and 3

| Factor | Sample 1 (n = 196) | Sample 2 (n = 138) | Sample 3 (n = 188) |
|--------|-------------------|-------------------|-------------------|
|        | Variance (%)      | Cumulative (%)    | Variance (%)      | Cumulative (%)    | Variance (%)      | Cumulative (%)    |
| 1      | 3.50              | 38.86             | 4.66              | 51.81             | 51.81             | 3.16              | 35.13             | 35.13             |
| 2      | 1.19              | 13.23             | 1.03              | 11.38             | 63.19             | 1.36              | 15.06             | 50.19             |
| 3      | 1.16              | 12.91             | 0.93              | 10.32             | 73.51             | 0.98              | 10.91             | 60.10             |
| 4      | 0.80              | 8.83              | 0.62              | 6.86              | 80.36             | 0.92              | 10.19             | 71.29             |
| 5      | 0.70              | 7.79              | 0.62              | 6.83              | 87.19             | 0.82              | 9.08              | 80.37             |
| 6      | 0.54              | 5.96              | 0.52              | 5.76              | 92.95             | 0.69              | 7.70              | 88.07             |
| 7      | 0.51              | 5.61              | 0.42              | 4.63              | 97.58             | 0.63              | 7.00              | 95.07             |
| 8      | 0.44              | 1.87              | 0.17              | 1.85              | 99.42             | 0.29              | 3.25              | 98.31             |
| 9      | 0.18              | 1.94              | 0.05              | 0.58              | 100.00            | 0.15              | 1.69              | 100.00            |
KMO Test shows moderate sampling adequacy for Sample 1, a poor adequacy for Sample 2 (but with a value higher than 0.5 still sufficient enough for the analysis), and a good adequacy for Sample 3.

In Sample 1, four items (2, 3, 4, and 8) exhibit communalities under 0.4, Item 9 shows a communality under 0.3. Mundform, Shaw, and Ke (2005) showed that more the variables are being measured per factor and the higher the communalities are, the smaller the sample size can be without distorting the factor solution. A sample size of 200 participants and a minimum of five items per factor were defined as good for low communalities (0.2–0.4) and would apply to Sample 1 (n = 196). Sample 2 is smaller (n = 138) but holds all communalities higher than 0.5. Mundform et al. defined a sample size of 130 and a minimum of eight items per factor as excellent for middle communalities (0.4–0.6). Sample 3 shows one communality larger than 0.6 (Item 8), four larger than 0.5 (Items 2, 3, 5, and 6), one larger than 0.4 (Item 7), one larger than 0.3 (Item 1), and two above 0.2 (Items 4 and 9). The considerably low communalities of some items show that the representation of the factor solution is below average for those items. It could indicate a different variance of the concerning items in comparison to the others, which could be explained by a different answering pattern throughout Sample 3. Overall, the low communalities of Sample 3 cannot be compensated by the number of total items or the sample size (n = 188). Communalities are an essential criterion for evaluating the quality of a factor solution. On account of the partially poor communalities, the results of the factor analysis on Sample 3 might show random loading patterns. Therefore, the results are interpreted in an exploratory way and should be replicated in following studies.

Table 3. Initial communalities and factor loadings for Samples 1, 2, and 3

| Criteria | Sample 1 (n = 196) | Sample 2 (n = 138) | Sample 3 (n = 188) |
|----------|---------------------|---------------------|---------------------|
|          | Initial communalities | Initial communalities | Initial communalities |
|          | Factor loadings     | Factor loadings     | Factor loadings     |
| 1        | 0.44 0.69           | 0.75 0.87           | 0.35 0.54           |
| 2        | 0.32 0.53           | 0.56 0.57           | 0.56 0.58           |
| 3        | 0.37 0.52           | 0.58 0.50           | 0.52 0.71           |
| 4        | 0.34 0.46           | 0.50 0.70           | 0.24 0.43           |
| 5        | 0.68 0.82           | 0.80 0.77           | 0.53 0.52           |
| 6        | 0.55 0.58           | 0.67 0.66           | 0.58 0.58           |
| 7        | 0.44 0.44           | 0.79 0.62           | 0.46 0.50           |
| 8        | 0.40 0.55           | 0.88 0.78           | 0.60 0.32           |
| 9        | 0.28 0.41           | 0.69 0.63           | 0.27 0.45           |
| Explained variance | 6.00 38.86 | 1.00 51.81 | 3.00 35.13 |

Note. 1: preoccupation; 2: withdrawal; 3: tolerance; 4: unsuccessful attempts to stop or reduce; 5: loss of interest in other hobbies; 6: excessive use despite problems; 7: deception; 8: escape from negative mood; 9: jeopardized relationships or job opportunities.

Figure 1. Endorsement rates for Samples 1, 2, and 3 relating to DSM-5 criteria. 1: preoccupation; 2: withdrawal; 3: tolerance; 4: unsuccessful attempts to stop or reduce; 5: loss of interest in other hobbies; 6: excessive use despite problems; 7: deception; 8: escape from negative mood; 9: jeopardized relationships or job opportunities. *p = .027. **p = .013
The findings of the factor analyses clearly favor a single-factor structure of the DSM-5 criteria adapted for IA in both of the general population samples (Samples 1 and 2). These results replicate findings of Pontes and Griffiths (2015) as well as Sarda, Begue, Bry, and Gentile (2016), who both found single-factor structures using EFA and CFA on convenience samples recruited via online gaming forums or Facebook. In our samples, this factor explains 38.86% of the total variance in Sample 1, 51.81% in Sample 2, and 35.13% in Sample 3. The total variance that can be explained by the factor “IA” in all three samples is not optimal, which might be a result of the fact that the samples are rather homogenous.

The eigenvalues of the factors for the analysis of Sample 3 (student sample) indicate a two-factor structure. This might be explained by the significantly lower age of the participants: the performance of the criteria might be different in younger age groups. The interpretation of the two-factor solution is difficult. Only one item (escape from a negative mood) can be assigned to the second factor. One way of interpreting this result is explained by Van Rooij and Prause (2014) who critically discussed the item as a criterion for IA and even suggested the Internet to be an affective coping tool for mood modification that does not necessarily have a negative effect. Considering the young age of the participants of this sample, the use of the Internet for mood modification might be an age-sensitive effect that should be considered in further studies. Simultaneously, the criterion seems to be less associated with pathological Internet use in younger age groups. It can be suggested that escaping from a negative mood using the Internet is very common among young people.

The analysis of endorsement rates shows that differences with respect to age are likely, since the student sample is significantly younger than both other general population-based samples. Despite controlling the effect of age and sex by including them as covariates, a significantly lower endorsement rate of criterion no. 5 (loss of interests) can be found in Sample 3. This could indicate a specific weakness of the sample. Overall, all three samples show the highest endorsement rate for criterion no. 8 (escape from a negative mood). This indicates the criterion to have low severity and could indicate an insufficient discrimination between problematic and non-problematic Internet use. This is in line with data from Rehbein et al. (2015) who evaluated how endorsement of single criteria correspond to meeting five or more criteria of IA in a sample of ninth graders. The authors state that the criterion was endorsed at high rates, but was weak in predicting IA.

Restrictively, all three samples that were used in this study have been preselected through screening for problematic Internet use. It can be assumed that the endorsement rates would turn out to be lower in non-preselected samples in general. Therefore, future studies should ascertain the high endorsement rate of criterion no. 8 that has been found in this study to be a ceiling effect using non-preselected samples.

To see whether the clinical interviews measured the same in the three different samples, it would have been beneficial to test for measurement invariance via CFA. Unfortunately, it is not recommended to perform CFA on the same sample as EFA. To randomly split the samples in half to perform both analyses was not possible due to low sample sizes. Verifying the detected factor structure via CFA was not possible for the same reason.

CONCLUSIONS

In general, the data confirm the usefulness of the DSM-5 approach in terms of factorial validity in adults. EFAs indicate the measurement of IA using DSM-5 criteria to be represented one dimensionally, although there is evidence for other contributing parameters that have not been identified yet. The criteria are strongly associated with each other. This result provides valuable guidance in understanding the underlying nosology as part of the general study of the DSM-5 criteria. Overall, the highest endorsement rate was found for the “escape from a negative mood” criterion, which could be a ceiling effect. In this case, the criterion would not be suitable for diagnostic approaches and should be discarded.

The study provides indications of possible differences in the appearance of this disorder in younger age groups. This is reflected particularly in the criterion mood modification, which occurs frequently in the younger cohort and simultaneously is less associated with pathological behavior. Provided that other studies find similar results and may also classify this criterion as inadequate for diagnostic purposes in younger cohorts, the diagnostic criteria should be tailored to match the specific characteristics of younger age groups.

To examine differences between the nine criteria as well as other contributing parameters and to verify the single-factor approach, further studies are needed, in particular with respect to the performance of the criteria in different age groups as well as in non-preselected samples. This would also facilitate the identification of possible ceiling effects. To analyze and understand data on a different level, it would be interesting to focus on specific item-level characteristics by performing item-response theory for both lifetime and past-year data.

Funding sources: This work was supported by German Federal States and German Federal Ministry of Health.

Authors’ contribution: BB contributed to data gathering, statistical analysis, interpretation of findings, and preparation of manuscript draft. LL contributed to statistical analysis and design. GB contributed to study concept, statistical analysis, and interpretation of findings. AB contributed to study concept, design, and data gathering. H-JR contributed to study concept, design, obtained funding, and interpretation of findings. All authors had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of interest: All authors declare no conflict of interest with regard to this article.
van Rooij, A. J., Schoenmakers, T. M., van de Eijnden, R., & van de Mheen, D. (2010). Compulsive Internet use: The role of online gaming and other Internet applications. *Journal of Adolescent Health, 47*(1), 51–57. doi:10.1016/j.jadohealth.2009.12.021

Wittchen, H.-U. (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research, 28*(1), 57–84. doi:10.1016/0022-3956(94)90036-1

Zadra, S., Bischof, G., Besser, B., Bischof, A., Meyer, C., John, U., & Rumpf, H. J. (2016). The association between Internet addiction and personality disorders in a general population-based sample. *Journal of Behavioral Addictions, 5*(4), 691–699. doi:10.1556/2006.5.2016.086