

A clinical evaluation of the DSM-5 criteria for Internet Gaming Disorder and a pilot study on their applicability to further Internet-related disorders

KAI W. MÜLLER^{1*}, MANFRED E. BEUTEL², MICHAEL DREIER¹ and KLAUS WÖFLING¹

¹Outpatient Clinic for Behavioral Addiction, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center, Johannes Gutenberg University Mainz, Mainz, Germany

²Department of Psychosomatic Medicine and Psychotherapy, University Medical Centre, Johannes Gutenberg University Mainz, Mainz, Germany

(Received: March 23, 2018; revised manuscript received: October 30, 2018; second revised manuscript received: November 28, 2018; accepted: December 2, 2018)

Background and aims: Internet Gaming Disorder (IGD) and other Internet-related disorders (IRDs) have become growing health concerns in our today's lives. Based on defined diagnostic criteria, IGD has been recognized as a condition for further research in the DSM-5; however, other IRDs have been excluded. Since the release of the DSM-5, representativeness and appropriateness of the nine diagnostic criteria have been debated. Although some first evidence has been published to evaluate these criteria, our knowledge is still limited. Thus, the purpose of this study was to provide data on the clinical validity of the DSM-5 criteria for IGD and other types of IRD. We were also interested in examining the additional diagnostic validity of craving that is currently not being considered in the DSM-5. *Methods:* Analyses on a sample of $n = 166$ treatment seekers for IRDs were performed. The clinician's diagnosis was used as a main reference for determining the DSM criteria's diagnostic performance. Secondary criteria (depression and anxiety) were defined as indicators for the construct validity. *Results:* The overall diagnostic accuracy ranged between 76.6% for deceiving and 92% for loss of control and craving. Considerable differences occurred in the degree of sensitivity and specificity between the single criteria. No particular differences were found for the applicability of the criteria to other forms of IRDs. *Discussion and conclusions:* Our results confirm the validity of the DSM criteria. However, the diagnostic utility of the criterion escaping aversive moods is critically discussed. Considering craving as an additional diagnostic indicator might be recommendable.

Keywords: clinical validity, diagnostic accuracy, DSM-5, Internet addiction, Internet Gaming Disorder, Internet-related disorders

INTRODUCTION

The 2013 release of the diagnostic criteria for Internet Gaming Disorder (IGD) has been recognized as a major step toward a better understanding of this new disorder (Griffiths et al., 2016; Petry et al., 2014). When defining a diagnostic framework for IGD in the fifth edition of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association [APA], 2013), the APA primarily relied on criteria known from Gambling Disorder. These shared diagnostic criteria include preoccupation, loss of control, tolerance, withdrawal, deceiving about the behavior, and jeopardizing important relationships or perspectives by the behavior that is continued despite negative consequences. Several studies have pointed out neurobiological and neuropsychological parallels between IGD and Gambling Disorder, and given the growing consensus that both conditions may be considered as behavioral addictions (e.g., Fauth-Bühler & Mann, 2017), the similarity between the diagnostic criteria of both disorders is comprehensible.

Yet, although specifying a diagnostic concept of IGD has been appreciated (Petry et al., 2014), researchers have also pointed out that there is space left for amendments (Griffiths et al., 2016). Main points of criticism regarded vague definition of some of these criteria, for instance, in the case of deceiving others in order to conceal the behavior or using the game for the regulation of emotions. Others claimed that some of the criteria are inappropriate for a modern lifestyle activity that gaming has become during the past decades and that individuals merely joining a modern recreational activity are running the risk of being pathologized (Aarseth et al., 2017).

While empirical findings have contradicted the latter perspective by demonstrating that IGD is accompanied by

* Corresponding author: Dr. Kai W. Müller; Outpatient Clinic for Behavioral Addiction, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center, Johannes Gutenberg University Mainz, Untere Zahlbacher Straße 8, 55131 Mainz, Germany; Phone: +49 6131 174039; Fax: +49 6131 4287; E-mail: muellka@uni-mainz.de

This is an open-access article distributed under the terms of the [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/), 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.

severe psychopathological symptoms and a loss of psychosocial functioning (e.g., Gentile et al., 2011; Ko et al., 2014; Müller et al., 2017; Rehbein, Kliem, Baier, Mößle, & Petry, 2015), the question remains if the DSM criteria are solid enough to capture clinical reality (e.g., Billieux, Schimmenti, Khazaal, Maurage, & Heeren, 2015). Several authors have proposed to either emphasize single criteria in order to achieve higher diagnostic accuracy (DA) or to add further criteria (e.g., Ko et al., 2014; Müller, Glaesmer, Brähler, Wölfling, & Beutel, 2014). In particular, the concept of craving has repeatedly been proposed as an additional criterion. Craving is commonly perceived as a motivational state initiating drug-seeking behavior. Research from substance use disorders has indicated that craving correlates with the disorder's severity and that it adds diagnostic validity in predicting alcohol dependence (Keyes, Krueger, Grant, & Hasin, 2011). There is first evidence that this may also hold true for IGD. For instance, Ko et al. (2014) published an early evaluation study of the DSM-5 criteria. They clinically assessed the presence of the nine criteria in young adults with IGD, individuals formerly displaying IGD, and healthy controls. In addition, they included two further criteria: irritability and craving. While irritability turned out not to have additional diagnostic value, craving had a DA of 89% and therefore was one of the best performing criteria in assessing IGD.

Subsequently, further approaches to examine the validity of the DSM criteria have been published. However, unlike Ko et al. (2014), these studies did not rely on a clinical evaluation of the IGD status. For example, Király, Slezcka, et al. (2017) conducted an online survey based on a sample of $n = 4,887$ gamers. They statistically examined the performance of the DSM criteria and found them generally satisfying. Interestingly, they also demonstrated that the criteria were differentially associated with the severity of IGD. Similarly, Rehbein et al. (2015) presented a population-based statistical examination of the criteria in $n = 11,003$ adolescents. They found the criteria escape negative mood states and preoccupation to be poorly related to IGD, whereas loss of interest and tolerance were the best predictors for IGD. Of particular interest is that the authors showed that specific combinations of criteria were enhancing the DA. Finally, Lemmens, Valkenburg, and Gentile (2015) reported findings on the DA of the DSM criteria from an epidemiological survey. By defining external criteria like loneliness, life satisfaction, and self-esteem, they demonstrated high sensitivity for the criteria withdrawal and continued use and high specificity for loss of control and jeopardizing important relationships.

Despite these promising approaches, our knowledge on the DA of the DSM criteria and their applicability in clinical contexts is still limited. Crucial questions regard the clinical validity of these criteria and potential beneficial effects of additional criteria.

An unresolved question concerns the applicability of these DSM criteria to other forms of Internet-related disorders (IRD; also called Internet addiction). There is an ongoing debate if addictive behaviors also occur in relation to online activities other than gaming (e.g., online-pornography and social networking sites). Some studies have dealt with specific subtypes of Internet addiction

(e.g., Love, Laier, Brand, Hatch, & Hajela, 2015; Müller et al., 2016; Pawlikowski, Nader, Burger, Stieger, & Brand, 2014; Rosenkranz, Müller, Dreier, Beutel, & Wölfling, 2017). However, the DSM-5 excludes these subtypes because of lacking evidence. This poses the question, if the DSM criteria originally defined for assessing IGD can be transferred to other types of IRDs.

With regard to the aforementioned aspects, we intend to examine the diagnostic performance of the DSM-5 criteria among individuals presenting in a clinical context because of excessive use of the internet. Our central research questions are as follows: (a) Are there differences between the diagnostic criteria proposed in the DSM-5 in assessing IGDs and other types of IRDs? Based on previous evaluation studies, we expect that the criteria will differ regarding their diagnostic properties (sensitivity, specificity, positive predictive rate, and negative predictive rate). (b) The DSM criteria have been defined for IGD but are currently used also to assess different subtypes of IRD. Thus, we wanted to provide first comparative data on their applicability to other types of IRD. The data presented here are part of a larger project "Diagnostics in Internet-Related Disorders" that aims to evaluate different diagnostic approaches in IRD and to identify conditions influencing DA.

We defined as a third research question to examine the usefulness of adding craving as a further diagnostic criterion. As depicted above, there is preliminary evidence for considering craving in the diagnosis of addictive behaviors and we were interested in assessing this criterion's diagnostic performance.

Finally, we wanted to address the question of potential discrepancies between externally assessed DSM criteria (e.g., by a clinician) and data from the patients' self-report. To that purpose, we compared the evaluation of the DSM criteria by a clinician with the standardized answers of the patients provided in a self-assessment tool (AICA-S; Wölfling, Beutel, & Müller, 2016).

METHODS

Pilot study

In a first step, we conducted a pilot study consisting of $n = 20$ patients classified with IRD (including IGD). With respect to the nine DSM criteria, any therapeutic material available (e.g., therapist's notes, clinical reports, and video recordings) was examined by trained psychologists. The goal was to identify indicators for each of the DSM criteria, test for their applicability, and to establish a blueprint for the subsequent analyses of new cases (examples for the indicators that were identified in this stage can be derived from Supplement 1).

Participants and procedure

The sampling procedure was based on a cohort of consecutive treatment seekers aged 16 years or older presenting between 2013 and 2015 ($n = 220$) in a specialized outpatient clinic. The reason for entering there was an overuse of computer games or Internet-based contents and therefore

suspected IGD or IRD. Any treatment seeker was asked to give written informed consent for recording, analyzing, and publishing their data for scientific purpose and the study was approved by the local ethics committee. A total of $n = 54$ individuals did not agree leading to a total sample of $n = 166$. The treatment seekers then underwent the standard procedure for initial diagnostics: they were asked to fill in questionnaires (see next paragraph) and afterward attended a diagnostic interview by a trained clinician. In the aftermath, the clinician made a diagnostic evaluation on the condition of IGD or IRD of the treatment seeker (criteria for IRD met vs. not met) that was based on the exploration of the IGD/IRD symptoms as assessed in the diagnostic interview (for details, see section “Diagnostic interview”). The clinician’s diagnosis of the IGD/IRD status was considered as a gold standard for the subsequent analyses of this study. A total of $n = 100$ treatment seekers presenting for supposed IGD and $n = 66$ presenting for other types of IRD. Supplement 2 illustrates the enrollment procedure and the subtype of IRD assessed.

In 2017, the data of these individuals were re-evaluated by another trained professional who was unaware of the clinician’s diagnosis (blinded evaluation). Again, these qualitative analyses were based on different sources, mainly on the clinical report, the clinician’s notes from the initial interview, and – if available – video recordings of the diagnostic interview. Based on this material, a standardized rating of the DSM criteria for IGD that was based on the findings of the pilot study was performed. The rating scale was related to the rationale of the Structured Clinical Interview for DSM-IV and contained the following categories: 0 = criterion not met; 1 = criterion met but not in a clinical sense; 2 = criterion fully met; and 9 = not enough information available. The last category was used when the clinical material available did not contain any reliable information on the specific criterion under investigation.

Questionnaires

Scale for the Assessment of Internet and Computer game Addiction (AICA-S). This self-report is based on the criteria loss of control, tolerance, withdrawal, continued use despite negative consequences, loss of interest, emotion regulation, jeopardizing perspectives, and craving to classify Internet usage according to its clinical relevance (Wölfling et al., 2016). A total of 13 items (ranging from 0 = *not at all/never* to 4 = *very often/very strongly*) and additional six items in a dichotomous format are weighted and summed for the total score. The psychometric properties, criterion and constructs validity, as well as clinical validity have been previously demonstrated (Kuss, Shorter, van Rooij, Griffiths, & Schoenmakers, 2013; Müller, Beutel, & Wölfling, 2014).

Patient Health Questionnaire, Depression Module (PHQ-9). This nine-item depression scale of the Patient Health Questionnaire (PHQ) is a self-report for assessing major depression and its severity. The items correspond to the nine diagnostic criteria for major depression disorder in DSM-IV and are rated on a 4-point rating scale ranging from 0 (*not at all*) to 3 (*nearly every day*) covering a period of the past 2 weeks. Using a cutoff score of 10, the PHQ-9 reached

a sensitivity of 89% and a specificity of 88% (Kroenke, Spitzer, & Williams, 2001).

Generalized Anxiety Disorder Scale (GAD-7). The GAD-7 is a self-report instrument for screening for general anxiety disorder and assessing its severity. The items are rated on a 4-point rating scale ranging from 0 (*not at all*) to 3 (*nearly every day*) for a period of the past 2 weeks. Using a cutoff score of 10, the GAD-7 demonstrated a sensitivity of 89% and specificity of 82% (Spitzer, Kroenke, Williams, & Lowe, 2006).

Mini-Social Phobia Inventory (Mini-SPIN). The Mini-SPIN is a self-administered screening instrument for social anxiety disorder. It is composed of three items derived from the long version of the Social Phobia Inventory that reached the highest sensitivity and specificity for the diagnosis of social anxiety disorders. The items are rated on a 5-point rating scale ranging from 0 (*not at all*) to 4 (*extremely*). Using a cutoff score of 6, the Mini-SPIN attained a sensitivity of 88.7%, specificity of 90.0%, and diagnostic efficiency of 89.9% (Weeks, Spokas, & Heimberg, 2007).

Global assessment of functioning (GAF). This expert rating based on axis V of the DSM-IV is covering psychosocial impairment. Its score ranges from 0 (*inadequate information*) to 100 (*superior functioning*) and its validity and interrater reliability have been demonstrated (Söderberg, Tungström, & Armelius, 2005). In this study, we only made use of the overall GAF score.

Diagnostic interview. Each treatment seeker underwent a semi-structured clinical exploration of the IGD/IRD symptoms that also consisted of a standardized clinical assessment of IRD symptoms (Checklist for the AICA-C; Wölfling, Beutel, & Müller, 2012). This clinical checklist assesses six of the nine DSM criteria. Its clinically validated cut-off score allows for a clinical diagnosis of IGD/IRD and its DA has been evaluated before (Wölfling et al., 2012). Main contents of the exploration consisted of assessing negative consequences arising from the online behavior, effects on the individual’s level of functioning or resulting psychosocial impairment. As a reference period, the past 12 months were defined. This exploration was conducted by a trained professional and also addressed the current Internet usage habits, the contents and extent of using the Internet. The clinician’s diagnosis on the presence of IGD/IRD was based on his or her global impression on the symptoms reported by the treatment-seeking individual. In the interview, a particular emphasize was put on the degree of impairment arising from the online behavior and the level of psychosocial functioning. This exploration was the primary basis for setting the diagnosis of IGD/IRD.

Statistical analyses

The main reference for an external criterion indicating IGD/IRD was the clinician’s evaluation after the diagnostic interview. For determining the construct and criterion validity of the clinician’s evaluation on the IRD status group, comparisons and correlations with external criteria (e.g., PHQ-9 and GAF) were performed.

As primary indicators for the diagnostic performance of the DSM criteria sensitivity (rate of true positive cases) and specificity (rate of true negative cases) were determined.

Positive prediction rate (PPR; patients with a positive test who actually have the disease), negative prediction rate (patients with a negative test who do not have the disease), and the DA (ability of the test to discriminate between positive and negative cases) were calculated as well. In order to estimate associations between the clinically assessed DSM criteria and the criteria reported by the patient in a self-report measure for IRD (AICA-S), contingency analyses were performed. The contingency coefficient (CC) served as an indicator for the strength of relationships with a threshold of $CC \geq 0.30$ indicating strong effects.

Ethics

Every participant was informed about the study’s background and that participation was voluntarily and finally asked to give written informed consent. Thus, the procedure was in accordance with the Declaration of Helsinki and the local ethical commission approved the study.

RESULTS

Characteristics of the selected treatment seekers

Table 1 displays the demographic characteristics of the treatment seekers for IGD and other types of IRD.

According to the clinician’s evaluation based on the diagnostic interview, a total of $n = 66$ (66.0%) individuals presenting for an overuse of computer games were diagnosed with IGD (IGD patients); thus, in $n = 34$ (34.0%), the

diagnosis IGD was not met (IGD-exclude-group). The majority of the IGD-treatment seekers reported an overuse of online games ($n = 98$; 98%). Among those presenting for other IRD subtypes, $n = 48$ (72.7%) met the diagnostic criteria (IRD patients), which means that $n = 18$ (27.3%) were not diagnosed with IRD (IRD-exclude-group). Those IRD-treatment seekers primarily displayed an overuse of online pornography ($n = 14$, 21.2%), social networking sites ($n = 13$, 19.7%), online streaming ($n = 6$, 9.1%), or other online contents ($n = 6$, 9.1%). Unspecified Internet use habits were assessed in 40.9% ($n = 27$) of the individuals, meaning that there was no clear preference for a specific online activity (e.g., streaming portals and social media were used excessively).

Validity of the clinician’s diagnosis

In order to evaluate the clinician’s assessment on the IGD/IRD status, we compared patients meeting criteria for IGD/IRD with those not being diagnosed with IGD/IRD. IGD/IRD patients ($M = 7.8$, $SD = 4.35$) spent significantly more hours online on an average week day compared to the IGD/IRD-exclude-group [$M = 6.2$, $SD = 3.81$, $t(146) = 2.18$, $p = .031$, $d = 0.16$]. Further significant differences were found for the AICA-S score. Here, IGD/IRD patients ($M = 12.7$, $SD = 5.91$) had significantly higher scores than the IGD/IRD-exclude-group [$M = 6.06$, $SD = 3.10$, $t(149) = 9.18$, $p = .001$, $d = 0.99$].

Analyses of the PHQ-9 revealed that IGD/IRD patients ($M = 9.9$, $SD = 6.06$) had significantly higher depression scores than the IGD/IRD-exclude-group

Table 1. Comparison of the demographics of the treatment seekers presenting for Internet Gaming Disorder or other Internet-related disorders

Demographics	Reason for seeking treatment		Statistics
	(Online) gaming ($n = 100$)	Other Internet-based activities ($n = 66$)	
Gender (% , n)			
Male	91.0% (91)	77.3% (51)	$\chi^2(1) = 6.06$, $p = .014$
Female	9.0% (9)	22.7% (15)	
Age (M , SD)	23.8 (6.72)	32.0 (11.71)	$t(163) = 5.68$, $p = .001$
Level of education (% , n)			
Still at school	13.3% (13)	8.1% (5)	n.s.
<10th grade	42.9% (42)	40.3% (25)	
>10th grade	38.8% (38)	51.5% (32)	
No graduation	5.1% (5)	0.0% (0)	
Occupational status (% , n)			
Full/part time	25.0% (23)	43.3% (26)	n.s.
Unemployed	30.4% (28)	21.7% (13)	
College/university	35.9% (33)	25.0% (15)	
Other	8.7% (8)	10.0% (6)	
Living situation (% , n)			
Living alone	32.0% (32)	30.2% (19)	n.s.
Living with partner	13.0% (13)	27.0% (17)	
Living with parents	49.0% (49)	34.9% (22)	
Other	6.0% (6)	7.9% (5)	
Partnership (% , n)			
Yes	30.0% (30)	42.2% (27)	n.s.
No	70.0% (70)	57.8% (37)	

Note. Missing values for level of education, occupational status, living situation, and partnership ranged between 2 and 14 cases. Degrees of freedom are represented in brackets. M : mean; SD : standard deviation; χ^2 : chi-square; p : p value (level of significance); t : t -value; n.s.: not significant.

[$M = 6.0, SD = 5.90, t(441) = 5.69, p = .001, d = 0.64$]. The same was true for the GAD-7 score [IGD/IRD-patients: $M = 8.1, SD = 5.18$; IGD/IRD-exclude-group: $M = 4.1, SD = 4.07, t(97) = 4.94, p = .001, d = 0.60$] and the Mini-SPIN [IGD/IRD-patients: $M = 4.7, SD = 3.30$; IGD/IRD-exclude-group: $M = 3.2, SD = 3.50, t(142) = 2.50, p = .014, d = 0.48$]. Finally, the global GAF score of IGD/IRD patients ($M = 64.3, SD = 12.46$) was significantly decreased compared to the IGD/IRD-exclude-group [$M = 77.1, SD = 14.12, t(153) = 5.64, p = .001, d = 0.86$].

Diagnostic performance of the DSM-5 criteria

Table 2 gives an overview on the diagnostic performance of the DSM criteria for IGD including craving. The initial clinician’s diagnostic evaluation on the IGD/IRD status was used as the main reference. Following the procedure of the pilot study, the DSM criteria extracted from the qualitative analyses of the clinical material were transformed in a dichotomous format (criterion met vs. not met). The former categories of 0 (criterion not met) and 1 (criterion met but not in a clinical sense) were combined to 0 (criterion not met); cases with category 9 (not enough information

available) were not part of this analyses, but their frequency was documented [see column “n.a. (%)” of Table 2].

The criteria jeopardizing relationships or perspectives, mood regulation, tolerance, and continued use displayed the best sensitivity (>90%). The best specificity was found for craving, withdrawal, and loss of control (>90%). DA was good to acceptable for almost every criterion, except for the criteria escaping aversive moods and deceiving.

The results also show that there were particular differences in the assessability of the criteria. In almost three thirds of the cases, no diagnostic information could be found for deceiving. Similarly, mood regulation, withdrawal, and craving could be explored in only about half of the treatment seekers.

Validity of the DSM-5 criteria for the global construct of IRDs

In a second step, it was examined if there are differences in the DSM criteria depending upon the type of IRD. To that purpose, we only considered patients either being diagnosed with IGD or IRD by the clinician. Comparisons of the differences in the presence of the diagnostic criteria were conducted using chi-square tests (Table 3).

Table 2. Diagnostic parameters of the DSM criteria and craving

	Sensitivity	Specificity	DA	PPR	NPR	n.a. (%)
Craving	88.3	100.0	92.0	1.000	0.800	47.0
Loss of control	91.5	92.7	91.8	0.970	0.809	11.4
Withdrawal	78.4	97.3	87.8	0.967	0.818	55.4
Jeopardizing	96.8	62.9	87.6	0.875	0.880	22.3
Loss of interest	86.8	84.6	86.1	0.917	0.767	30.7
Tolerance	95.5	61.8	86.1	0.866	0.840	26.5
Continued use	94.0	64.1	85.6	0.870	0.806	16.3
Preoccupation	80.7	82.7	81.3	0.911	0.662	10.2
Mood regulation	96.6	33.3	79.7	0.800	0.778	52.4
Deceiving	82.8	66.7	76.6	0.800	0.706	60.3

Note. $N = 166$; Diagnostic criteria are sorted by diagnostic accuracy, specificity (true positive/true positive and false negative), sensitivity (true negative/true negative and false positive). DA: diagnostic accuracy (true positive and true negative/all); PPR: positive predictive rate (true positive/true positive and false positive); NPR: negative predictive rate (true negative/true negative and false negative); n.a.: not applicable (there was no sufficient information to decide upon this criterion).

Table 3. Comparison of the diagnostic criteria in patients with Internet Gaming Disorder (IGD) and patients with other Internet-related disorders (IRDs)

Presence of confirmed diagnostic criteria (%)	IGD patients ($n = 66$)	IRD patients ($n = 48$)	Statistics
Preoccupation	54 (87.1%)	42 (91.3%)	n.s.
Craving	30 (88.2%)	24 (88.9%)	n.s.
Tolerance	53 (94.6%)	35 (97.2%)	n.s.
Loss of control	55 (85.9%)	46 (97.9%)	$\chi^2 = 4.71, p = .030, \phi = .206$
Withdrawal	13 (61.9%)	19 (100%)	$\chi^2 = 9.05, p = .003, \phi = .476$
continued use	61 (95.3%)	36 (92.3%)	n.s.
Loss of interest	44 (88.0%)	28 (83.9%)	n.s.
Mood regulation	33 (94.3%)	26 (100%)	n.s.
Jeopardizing	58 (95.1%)	37 (97.4%)	n.s.
Deceiving	6 (60.0%)	13 (92.9%)	$\chi^2 = 3.82, p = .051, \phi = .399$

Note. IGD patients: treatment seekers meeting criteria for Internet Gaming Disorder according to the clinical interview; IRD patients: treatment seekers meeting criteria for other types of Internet-related disorders according to the clinical interview; n.s.: not significant; χ^2 : chi-square; p : p value (level of significance); ϕ : Phi (effect size).

The only significant differences in the diagnostic criteria between both groups concerned loss of control and withdrawal; for both criteria with medium effects sizes were found and both were more often present in patients with other forms of IRD. The differences of the criterion deceiving failed to reach statistical significance, yet yielded a medium effect size. Thus, given the small sample size, there was a significant trend.

Associations between self-reported and clinically assessed DSM criteria

Finally, we were interested in potential overlaps of the DSM criteria either assessed in the diagnostic interview or by self-report. As a self-report, we used AICA-S, a validated instrument for IRDs. AICA-S is based on eight of the nine DSM criteria (the criterion deceiving is not included) and additionally contains craving.

Except for mood regulation ($p = .052$), the contingency analyses yielded significant relationships for each of the remaining criteria. Strong effects were found for loss of interest ($CC = 0.578, p = .001$), loss of control ($CC = 0.530, p = .001$), withdrawal ($CC = 0.499, p = .001$), craving ($CC = 0.441, p = .001$), preoccupation ($CC = 0.432, p = .001$), continued use ($CC = 0.430, p = .001$), and tolerance ($CC = 0.375, p = .001$). For jeopardizing relationships or perspectives, a medium effect of $CC = 0.266 (p = .001)$ was found.

DISCUSSION

The release of the DSM criteria for IGD was appreciated as a big step toward a better understanding of this new clinical phenomenon. Yet, it also caused skepticism among scholars and provoked heated debates on the validity of those criteria (e.g., Aarseth et al., 2017; Griffiths et al., 2016; Kuss, Griffiths, & Pontes, 2017; Petry et al., 2014). Undoubtedly, there is a need for sound scientific research on the utility and validity of these criteria proposed. A reliable scientific basis is also needed to objectify some of these debates (e.g., Aarseth et al., 2017).

Thus, our aim was to add knowledge on the clinical validity of the DSM criteria for IGD from a clinical perspective. Based on a sample of treatment seekers, we intended to provide data on their DA and to give a first impression if they might be appropriate to also cover different types of IRDs. Finally, we aimed to broaden the current diagnostic perspective by adding craving as a potential additional diagnostic criterion.

To that purpose, we relied on the diagnostic evaluation of experienced clinicians as a gold standard for assessing IGD and IRD. Furthermore, we made use of independent qualitative analyses performed by trained professionals in order to confirm the presence of the single diagnostic criteria.

In order to evaluate the clinicians' diagnosis, we defined several constructs indicative for the criterion and construct validity. Compared to those individuals without IGD/IRD, we found that patients diagnosed with IGD or IRD spent significantly more time online, had significantly higher

scores in depression and anxiety, and displayed a significantly decreased level of functioning. For most of the indicators, medium to large effect sizes were found, except for time spent online with only a small effect size. Thus, these findings confirm the validity of the clinicians' evaluation on the IRD status of the individuals.

Our main results indicate that the DSM criteria are generally appropriate in clinically assessing IGD and IRD. However, some of these criteria showed a better diagnostic performance than others and likewise differences occurred in their appropriateness in either reliably identifying patients (sensitivity ranged between 78% and 97%) or healthy individuals (specificity varying between 33% and 100%). Although most of the criteria showed a DA higher than 80%, the criteria deceiving and escaping adverse moods failed to skip that threshold.

Generally, our results show quite some parallels but also differences compared to the findings provided by Ko et al. (2014). We also found considerable differences in the diagnostic performance of the criteria. Deceiving on the amount of using the Internet had the lowest DA. Ko et al. (2014) hypothesized that this might be due to social isolation of patients suffering from IGD. Indeed, this explanation might be accurate since different studies have shown that IGD and IRD are associated with loneliness and social isolation (Lemmens et al., 2015; Müller et al., 2017). Accordingly, about one third of the treatment seekers of this study reported living alone and more than a half of them were without partnership. Thus, social isolation might indeed impede the exploration of the criterion deceiving what should be considered in the clinical exploration of it.

In accordance with Ko et al. (2014), we also found the criterion craving to have a good diagnostic performance with a DA of 92%. This also matches findings from studies regarding substance use disorders, where craving has been discussed as a latent variable and as an indicator of severity (e.g., Hasin, Fenton, Beseler, Park, & Wall, 2012). Although more research is needed here, there seems reason to assume that craving can be considered as an additional diagnostic criterion for IGD and IRD.

We also found some results contradicting the study of Ko et al. (2014). For example, while Ko et al. (2014) reported quite low sensitivity for loss of interest (59%) and escaping negative moods (69%), we found both criteria to have excellent sensitivity. Similar differences concerned the specificity of the criterion jeopardizing where we found only a modest specificity of 63%, whereas Ko et al. (2014) reported a specificity of 97%. The differences found might be explainable by the different samples examined or cultural influences. In contrast to Ko et al. (2014), our analyses were exclusively based on a clinical sample. Moreover, individuals presenting to our service without being diagnosed with IGD or IRD were expressing a high intensity of media usage. Thus, our control group consisted of intense users and cannot be regarded representative for the average healthy user of modern technologies.

As mentioned before, some further studies provided information on the diagnostic validity of the DSM criteria. Rehbein et al. (2015) found loss of interest and tolerance to be the best predictors for IGD, whereas in our investigation, both criteria had an average scoring of DA. On the

contrary, Rehbein et al. (2015) found that preoccupation and escaping adverse mood were of lower validity. While this is basically confirmed by our results, the criterion preoccupation displayed a rather good PRP. Thus, individuals meeting this criterion have a heightened probability for IGD/IRD, which makes this criterion particularly relevant in confirming the diagnosis IGD/IRD in patients. It is important to note that we did not define preoccupation merely as being mentally preoccupied using the game or the Internet; rather we also considered the behavioral component of this criterion (e.g., gaming being a persistent and central activity in the person's life or having essential parts of one's life re-organized in order to allow for more time to use the game or the Internet). This might be essential, since simply thinking about a favorable activity should not be equated with a problem behavior. Of course, a further explanation regards general differences of the samples under examination. In our clinical sample, the individuals (regardless of being diagnosed with IRD or not) spent a significant amount of time being online, while in representative community samples, this is not the case. Thus, again, the specific characteristics of this clinical population play a role (cf. Griffiths, 2010; Király, Tóth, et al., 2017). Finally, the discrepancies between the study by Rehbein et al. (2015) and our investigations might also be due to the age difference of the samples examined. Rehbein et al. (2015) examined adolescents, whereas our sample encompassed late adolescents and adults. It might be that the diagnostic criteria are age-depending to some extent as has been hypothesized by others (Griffiths et al., 2016). Clearly, we need further studies on that possibility to guarantee high diagnostic standards also for children and adolescents. At present, only few attempts have been made here (e.g., Jo et al., 2017).

Also, Király, Slecza, et al. (2017) provided data on the diagnostic validity of the DSM criteria by examining a large sample of gamers. Preoccupation and escape adverse moods turned out to be the least apt criteria predicting IGD. Again, considering the behavioral component of the preoccupation criterion might be an explanation for that. Escaping adverse moods had also a poor diagnostic performance in our investigation, and this criterion has been conceptually criticized before (Griffiths et al., 2016). Therefore, it might be necessary to reconsider the nature of this criterion and to investigate potential additive associations between escaping adverse moods and other criteria (e.g., preoccupation or withdrawal).

An explorative aspect of this study regarded estimating the applicability of the IGD criteria to other types of IRD. As has been argued before, the term IRD encompasses different online activities that can run out of control. Moreover, it is reasonable to assume that different subtypes can be characterized by similar psychopathological features (Billieux, 2012; Brand, Young, & Laier, 2014; Müller, Beutel, et al., 2014). While we were only able to roughly address this important issue, our preliminary findings suggest that there are no striking differences in the applicability of the IGD criteria to other forms of IRD. Yet, it has to be noted critically that there were considerable demographic differences between our patients with IGD and IRD (e.g., higher age and more female patients in the IRD group). Of course,

these differences limit the comparability of both groups. Notwithstanding, more research is needed here.

Finally, we found quite strong relationships of the DSM criteria between self-reports and external assessment, which can be regarded as a further indicator of their criterion validity. Except for escaping aversive moods, we found strong associations between self-reports and clinical evaluation. This particularly holds for the examination of the DSM criteria clinical samples. It might also be an indicator of the criteria's validity as applied in epidemiological surveys; however, we shall not forget that the base rate of true positive cases in healthy samples is smaller than that in (specific) clinical populations; thus, more research is needed here (cf. Maráz, Király, & Demetrovics, 2015).

Unexpectedly, we found some of the criteria to be more challenging and demanding than others in terms of exploration. This becomes evident when comparing the missing rates for the single criteria. The lowest missing rates were obtained for preoccupation and loss of control, whereas much higher rates were found for deceiving and withdrawal. First, this variance has to be addressed as a limitation of this study. Second, it demonstrates that, in clinical practice, better guidelines are needed in order to explore the diagnostic criteria sufficiently. Such guidance could be best provided by structured clinical interviews assessing symptoms of IGD or IRD. In this respect, the analyses presented here provided a rich source for developing such a structured clinical interview (AICA-Structured Clinical Interview for Internet-Related Disorders; AICA-SCI:IRD). This has to be seen as a secondary outcome of the project "Diagnostics in Internet-Related Disorders."

A clear limitation of this study has to be seen in the small sample size that reduces the statistical possibilities for a further examination of the data. Moreover, as it is often the case in qualitative analyses, the procedure of evaluating the clinical data available was potentially prone to some biases. Although independent professionals were working on the material and even though a standardized framework for the criteria's examination had been defined, it cannot be ruled out that interpretation biases might have been occurred.

It has also to be noted that in this study, a rather specific sample was examined. Individuals from the IRD-exclude-group displayed an intense, even excessive online behavior that was also related to at least some negative repercussions. Yet, these negative consequences were not of sufficient severity to justify the diagnosis of IRD/IGD. Thus, there was no typical control group. Having a further control group consisting of non-clinical participants should be a goal of upcoming studies.

Nevertheless, this study contributes to the evaluation process of the DSM criteria for IGD and possibly IRD. Especially, the combination of qualitative and quantitative methods of data analyses, the clinical sample our analyses were based on, and the availability of an external clinical evaluation on the IRD status can be regarded as a good standard of quality.

Funding sources: No financial support was received for this study.

Authors' contribution: KWM: principal investigator and study concept and design, analysis and interpretation of data, and statistical analyses. MD: analysis and interpretation of data. KW: study supervision, analysis and interpretation of data, and statistical analyses. MEB: study supervision.

Conflict of interest: The authors report no financial or other relationship relevant to the subject of this article.

REFERENCES

- Aarseth, E., Bean, A. M., Boonen, H., Colder Carras, M., Coulson, M., Das, D., Deleuze, J., Dunkels, E., Edman, J., Ferguson, C. J., Haagsma, M. C., Helmersson Bergmark, K., Hussain, Z., Jansz, J., Kardefelt-Winther, D., Kutner, L., Markey, P., Nielsen, R. K. L., Prause, N., Przybylski, A., Quandt, T., Schimmenti, A., Starcevic, V., Stutman, G., Van Looy, J., & Van Rooij, A. J. (2017). Scholars' open debate paper on the World Health Organization ICD-11 Gaming Disorder proposal. *Journal of Behavioral Addictions, 6*(3), 267–270. doi:10.1556/2006.5.2016.088
- American Psychiatric Association [APA]. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)* (5th ed.). Arlington, TX: American Psychiatric Association.
- Billieux, J. (2012). Problematic use of the mobile phone: A literature review and a pathways model. *Current Psychiatry Reviews, 8*(4), 299–307. doi:10.2174/157340012803520522
- Billieux, J., Schimmenti, A., Khazaal, Y., Maurage, P., & Heeren, A. (2015). Are we overpathologizing everyday life? A tenable blueprint for behavioral addiction research. *Journal of Behavioral Addictions, 4*(3), 119–123. doi:10.1556/2006.4.2015.009
- Brand, M., Young, K. S., & Laier, C. (2014). Prefrontal control and Internet addiction: A theoretical model and review of neuropsychological and neuroimaging findings. *Frontiers in Human Neuroscience, 8*, 375. doi:10.3389/fnhum.2014.00375
- Fauth-Bühler, M., & Mann, K. (2017). Neurobiological correlates of Internet gaming disorder: Similarities to pathological gambling. *Addictive Behaviors, 64*, 349–356. doi:10.1016/j.addbeh.2015.11.004
- Gentile, D. A., Choo, H., Liau, A., Sim, T., Li, D., Fung, D., & Khoo, A. (2011). Pathological video game use among youths: A two-year longitudinal study. *Pediatrics, 127*(2), e319–e329. doi:10.1542/peds.2010-1353
- Griffiths, M. D. (2010). The role of context in online gaming excess and addiction: Some case study evidence. *International Journal of Mental Health and Addiction, 8*(1), 119–125. doi:10.1007/s11469-009-9229-x
- Griffiths, M. D., Van Rooij, A. J., Kardefelt-Winther, D., Starcevic, V., Király, O., Pallesen, S., Müller, K. W., Dreier, M., Carras, M., Prause, N., King, D. L., Aboujaoude, E., Kuss, D. J., Pontes, H. M., Fernandez, O. L., Nagygyorgy, K., Achab, S., Billieux, J., Quandt, T., Carbonell, X., Ferguson, C., Hoff, R. A., Derevensky, J., Haagsma, M., Delfabbro, P., Coulson, M., Hussain, Z., & Demetrovics, Z. (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. doi:10.1111/add.13057
- Hasin, D. S., Fenton, M. C., Beseler, C., Park, J. Y., & Wall, M. M. (2012). Analyses related to the development of DSM-5 criteria for substance use related disorders: 2. Proposed DSM-5 criteria for alcohol, cannabis, cocaine and heroin disorders in 663 substance abuse patients. *Drug and Alcohol Dependence, 122*(1–2), 28–37. doi:10.1016/j.drugalcdep.2011.09.005
- Jo, S. J., Yim, H. W., Lee, H. K., Lee, H. C., Choi, J. S., & Baek, K. Y. (2017). The Internet Game Use-Elicited Symptom Screen proved to be a valid tool for adolescents aged 10–19 years. *Acta Paediatrica, 107*(3), 511–516. doi:10.1111/apa.14087
- Keyes, K. M., Krueger, R. F., Grant, B. F., & Hasin, D. S. (2011). Alcohol craving and the dimensionality of alcohol disorders. *Psychological Medicine, 41*(3), 629–640. doi:10.1017/S003329171000053X
- 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. doi:10.1016/j.addbeh.2015.11.005
- Király, O., Tóth, D., Urbán, R., Demetrovics, Z., & Maraz, A. (2017). Intense video gaming is not essentially problematic. *Psychology of Addictive Behaviors, 31*(7), 807–817. doi:10.1037/adb000031610.1037/adb0000316
- Ko, C. H., Yen, J. Y., Chen, S. H., Wang, P. W., Chen, C. S., & 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
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9. Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613. doi:10.1046/j.1525-1497.2001.016009606.x
- Kuss, D. J., Griffiths, M. D., & Pontes, H. M. (2017). Chaos and confusion in DSM-5 diagnosis of Internet Gaming Disorder: Issues, concerns, and recommendations for clarity in the field. *Journal of Behavioral Addictions, 6*(2), 103–109. doi:10.1556/2006.5.2016.062
- Kuss, D. J., Shorter, G. W., van Rooij, A. J., Griffiths, M. D., & Schoenmakers, T. (2013). Assessing Internet addiction using the parsimonious Internet addiction components model – A preliminary study. *International Journal of Mental Health and Addiction, 12*(3), 351–366. doi:10.1007/s11469-013-9459-9
- 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
- Love, T., Laier, C., Brand, M., Hatch, L., & Hajela, R. (2015). Neuroscience of Internet pornography addiction: A review and update. *Behavioral Sciences, 5*(3), 388–433. doi:10.3390/bs5030388
- Maráz, A., Király, O., & Demetrovics, Z. (2015). The diagnostic pitfalls of surveys: If you score positive on a test of addiction, you still have a good chance not to be addicted. A response to Billieux et al. 2015. *Journal of Behavioral Addictions, 4*(3), 151–154. doi:10.1556/2006.4.2015.026
- Müller, K. W., Beutel, M. E., & Wölfling, K. (2014). A contribution to the clinical characterization of Internet addiction in a sample of treatment seekers: Validity of assessment, severity of psychopathology and type of co-morbidity. *Comprehensive Psychiatry, 55*(4), 770–777. doi:10.1016/j.comppsy.2014.01.010
- Müller, K. W., Dreier, M., Duven, E., Giral, S., Beutel, M. E., & Wölfling, K. (2016). A hidden type of Internet addiction? Intense and addictive use of social networking sites in

- adolescents. *Computers in Human Behavior*, 55, 172–177. doi:[10.1016/j.chb.2015.09.007](https://doi.org/10.1016/j.chb.2015.09.007)
- Müller, K. W., Dreier, M., Duven, E., Giral, S., Beutel, M. E., & Wölfling, K. (2017). Adding clinical validity to the statistical power of large-scale epidemiological surveys on Internet addiction in adolescence: A combined approach to investigate psychopathology and development-specific personality traits associated with Internet addiction. *Journal of Clinical Psychiatry*, 78(3), e244–e251. doi:[10.4088/JCP.15m10447](https://doi.org/10.4088/JCP.15m10447)
- Müller, K. W., Glaesmer, H., Brähler, E., Wölfling, K., & Beutel, M. E. (2014). Prevalence of Internet addiction in the general population: Results from a German population-based survey. *Behaviour and Information Technology*, 33(7), 757–766. doi:[10.1080/0144929X.2013.810778](https://doi.org/10.1080/0144929X.2013.810778)
- Pawlikowski, M., Nader, I. W., Burger, C., Stieger, S., & Brand, M. (2014). Pathological Internet use – It is a multidimensional and not a unidimensional construct. *Addiction Research & Theory*, 22(2), 166–175. doi:[10.3109/16066359.2013.793313](https://doi.org/10.3109/16066359.2013.793313)
- Petry, N. M., Rehbein, F., Gentile, D. A., Lemmens, J. S., Rumpf, H. J., Mossle, T., Bischof, G., Tao, R., Fung, D. S. S., Borges, G., Auriacombe, M., Ibáñez, A. G., Tam, P., & 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](https://doi.org/10.1111/add.12457)
- Rehbein, F., Kliem, S., Baier, D., Mößle, T., & Petry, N. M. (2015). Prevalence of Internet gaming disorder in German adolescents: Diagnostic contribution of the nine DSM-5 criteria in a state-wide representative sample. *Addiction*, 110(5), 842–851. doi:[10.1111/add.12849](https://doi.org/10.1111/add.12849)
- Rosenkranz, T., Müller, K. W., Dreier, M., Beutel, M. E., & Wölfling, K. (2017). Addictive potential of internet applications and differential correlates of problematic use in Internet gamers versus generalized Internet users in a representative sample of adolescents. *European Addiction Research*, 23(3), 148–156. doi:[10.1159/000475984](https://doi.org/10.1159/000475984)
- Söderberg, P., Tungström, S., & Armelius, B.A. (2005). Special section on the GAF: Reliability of global assessment of functioning ratings made by clinical psychiatric staff. *Psychiatric Services*, 56, 434–438. doi:[10.1176/appi.ps.56.4.434](https://doi.org/10.1176/appi.ps.56.4.434)
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Lowe, B. (2006). A brief measure for assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097. doi:[10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)
- Weeks, J. W., Spokas, M. E., & Heimberg, R. G. (2007). Psychometric evaluation of the Mini-Social Phobia Inventory (Mini-Spin) in a treatment-seeking sample. *Depression and Anxiety*, 24(6), 382–391. doi:[10.1002/da.20250](https://doi.org/10.1002/da.20250)
- Wölfling, K., Beutel, M. E., & Müller, K. W. (2012). Construction of a standardized clinical interview to assess Internet addiction: First findings regarding the usefulness of AICA-C. *Journal of Addiction Research and Therapy*, S6, 1–7. doi:[10.4172/2155-6105.S6-003](https://doi.org/10.4172/2155-6105.S6-003)
- Wölfling, K., Beutel, M. E., & Müller, K. W. (2016). OSV-S–Skala zum Onlinesuchtverhalten [AICA-S: Scale for the assessment of Internet and computer game addiction]. In K. Geue, B. Strauß, & E. Brähler (Eds.), *Diagnostische Verfahren in der Psychotherapie (Diagnostik für Klinik und Praxis)* [Diagnostic measures for psychotherapy] (pp. 362–366). Göttingen, Germany: Hogrefe.