

INTERNET GAMING DISORDER AS A FORMATIVE CONSTRUCT: IMPLICATIONS
FOR CONCEPTUALIZATION AND MEASUREMENT.

Manuscript in-press, cite as:

Van Rooij, A. J., Van Looy, J., & Billieux, J. (in-press). Internet Gaming Disorder as a formative construct: Implications for conceptualization and measurement. *Psychiatry and Clinical Neurosciences*.

Authors:

Antonius J. van Rooij, Ph.D.¹

Jan van Looy, Ph.D.¹

Joël Billieux, Ph.D.^{2,3}

¹ iMinds-MICT-Ghent University, Department of Communication Sciences, Ghent, Belgium

² Laboratory for Experimental Psychopathology (LEP), Psychological Science Research Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

³ Internet and Gambling Disorders Clinic, Department of Adult Psychiatry, Cliniques Universitaires Saint-Luc, Brussels, Belgium

Corresponding Author:

Antonius J. van Rooij, Ph.D.

Ghent University, Campus Aula, Department of Communication Studies, iMinds-MICT
Korte Meer 7-9-11, 9000 Ghent, Belgium

Mail: Tony.vanRooij@UGent.be, Telephone: +32 9 264 84 76

Submitted: January, 2016

Resubmitted: April 2016

In-press: May 2016

Abstract

Background. Some people have serious problems controlling their internet and video game use. The DSM-5 now includes a proposal for ‘Internet Gaming Disorder’ as a condition in need of further study. Various studies aim to validate the proposed diagnostic criteria for Internet Gaming Disorder and multiple new scales have been introduced that cover the suggested criteria.

Approach. Using a structured approach, we demonstrate that Internet Gaming Disorder might be better interpreted as a formative construct, as opposed to the current practice of conceptualizing it as a reflective construct. Incorrectly approaching a formative construct as a reflective one causes serious problems in scale development including (a) incorrect reliance on item-to-total scale correlation to exclude items and incorrectly relying on indices of inter-item reliability that do not fit the measurement model (e.g., Cronbach’s α) (b) incorrect interpretation of composite or mean scores that assume all items are equal in contributing value to a sum score, and (c) biased estimation of model parameters in statistical models.

Implications. We show that these issues are impacting current validation efforts through two recent examples. A reinterpretation of Internet Gaming Disorder as a formative construct has broad consequences for current validation efforts and provides opportunities to reanalyze existing data. We discuss three broad implications for current research: (1) Composite latent constructs should be defined and used in models, (2) Item exclusion and selection should *not* rely on item-to-total scale correlations, and (3) Existing definitions of Internet Gaming Disorder should be enriched further.

Key words:

Internet Gaming Disorder; Formative measurement model; Validation; Reflective measurement model; Psychometrics

Internet (video) gaming disorder: new condition, same debate

Various studies have shown that some people have serious problems controlling their internet and video game use.¹⁻⁴ It is currently being debated if these problems warrant a new clinical diagnosis and controversy on this subject continues.⁵⁻⁷ Despite this ongoing debate, the *American Psychiatric Association* has already made a move in the direction of a new disorder. Whereas an initial online draft for the 5th *Diagnostic and Statistical Manual of Mental Disorders* (DSM) only discussed a concept of ‘Internet Use Disorder’,⁸ the manual ultimately included a proposal for ‘Internet Gaming Disorder’ (IGD) in the appendix (section 3), as a condition in need of further study.⁹

IGD criteria were directly transposed from those classifying and diagnosing other addictive disorders, i.e. substance use and gambling disorders. The condition is summarized as the ‘Persistent and recurrent use of the Internet to engage in games, often with other players, leading to clinically significant impairment or distress’, as indicated by five or more out of nine proposed items. This specific focus on online gaming was driven both by the lack of evidence regarding the etiology and course of other Internet-related disorders¹⁰ and the fact that the majority of treatment-seeking individuals displaying “Internet addiction” symptoms actually consists of video game players.^{3,4,11}

The advent of a potential new diagnosis has fueled the existing debate.¹²⁻¹⁴ We believe that both the individually suggested criteria in the IGD proposal¹⁵ and the somewhat confirmatory approach consisting in applying substance abuse criteria to IGD⁶ warrant further critical discussion. However, this debate is already happening elsewhere and will not be reiterated here.¹²⁻¹⁴ The DSM-5 proposal, however imperfect, at least provides the field with an opportunity to move forward by centering the debate on a common point of reference. While mindful of the atheoretical nature and controversial status of the criteria, our

aim in the current article is to extend our thinking on IGD conceptualization and measurement by starting from the tentative criteria proposed in Section 3 of the DSM-5.

The proliferation of scales assessing IGD

Before the DSM-5 IGD proposal, the majority of empirical internet and videogame ‘addiction’ research papers have conceptualized these conditions using criteria derived from the DSM-IV criteria for pathological gambling (which were themselves transposed from substance abuse criteria¹⁶). Through this lens, various ‘components’ were identified and applied to the measurement of internet and videogame disorders. Some examples include Tao’s criteria, Griffith’s components model, and Young’s Internet Addiction Test.^{15–19} While these sets differ slightly, they are all grounded in the same *a priori* assumption, i.e. that the DSM-IV criteria for pathological gambling can validly be extended to other behavioral addictions. The vast majority of previous studies used these ‘components’ *and* collected large scale survey data.²⁰ These clinical checklists (with yes/no questions) are regularly translated into scales with dimensional items to better fit research in survey samples (e.g., 5-point Likert scales). However, this translation to a new research method might generate certain issues. For example, multiple authors have struggled with the determination of cut-off scores.^{21–23} In the absence of robust normative data, which are unavailable for these instruments, employed cut-off scores and population prevalence estimates have remained dubious at best. For example, numerous studies conducted with Young’s Internet Addiction Test used and reproduced cut-off scores that were not based on empirical data.^{1,24} Thus, their numerical estimates of pathological cases have low validity and might be inflated.^{15,25}

The DSM-5 IGD proposal has mainly spawned research dedicated to the validation of the proposed diagnostic criteria for IGD in a wider population setting.^{26–30} Multiple new scales designed to assess the IGD criteria have emerged since 2014. Indeed, a literature search conducted on Google Scholar reveals the existence of six distinct measurement

instruments already (*terms: [allintitle: (gaming disorder OR IGD) AND (scale OR test OR assessment OR tool)]*; 11 Jan. 2016). The 20 item *Internet Gaming Disorder Test* (IGD-20) was tested among an online sample of 1003 gamers^{29,31} and a separate nine item *Internet Gaming Disorder Scale–Short-Form* (IGDS9-SF) was tested in the same sample by UK researchers.³⁰ A Hungarian study presented and tested an *Internet Gaming Disorder Test (IGDT-10)* in an online sample of 4887 gamers.²⁶ A Dutch study employed a representative sample of 2444 respondents to test a 27-item and a 9-item version of an instrument titled *Internet Gaming Disorder Scale (IGDS-9)*.³² Another group working from the Netherlands presented a clinical assessment instrument which covers the IGD criteria and was field-tested among a sample of treatment-seeking online gamers: the *Clinical-Video game Addiction Test (C-VAT 2.0)*.³³ Finally, a rather unconventional *hybrid scale* with questions pertaining to both Internet and online gaming³⁴ was tested among 225 Korean college students: The mixing of items targeting two different domains and the small sample size of the study question the validity of this scale, which will not be further considered here. This proliferation of scales will undoubtedly further complicate comparison between studies and might ultimately negatively influence the field, which repeats the earlier confusion with a large number of pre-IGD proposal scales.¹⁶

In these validation studies, an implicit assumption is generally made that the hypothesized construct of IGD *causes* the various criteria, such as a loss of control over gaming behavior. This approach is known as *reflective measurement*, and it implies that items are expected to co-vary and to be mutually interchangeable.³⁵ In the present article, we will develop the idea that some, if not all, criteria used to measure IGD (the focal construct) should be interpreted as formative sub-constructs instead. In a formative measurement model causality is reversed and items are thought to compose or *cause* the focal construct.³⁵ This reversal in thinking questions several common practices in the field, including the reporting

of reliability coefficients, the use of cut-off scores that assume a reflective model, and the estimation of relationships between latent constructs in structural models.³⁶

Reflective and formative constructs

In an important yet underused³⁷ theoretical paper, Bollen and Lennox (1991) suggested that theory development typically neglects the assessment of the relationship between construct and measure (usually, these measures are questionnaire items).³⁸ Constructs, such as ‘depression’ or ‘videogame addiction’, are generally not directly observable. Measures, however, are both observable and related to the construct. Generally, the measures are assumed to be an effect of the construct, e.g. they are caused by it. One of the key points raised by Bollen and Lennox is that a causally reversed way of thinking sometimes better corresponds to the relationship between construct and measure.³⁸

If the construct is thought to cause the measures it is referred to as a ‘reflective’ construct. This approach is dominant in the current psychology and psychiatry literature.^{35,36,39} Some examples would be the idea that intelligence *causes* IQ test performance, that having an extraverted personality type *causes* social behavior, or that impaired inhibitory control *causes* substance abuse. Let us assume for the sake of argument that intelligence is a simple, one-dimensional construct. We might aim to uncover the non-observable construct of intelligence with measures that include a visual puzzle, a maze, or a math problem. With a reflective measurement approach, removing an individual measure from the set would not change the nature of the hypothesized intelligence construct. The measures are thought to be associated with each other. This has consequences for reliability assessment, where one might drop an item with low-item-total score correlations from a scale to enhance internal consistency.⁴⁰

The opposite direction, whereby the measures *cause* or constitute the construct is known as a formative construct: the measures *form* the construct.³⁵ A generalized price index,

for example, might include a composite score that includes the pricing of bread, smartphones, and cinema tickets. The individual prices add to the price index, but the price index does not *cause* the prices in any way. The prices included determine the index rather than the other way around. Moreover, removing a measure from the composite changes the nature of the construct itself. For instance, removing cinema tickets might mean that the price index is now, perhaps, better viewed as a ‘bare life necessities’ price index. Correlation between measures is *possible* but is *not a prerequisite*.³⁵ This complicates decisions on the inclusion or exclusion of a measure in a construct as it cannot always be done on purely statistical (correlational) grounds: theoretical and conceptual reasoning is required.⁴¹

Reflective versus formative: why does it matter?

There are several key differences between reflective and formative constructs which, when disregarded, may cause methodological errors and lead to biased conclusions^{35,36,38,42}

The main differences are the following:

1. Removal of a single measure changes the nature of a formative construct itself, while removal of a single measure from a reflective construct has no influence on construct validity, and is even sometimes warranted when one measure is not sufficiently correlated with the construct.^{35,38}
2. Using reliability indicators that focus on internal consistency, like Cronbach’s α , and working with item-to-total scale correlation or inter-item correlation is less relevant in formative measurement models. Correlation between items is not necessarily expected because each item adds conceptually to the whole construct and is not mutually interchangeable with other items.^{40,41,43}
3. Estimation of structural relationships for models incorrectly specified as reflective has been shown to result in inflated estimates of structural parameters when building

explanatory theoretical models. Moreover, these mistakes are not likely to be detected by commonly used goodness-of-fit indices.^{36,44}

Some of these issues are more relevant (e.g. estimation of structural relationships) and pronounced when doing survey research than when using a checklist approach (yes/no questions) in a clinical context. For example, the clinical setting generally has a relatively strong measure of validity in the form of observations by a therapist, and people reporting for treatment are very likely to have some form of impairment or life disruption.³³ This anchor does not exist in survey research with mostly healthy respondents. Accordingly, any attempt to validate the disorder by relying exclusively on survey data, using a confirmative and reflective approach to measurement, suffers from weak validity. Even more so if a formative interpretation fits better with the proposed ‘disorder’, as that would necessitate that all measures included are individually relevant as they are not interchangeable. Nonetheless, as reported in Table 1, most of available studies have relied on survey samples.

Relevance of the reflective/formative distinction for IGD

To the best of our knowledge, ongoing validation efforts of IGD criteria have not included deliberation on the formative or reflective nature of the focal construct (IGD) and the sub-constructs (criteria). This might be fruitful, as this distinction has major implications in terms of scale construction, validation, and (composite) scale score interpretation. Crucially, it would be incorrect to presuppose, in the case of a formative construct, that a higher number of items endorsed reflects a greater severity or impact.⁴⁵ Using a structured approach,^{36,44} the current study evaluates to what extent it might be appropriate to employ formative constructs in the definition and validation of IGD.

Evaluation of current conceptual approach to the measurement of IGD

While the concept of formative measurement has been available for a long time^{35,38}, its impact has been limited as most papers automatically imply and apply a reflective approach to measurement. MacKenzie et al. suspect this might be the case because a comprehensive set of recommendations on formative scale validation was previously lacking, and therefore proposed a stepwise approach.⁴⁰ In the present paper, the first step of their approach, which refers to a *clear conceptualization of the nature and dimensionality of the focal construct and its sub constructs*, is applied. This initial step comprises conceptualization of the construct, with a specific focus on both the *dimensional* and *reflective / formative* nature of the possible sub-constructs. Subsequent steps (not addressed here) would involve the development of items (step 2), assessment of content validity of items (step 3), and specification of a measurement model (step 4). However, as we propose a conceptual critique of the current validation approach and related assessment tools developed in the IGD research field, these subsequent steps are beyond the scope of this article. The remaining six steps, which will not be considered here, deal with further validation procedures, such as scale evaluation and refinement (step 5, 6), formal validation (steps 7, 8, 9), and normative data development (step 10).

Table 1 summarizes the procedures used to validate the IGD scales mentioned earlier. Four scales were included (a) preferring author-recommended brief versions of scales, and (b) the most recent work by authors if they put out multiple scales over time. The final selection includes the IGDS-9,³⁰ the IGDT-10,²⁶ the IGDS9-SF,³² and the C-VAT 2.0.³³ Table 2 reports the item phrasing for each of these scales and contrasts them with the DSM-5 criteria.

Table 1. Summary of validation procedure, DSM-5 proposal and four IGD scales.

	DSM-5 Proposal, IGD⁹	IGDS9-SF³⁰	IGDT-10²⁶	IGDS-9 item short scale version³²	Clinical-Video Game Addiction Test 2.0³³
Item scale (answer options)	Dichotomous: yes/no	5-point Likert scale: 1 (“Never”), 2 (“Rarely”), 3 (“Sometimes”), 4 (“Often”), and 5 (“Very Often”).	3-point scale: Never (0), Sometimes (1), Often (2)	Half the sample: <ul style="list-style-type: none"> • Polytomous: Never (0) to Every Day (5), Half the sample: <ul style="list-style-type: none"> • Dichotomous (Yes/No) 	Dichotomous: yes/no
Temporal reference	The following questions refer to the past year (the past 12 months)	Last 12 months	Last 12 months	During the last year...	The following questions refer to the past year (the past 12 months)
Measurement model analysis	n/a	Reflective measurement model: One factor structure found in Exploratory and Confirmatory Factor Analyses.	Reflective measurement model: One factor structure found in Confirmatory Factor Analysis.	Reflective measurement model: One factor structure found in Confirmatory Factor Analysis for both scales.	n/a, study explored validity of the 5/9 cut-off score proposal by DSM in a clinical setting.
Reported reliability	n/a	Cronbach's $\alpha = .88$ full sample, over 9 items	Cronbach's $\alpha = .79$, over 10 items (Cronbach's $\alpha = .68$, over 9 binary items)	Cronbach's $\alpha = .83$ (dichotomous) and $.95$ (polytomous), over 9 items	n/a
Sample	n/a	N=1061, online opportunity sample of gamers	N=4887, online opportunity sample of gamers.	Representative population sample with n=1193 for the polytomous items and n=1251 for the dichotomous items	N=32 young patients in treatment for problems with video games

Table 2. Item phrasing for four proposed Internet Gaming Disorder scales

DSM-5 Proposal, IGD⁹	IGDS9-SF³⁰	IGDT-10²⁶	IGDS-9 item short scale version³²	C-VAT 2.0³³
1. Preoccupation with Internet games. (The individual thinks about previous gaming activity or anticipates playing the next game; Internet gaming becomes the dominant activity in daily life). Note: This disorder is distinct from Internet gambling, which is included under gambling disorder.	1. Do you feel preoccupied with your gaming behaviour? (Some examples: Do you think about previous gaming activity or anticipate the next gaming session? Do you think gaming has become the dominant activity in your daily life?)	1. When you were not playing, how often have you fantasized about gaming, thought of previous gaming sessions, and/or anticipated the next game?	...have there been periods when all you could think of was the moment that you could play a game?	1. Could you hardly think about anything else than playing games when you were not gaming?
2. Withdrawal symptoms when Internet gaming is taken away. (These symptoms are typically described as irritability, anxiety, or sadness, but there are no physical signs of pharmacological withdrawal.)	2. Do you feel more irritability, anxiety or even sadness when you try to either reduce or stop your gaming activity?	2. How often have you felt restless, irritable, anxious and/or sad when you were unable to play or played less than usual?	...have you been feeling miserable when you were unable to play a game?	2. Did you feel stressed, annoyed, or angry if you were not allowed or could not play games?
3. Tolerance—the need to spend increasing amounts of time engaged in Internet games.	3. Do you feel the need to spend increasing amount of time engaged gaming in order to achieve satisfaction or pleasure?	3. Have you ever in the past 12 months felt the need to play more often or played for longer periods to feel that you have played enough?	...have you felt unsatisfied because you wanted to play more?	3. Did you spend more and more time on playing videogames?
4. Unsuccessful attempts to control the participation in Internet games.	4. Do you systematically fail when trying to control or cease your gaming activity?	4. Have you ever in the past 12 months unsuccessfully tried to reduce the time spent on gaming?	...were you unable to reduce your time playing games, after others had repeatedly told you to play less?	4. Did you unsuccessfully try to spend less time on games?
5. Loss of interests in previous hobbies and entertainment as a result of, and with the exception of, Internet games.	5. Have you lost interests in previous hobbies and other entertainment activities as a result of your engagement with the game?	5. Have you ever in the past 12 months played games rather than meet your friends or participate in hobbies and pastimes that you used to enjoy before?	...have you lost interest in hobbies or other activities because gaming is all you wanted to do?	5. Did you have to give up or strongly reduce important activities because of gaming? Examples: sports, work, or seeing friends/family

6. Continued excessive use of Internet games despite knowledge of psychosocial problems.	6. Have you continued your gaming activity despite knowing it was causing problems between you and other people?	6. Have you played a lot despite negative consequences (for instance losing sleep, not being able to do well in school or work, having arguments with your family or friends, and/or neglecting important duties)?	...have you had arguments with others about the consequences of your gaming behavior?	6. Did you regularly neglect important commitments or persons in order to play videogames? (Examples: social relationships in real-life/offline, (home) work, other hobbies, school, or work)
7. Has deceived family members, therapists, or others regarding the amount of Internet gaming.	7. Have you deceived any of your family members, therapists or others because the amount of your gaming activity?	7. Have you tried to keep your family, friends or other important people from knowing how much you were gaming or have you lied to them regarding your gaming?	...have you hidden the time you spend on games from others?	7. Did you sometimes lie to others about the amount of time you spend on video games?
8. Use of Internet games to escape or relieve a negative mood (e.g., feelings of helplessness, guilt, anxiety).	8. Do you play in order to temporarily escape or relieve a negative mood (e.g., helplessness, guilt, anxiety)?	8. Have you played to relieve a negative mood (for instance helplessness, guilt, or anxiety)?	...have you played games so that you would not have to think about annoying things?	8. Did you regularly play videogames to avoid thinking about problems (difficulties)
9. Has jeopardized or lost a significant relationship, job, or educational or career opportunity because of participation in Internet games.	9. Have you jeopardised or lost an important relationship, job or an educational or career opportunity because of your gaming activity?	9. Have you risked or lost a significant relationship because of gaming? 10. Have you ever in the past 12 month jeopardized your school or work performance because of gaming?	...have you experienced serious conflicts with family, friends or partner because of gaming?	9. Did you play games even though you knew this was causing problems with your family, friends, at work, or at school?

Establish the nature and dimensionality of the construct

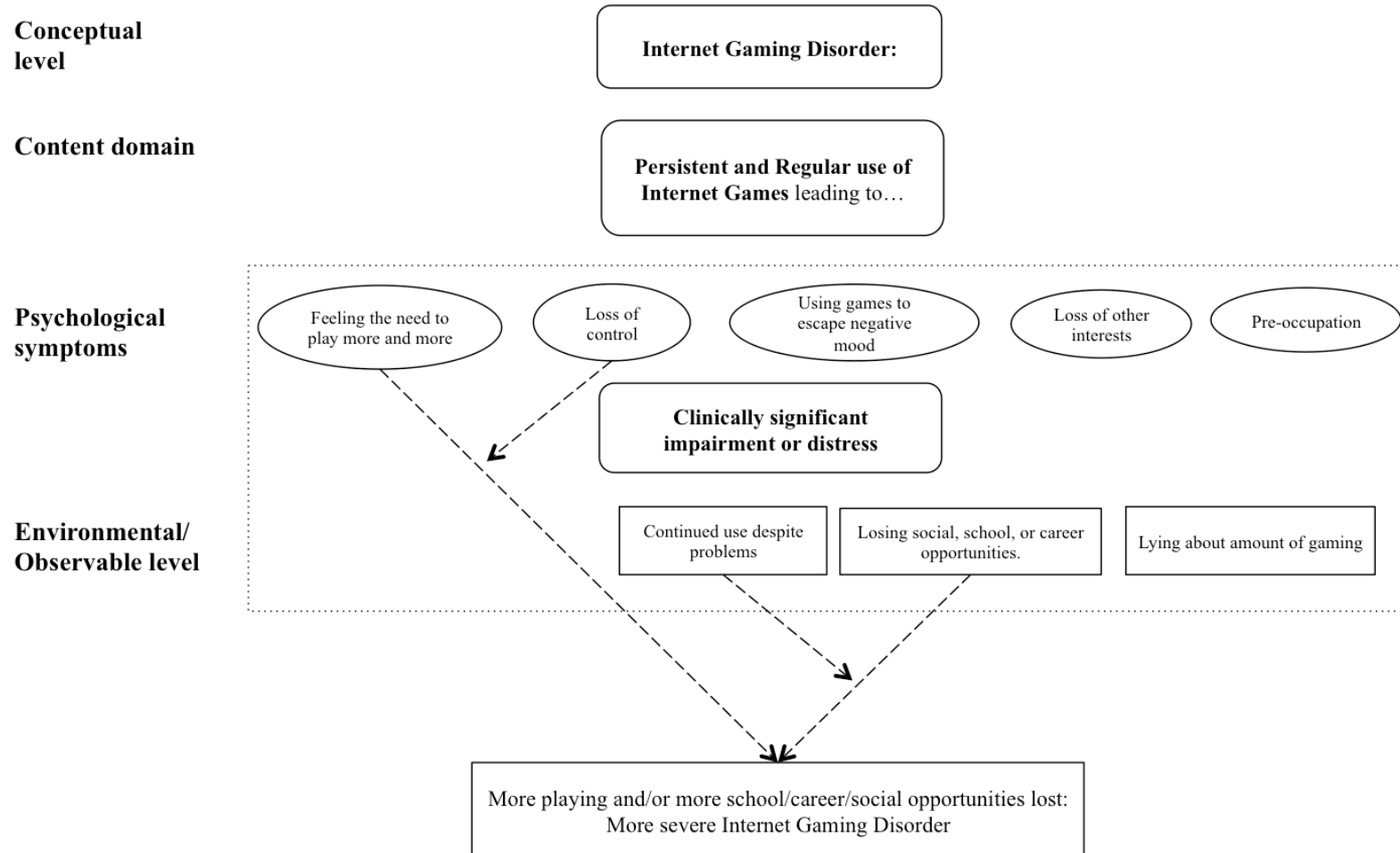
The DSM-5⁹ proposal delineates the construct domain for IGD through the following description (p. 795):

‘Persistent and recurrent use of the Internet to engage in games, often with other players, leading to clinically significant impairment or distress’, as indicated by 5 or more out of 9 proposed items, which might be summarized as: preoccupation, withdrawal symptoms, tolerance or the need to play more and more, loss of control, loss of other interests, continued use despite problems, deceiving others about amount of gaming, using games to escape negative mood states, and losing social, school, or career opportunities. Those with severe Internet gaming disorder will have more hours spent on the computer and more severe loss of relationships or career or school opportunities.

The DSM-5 text limits the subject matter to Internet games: online gambling, social networking activities, and sexual online activities are excluded. The first part of this proposal establishes that some types of persistent and regular use are required in association with clinically significant impairment or distress. A more detailed checklist is suggested through the nine proposed items that supposedly cover the two above-mentioned domains. In the small print, negative outcomes are connected to two rather different outcomes, i.e. increased in the hours played (often considered as a proxy measure of ‘tolerance’) and more loss of opportunities. See Figure 1 for a visual representation of the IGD definition.

When defining the conceptual domain, it is also helpful to distinguish the object or entity to which a measure applies.⁴⁰ From Table 2, it appears that the majority of criteria refer to psychological symptoms (e.g. anxiety, pre-occupation and thinking about games, loss of interest in other hobbies, loss of control, *need* to spend more and more time on games), while a minority of criteria refers to socio-environmental states: deceiving others, losing opportunities, and continuing despite psycho-social problems.

Figure 1: Graphical representation of the DSM-5 IGD proposal



Evaluate the relationship between items and constructs (dimensionality)

According to MacKenzie and colleagues⁴⁰, it is important to determine whether the construct under investigation involves interrelated sub-constructs. According to these authors, two questions can guide the process of discovery here:

1. How distinctive are the essential characteristics from each other?
2. Can any of these characteristics be eliminated without altering the conceptual domain of the construct?

We shall now re-interpret the IGD proposal through these questions. The resulting framework (i.e. figure 2.) is not meant to be a definitive proposal for a formative interpretation of IGD, but merely provides one illustration of a formative interpretation of IGD that avoids some of the pitfalls of reflective modeling. The proposition formulated here should be tested in future empirical work, besides alternative hypotheses (such as a fully formative interpretation of IGD).

We assume that each of the nine criteria described in the IGD proposal forms an indication of clinically significant impairment or distress. If we evaluate the criteria for distinctiveness, only two seem mutually connected and interchangeable to some degree (i.e. pre-occupation and loss of other interests). This implies that grouping them or removing one of them does not alter the nature of the IGD construct. None of the other criteria seem mutually interchangeable. Combining the two criteria, we propose this integrated sub-construct:

- *Extreme focus on gaming*: This reflective sub-construct is proposed to regroup the *pre-occupation* and *loss of other interests* criteria and relates to *single-minded thinking*. The first mainly reflects an internal state and the second can be at least partly observable to outsiders. Either criterion might suffice to increase mental focus on gaming and both can be assumed to be interchangeable to a certain degree, which is indicative of a reflective sub-construct.

Dimensionality: relationships between focal and sub-constructs

The next phase of the process consists of evaluating the degree to which the construct of interest relates to its sub-constructs. Two questions can guide this process⁴⁰:

1. Are the sub-constructs manifestations (*reflective*), rather than defining characteristics (*formative*), of IGD?
2. Would a change in IGD severity be associated with changes in all the sub-dimensions (*reflective*), or can a change in IGD severity be restricted to a change in a subset of sub-dimensions (*formative*)?

Our view is that an increase in IGD severity will *not* necessarily be associated with an increase in all its diagnostic criteria by default. Let us assume that IGD increases, due to an adverse negative life event, a decrease in well-being, or the release of a long-expected new game. If a reflective model would be correct, an increase of the disorder would systematically increase the degree of single-mindedness, the tendency to use games to escape negative mood, the problems in various and divergent aspects of life, AND the amount of lying about use. As this does not make logical sense, it appears that the focal construct of IGD might have a formative nature to some extent.

The DSM-5 proposal also specifies that both an increase in hours spent and a more pronounced loss of opportunities (relationship, school, or work) are indications of more severe IGD. The escalation of hours spent on gaming implies both the *need* to be involved in gaming for more and more hours and a simultaneous *loss of control*. These two aspects are both required and not combinable or mutually interchangeable. We therefore propose the following sub-construct:

- *Escalated use*: This formative sub-construct includes both a ‘loss of control over playing’ and ‘the need to play more’.

A similar line of reasoning applies to the increase in missed opportunities / problems and persistence despite experiencing problems. Both seem required to reflect a more severe loss

of chances in the real life and both deal with problems quite directly. We thus propose to combine them into a formative sub-construct as well:

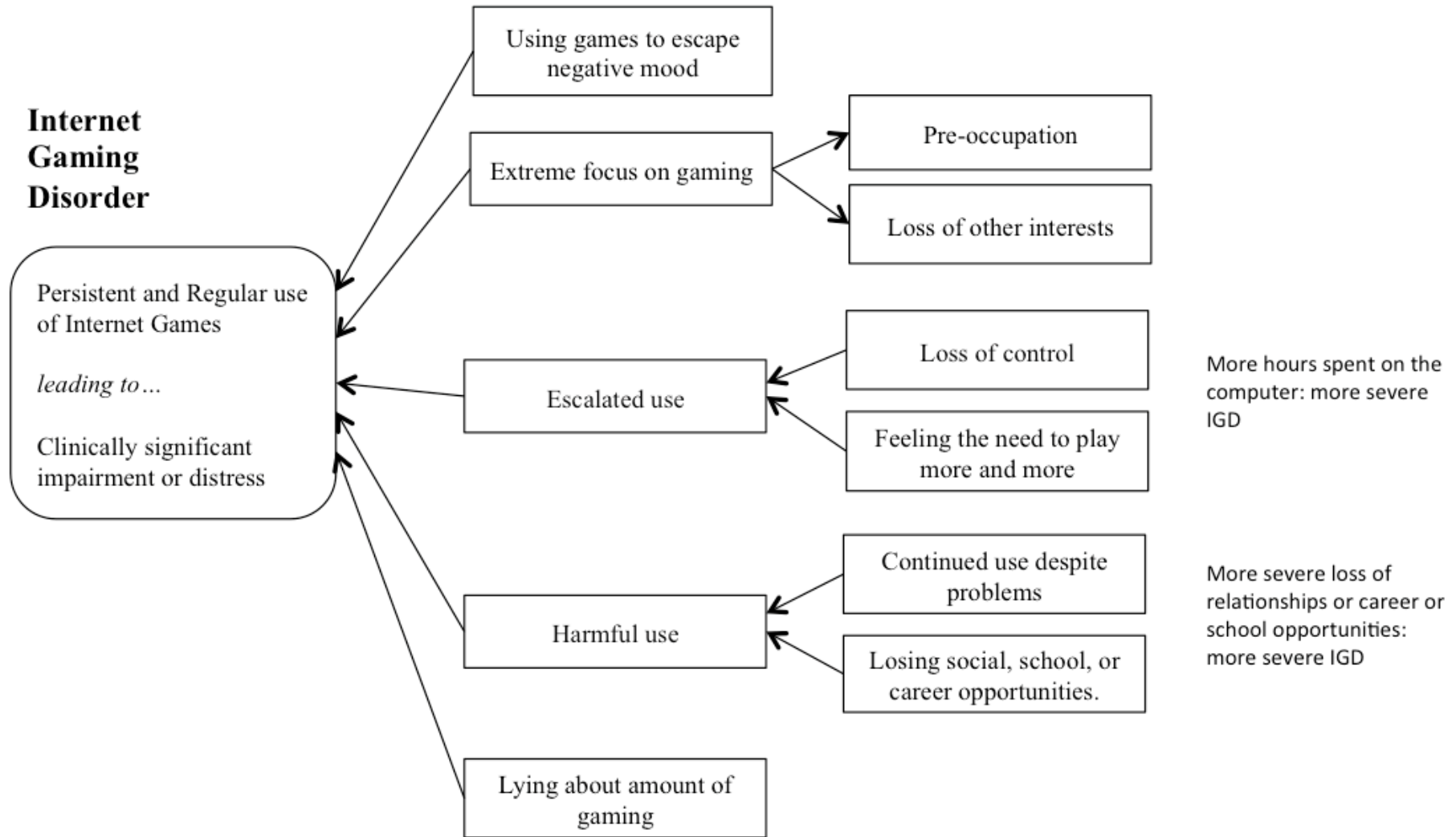
- *Harmful use*: This formative sub-construct includes the ‘continuation of gaming despite problems’ and the ‘experiencing of problems’, such as loss of opportunities in real life.

Figure 2 depicts IGD as a main focal construct, formatively determined by mood modification, extreme focus on gaming (predicting both mental and behavioral focus on gaming), lying about gaming, escalated use, and harmful use. It is notable that the last two constructs fit the quantifier suggested for severity in the DSM-5 phrasing: escalated use is associated with ‘more hours spent’ and harmful use is associated with ‘more severe loss of opportunities’.

IGD as a (mostly) formative construct

Using a structured approach,⁴⁰ we demonstrated that the IGD proposal might be best interpreted as a predominantly formative construct, as opposed to the current practice of conceptualizing it as a reflective construct. It is known that incorrectly approaching a formative construct as a reflective one causes serious problems in scale development. These problems include (a) incorrect reliance on item-to-total scale correlation to exclude items and incorrect reliance on indices of inter-item reliability that do not fit the measurement model (e.g., Cronbach’s α), (b) incorrect interpretation of composite or mean scores that assume all items are equal in contributing value to a sum score, and (c) biased estimation of parameters in statistical models. These issues impact survey research most directly, although scale validation problem such as dropping conceptually sound items through reliance on item-to-total-correlations also negatively impact clinical screening instruments. Figure 2 provides a suggestion of a formative interpretation of IGD. In Figure 2, the focal construct of IGD is

Figure 2: Hypothesized dimensionality and construct relationships in IGD



formatively composed out of five sub-constructs. Two of these sub-constructs have a formative nature themselves (escalated use and harmful use). One sub-construct (extreme focus on gaming) has a reflective nature. Obviously alternative formative formulations might be possible too, such as viewing each of the nine criteria as representing nine sub-constructs that formatively compose the first-order focal construct of IGD.

Implications for future research

Based on the analysis conducted, it appears that a (mostly) formative process of validation appears justified in the case of IGD. This has implications with regard to some key aspects of IGD-related scale validation. While a comprehensive review of the suggested approach to formative scale validation can be found elsewhere,^{36,40} it might be helpful here to discuss some important differences and contrast them with current practices in the field of IGD scale development.

1. Composite latent constructs should be defined and used in statistical models.

Confirmatory and exploratory statistical model building should define a composite construct. This manifests itself in a different specification of models and error terms in the models. For instance, measures (items) are not expected to individually have error terms. Rather, measurement error is associated with the construct as a whole (residual), and all variation in the individual measures is relevant to the formative construct.³⁸

The earlier examples of the price index and the IQ test might, again, be helpful for understanding this point. In the hypothetical IQ example, a scale with three measures was proposed: a visual puzzle, a maze, and a math problem. These measures were supposed to covary due to an underlying cause, *i.e.* ‘intelligence’. According to a reflective measurement approach, it is assumed that the score on any of the three measures contains two components: ‘intelligence’ and measurement error. The total scale is useful if the measurement error is quite low, and any observed variation in scores on the maze, puzzle, and math problem is

associated with the common variation over all three tests. In essence, the scale-score involves the overlapping ground between the scores on the three individual measures. If this overlap is small, and the error large, then the test is not a good one.

With the formative example construct, the price index, this approach is not very helpful. The prices of bread, smartphones, and cinema tickets might not share much common ground at all. Thus, it also becomes irrelevant to split the error from the ‘price index’ component for individual measures. Instead, the error is sought on the level of the construct itself, *i.e.* the price index. Bread and smartphones and cinema tickets might not fully and comprehensively cover all aspects of pricing. We might be missing the price of television sets, or healthcare, for example. In models, this missing information is included as a ‘residual’ of the construct. This residual is used as an indicator of scale quality. For instance, it is very ambiguous what is actually measured once the residual starts to explain more than half of (for example) the price index.⁴⁰

The synthesis provided in table 1 illustrates how the majority of existing validation articles have defined confirmatory models with individual error terms, which can, in light with the arguments described earlier, be viewed as incorrect. Violating formative measurement model assumptions might lead to, for example, inflated estimates of the relationship that IGD has with other constructs of interest, such as depressive mood, loneliness, parenting behavior, etcetera.³⁶

2. Item exclusion and selection should NOT rely on item-to-total scale correlations.

Inter-item correlation is possible, but not required, in a formative formulation. Thus, it should not be used as grounds for justifying scale modification or item drops. This problem is further compounded by the fact that entire concepts might drop from the construct through exclusive reliance on these statistical indicators. With all due respect to authors and the

amount of work involved in producing their contributions, we shall now re-interpret the validation methods from two studies:

Example 1. Development of the IGDS-9

Lemmens et al. drafted three items³² for each of the IGD proposal criteria, generating a pool of 27 items in total. They specifically refer to the value of breaking items into separate ‘discrete components’ to assess different ‘aspects of the DSM-criteria (e.g., relationship, job, education)’. This formulation seems formative in nature, but the authors proceed using a reflective analysis. IGD is interpreted as a reflective focal construct with 9 reflective sub-constructs, involving each three items with associated measurement errors. In order to create the short 9-item form, ‘the highest loading item from each criterion was selected to create nine-item versions of the scales that encompassed all criteria’ (p 8). There is some clear tension in this sentence and this approach: Maintaining all of the criteria implies formative thinking, but the analysis method and process of selection are based on a reflective approach, which by definition would result in mutually interchangeable items. This confusion creates various problems and some negative consequences of this confusion manifest in the process of item selection. Here, conceptual richness is lost due to choice based on reflective model reasoning. For instance, the continued use of Internet games despite psychosocial problems was translated into three items:

- have you skipped work or school so that you could play games?
- have you played throughout the night, or almost the whole night?
- have you had arguments with others about the consequences of your gaming behavior?

Based on factor loading the last item was suggested for use in the summarized 9-item scale. In a formative interpretation (which we have argued to be correct), this approach is problematic as it excludes the entire conceptual area of problems at school/work and playing all night through inappropriate statistical assumptions of item equality.

Example 2: Development of the IGDS9-SF

Pontes et al.³⁰ ran an exploratory factor analysis on half of their sample and found that the 9 items in their IGDS9-SF are best summarized by a single factor, which explains 45.4% of the total variance. This one-factor solution was confirmed through a confirmatory factor analysis in the other half of the sample. Both techniques rely on inter-item association to explain total variance and their graphical summary also confirms that a reflective approach is assumed (e.g., items have individual error terms). Statistically, two indicators support that this reflective approach might not be optimal. Firstly, the factor loadings on the items diverge quite widely (from .54 on pre-occupation with gaming to .78 on the item tapping irritability, anxiety, or sadness when attempting to reduce or stop gaming behavior). This is not expected behavior if items are mutually interchangeable. Secondly, the common factor explains less than 50% of the total variance.

The authors further state that *'the instrument was highly reliable across the samples since the Cronbach's alphas were very high and not possible to be increased by deleting any of the nine items of the scale'* (p 141). While the authors ultimately did not delete any items, this phrasing implies a reflective interpretation of the scale. Assuming a formative model, removing any of the conceptually rich items (e.g., the loss of control over gaming, or the experience of negative outcomes in general), is not warranted and the item variance that is now discounted as measurement error should be taken into account. In fact, applying this type of rationale regarding item contributions in a reflective interpretation consistently, one could conceivably end up with a diagnosis that involves only items that measure enthusiasm about gaming and not a clinical disorder *per se*.

The reflective model also manifests in the extrapolation of cut-off scores: the 5 out of 9 threshold proposed in the DSM-5 proposal becomes 36 out of 45 points based on a 5-point scale (ranging from "never" to "very often") with the IGDS9-SF. This approach implies

equivalence between measurement items, as they all contribute equally to the final score, with ‘higher scores being indicative of higher degrees of gaming disorder’ (p. 138). Such a statement also mismatches with the assumption of a formative model, in which scores are not necessarily cumulative and items are not equivalent.

3. *Existing definitions of IGD should be enriched.*

For a formative construct, the measures chosen (and their weights) determine the nature of the construct, and thus it is even more important to cover the intended construct completely and with items that have proper content validity.⁴⁶ The validation of IGD is an ongoing process, but the existing proposal advanced in the DSM-5 has the risk of locking thinking into a confirmatory approach.⁶ We have evaluated the nature of the existing conceptual definition and the existing proposed sub-constructs in what is essentially only the first step in a trajectory of proper scale validation.⁴⁰ An essential and complementary approach in future research would be to critically expand this first step, employing a more open perspective. This will allow the addressing of other key issues such as clarifying the degree to which the current IGD proposal is able to capture the clinical profile and explain the functional impairment of treatment-seeking individuals displaying disordered online gaming.

As illustrations, recent work conducted in clinical settings highlighted that other criteria could be relevant to diagnose IGD, such as craving (or urge to play video games), irritability, or poor hygiene.^{27,33} Irritability or poor hygiene might be interpreted as an additional (formative) manifestation of problems, and craving could be viewed as part of the *extreme focus on gaming* sub-construct defined earlier. Some specific criteria, such as lying about gaming or playing to modify one’s mood, appear to be unnecessary to account for the severity of IGD in the current proposal. It should also be studied if these criteria are relevant and specific enough to be included in the core IGD definition. For instance, lying about use

displayed a sensitivity of only 44% in one study.²⁷ Another study revealed a sub-group of people that rely on gaming and other enjoyable behaviors (or substances) to regulate their mood state in a non-addictive but potentially risky manner⁴⁷ and studies also highlight the role of emotion in media selection in general.⁴⁸

Aspects unique to (online) gaming might also be added to the diagnosis over the course of time, similarly to the developments that occurred with regard to gambling disorders (e.g., chasing losses). These unique aspects might involve, for example, the social nature of online gaming or game mechanics / design aspects that are related to the behavior. Any existing and proposed new (sub)constructs or measures should also be evaluated for possible spurious determination by underlying other causes. This is a particular issue for IGD, as negative attitudes about games are prevalent in many Western societies.⁴⁹ Some criteria might be spurious and be determined mostly by an underlying construct that influences both IGD and the specific criterion, e.g., highly engaged gaming.⁵⁰

Interpreting IGD as a formative construct might at first seem a primarily psychometric matter, but the change has clinical implications as well. For instance, a reflective interpretation of IGD implies a realist perspective involving that an underlying ‘disorder’ caused these symptoms and problematic behaviors. In contrast, a formative interpretation allows the broader interpretation that a cluster of observed behaviors and psychological issues might be pragmatically summarized under the header of ‘Internet Gaming Disorder’. Clinicians would benefit from measurement instruments that include a comprehensive construct for IGD, preferably a construct that does not exclude conceptually relevant criteria or overemphasize criteria that are better used to assess enthusiast or highly engaged gaming (i.e., a harmonious, non-pathological involvement).⁵⁰ In this sense, it seems the IGD disorder proposal has great room for improvement.

Acknowledgments

We thank Dr. Michelle Colder Carras for proofreading this manuscript and Halley Pontes for providing feedback on an earlier version of this manuscript.

References

1. Durkee T, Kaess M, Carli V, Parzer P, Wasserman C, Floderus B, et al. Prevalence of pathological internet use among adolescents in Europe: demographic and social factors. *Addiction*. 2012;107(12):2210–22.
2. Haagsma MC, Pieterse ME, Peters O. The Prevalence of Problematic Video Gamers in The Netherlands. *Cyberpsychology, Behav Soc Netw*. 2012;15(3):162–8.
3. Wisselink DJJ, Kuijpers WGT, Mol A. Key figures addiction care 2014. Houten: Informatie Voorziening Zorg (IVZ); 2015.
4. Müller KW, Beutel ME, Egloff B, Wölfling K. Investigating Risk Factors for Internet Gaming Disorder: A Comparison of Patients with Addictive Gaming, Pathological Gamblers and Healthy Controls regarding the Big Five Personality Traits. *Eur Addict Res*. 2014;20(3):129–36.
5. Kardefelt-Winther D. A critical account of DSM-5 criteria for internet gaming disorder. *Addict Res Theory*. 2015;23(2):93–8.
6. Billieux J, Schimmenti A, Khazaal Y, Maurage P, Heeren A. Are we overpathologizing everyday life? A tenable blueprint for behavioral addiction research. *J Behav Addict*. 2015;4(3):119–23.
7. Wood RTA. Problems with the concept of video game “addiction”: Some case study examples. *Int J Ment Health Addict*. 2008;6(2):169–78.
8. American Psychiatric Association. Internet Use Disorder [Internet]. DSM5.org. 2012. Available from: <https://web.archive.org/web/20120518213016/http://www.dsm5.org/ProposedRevision/Pages/proposedrevision.aspx?rid=573>. Accessed: 2015-10-14. (Archived by WebCite? at <http://www.webcitation.org/6cH82dDIL>)
9. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition: DSM-5*. Fifth. Arlington, VA: American Psychiatric Association; 2013.
10. Petry NM, O’Brien CP. Internet gaming disorder and the DSM-5. *Addiction*. 2013;108(7):1186–7.
11. Thorens G, Achab S, Billieux J, Khazaal Y, Khan R, Pivin E, et al. Characteristics and treatment response of self-identified problematic Internet users in a behavioral addiction outpatient clinic. *J Behav Addict*. 2014;3(1):78–81.
12. Griffiths MD, van Rooij AJ, Kardefelt-Winther D, Starcevic V, Király O, Pallesen S, et al. Working towards an international consensus on criteria for assessing internet gaming disorder: A critical commentary on Petry et al. (2014). *Addiction*. 2016;111(1):167–75.
13. Petry NM, Rehbein F, Gentile D a., Lemmens JS, Rumpf H-J, Mößle T, et al. An

- international consensus for assessing internet gaming disorder using the new DSM-5 approach. *Addiction*. 2014;109(9):1399–406.
14. Petry NM, Rehbein F, Gentile DA, Lemmens JS, Rumpf H-J, Möble T, et al. Griffiths et al.'s comments on the international consensus statement of internet gaming disorder: furthering consensus or hindering progress? *Addiction*. 2015;111:167–78.
 15. Van Rooij AJ, Prause N. A critical review of “Internet addiction” criteria with suggestions for the future. *J Behav Addict*. 2014;3(4):203–13.
 16. King DL, Haagsma MC, Delfabbro PH, Gradisar M, Griffiths MD. Toward a consensus definition of pathological video-gaming: A systematic review of psychometric assessment tools. *Clin Psychol Rev*. 2013;33(3):331–42.
 17. Tao R, Huang X, Wang J, Zhang H, Zhang Y, Li M. Proposed diagnostic criteria for internet addiction. *Addiction*. 2010;105(3):556–64.
 18. Griffiths MD. A “components” model of addiction within a biopsychosocial framework. *J Subst Use*. 2005;10(4):191–7.
 19. Young KS. Internet Addiction: The emergence of a new clinical disorder. *Cyberpsychology, Behav Soc Netw*. 1998;1(3):237–44.
 20. Kuss DJ, Griffiths MD. Internet gaming addiction: A systematic review of empirical research. *Int J Ment Health Addict*. 2012;10(2):278–96.
 21. Gentile DA, Choo H, Liau A, Sim T, Li D, Fung D, et al. Pathological Video Game Use Among Youths: A Two-Year Longitudinal Study. *Pediatrics*. 2011;127(2):e319–29.
 22. Lemmens JS, Valkenburg PM, Peter J. Development and Validation of a Game Addiction Scale for Adolescents. *Media Psychol*. 2009;12(1):77–95.
 23. Van Rooij AJ, Schoenmakers TM, Van den Eijnden RJJM, Vermulst AA, van de Mheen D. Video game addiction test: validity and psychometric characteristics. *Cyberpsychology, Behav Soc Netw*. 2012;15(9):507–11.
 24. Guertler D, Rumpf HJ, Bischof A, Kastirke N, Petersen KU, John U, et al. Assessment of problematic internet use by the compulsive internet use scale and the internet addiction test: A sample of problematic and pathological gamblers. *Eur Addict Res*. 2014;20(2):75–81.
 25. Lortie CL, Guitton MJ. Internet addiction assessment tools: dimensional structure and methodological status. *Addiction*. 2013;108(7):1207–16.
 26. Király O, Slezcka P, Pontes HM, Urbán R, Griffiths MD, Demetrovics Z. Validation of the Ten-Item Internet Gaming Disorder Test (IGDT-10) and evaluation of the nine DSM-5 Internet Gaming Disorder criteria. *Addict Behav*. 2015;
 27. Ko CH, Yen JY, Chen SH, Wang PW, Chen CS, Yen CF. Evaluation of the diagnostic

- criteria of Internet gaming disorder in the DSM-5 among young adults in Taiwan. *J Psychiatr Res.* 2014;53(1):103–10.
28. Rehbein F, Kliem S, Baier D, Mößle T, Petry NM. Prevalence of Internet Gaming Disorder in German adolescents: diagnostic contribution of the nine DSM-5 criteria in a statewide representative sample. *Addiction.* 2015;110(5):842–51.
 29. Fuster H, Carbonell X, Pontes HM, Griffiths MD. Spanish validation of the Internet Gaming Disorder-20 (IGD-20) Test. *Comput Human Behav.* 2016;56:215–24.
 30. Pontes HM, Griffiths MD. Measuring DSM-5 internet gaming disorder: Development and validation of a short psychometric scale. *Comput Human Behav.* 2015;45:137–43.
 31. Pontes HM, Király O, Demetrovics Z, Griffiths MD. The Conceptualisation and Measurement of DSM-5 Internet Gaming Disorder: The Development of the IGD-20 Test. *PLoS One.* 2014;9(10):e110137.
 32. Lemmens JS, Valkenburg PM, Gentile DA. The internet gaming disorder scale. *Psychol Assess.* 2015;27(2):567–82.
 33. Van Rooij AJ, Schoenmakers TM, Van de Mheen D. Clinical validation of the C-VAT 2.0 assessment tool for gaming disorder: A sensitivity analysis of the proposed DSM-5 criteria and the clinical characteristics of young patients with “video game addiction.” *Addict Behav.* 2015;
 34. Cho H, Kwon M, Choi J-H, Lee S-K, Choi JS, Choi S-W, et al. Development of the Internet addiction scale based on the Internet Gaming Disorder criteria suggested in DSM-5. *Addict Behav.* 2014;39(9):1361–6.
 35. Edwards JR, Bagozzi RP. On the nature and direction of relationships between constructs and measures. *Psychol Methods.* 2000;5(2):155–74.
 36. MacKenzie SB, Podsakoff PM, Jarvis CB. The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions. *J Appl Psychol.* 2005;90(4):710–30.
 37. Borsboom D. The attack of the psychometricians. *Psychometrika.* 2006;71(3):425–40.
 38. Bollen K, Lennox R. Conventional wisdom on measurement: A structural equation perspective. *Psychol Bull.* 1991;110(2):305–14.
 39. Howell RD, Breivik E, Wilcox JB. Is formative measurement really measurement? Reply to Bollen (2007) and Bagozzi (2007). *Psychol Methods.* 2007;12(2):238–45.
 40. Mackenzie SB, Podsakoff PM, Podsakoff NP. Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques. *MIS Q.* 2011;35(2):293–334.
 41. Diamantopoulos A, Winklhofer HM. Index Construction with Formative Indicators: An Alternative to Scale Development. *J Mark Res.* 2001;38(2):269–77.

42. Diamantopoulos A, Riefler P, Roth KP. Advancing formative measurement models. *J Bus Res.* 2008;61(12):1203–18.
43. MacKenzie SB. The Dangers of Poor Construct Conceptualization. *J Acad Mark Sci.* 2003;31(3):323–6.
44. Jarvis CB, MacKenzie SB, Podsakoff PM. A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research. *J Consum Res.* 2003;30(2):199–218.
45. Schellinck T, Schrans T, Schellinck H, Bliemel M. Construct Development for the FocaL Adult Gambling Screen (FLAGS): A Risk Measurement for Gambling Harm and Problem Gambling Associated with Electronic Gambling Machines. *J Gambl Issues.* 2015;30(30):140–73.
46. Petter S, Straub D, Rai A. Specifying Formative Constructs in Information Systems Research. *MIS Q.* 2015;31(4):623–56.
47. Deleuze J, Rochat L, Romo L, Van der Linden M, Achab S, Thorens G, et al. Prevalence and characteristics of addictive behaviors in a community sample: A latent class analysis. *Addict Behav Reports.* 2015;1:49–56.
48. Bartsch A, Vorderer P, Mangold R, Viehoff R. Appraisal of Emotions in Media Use: Toward a Process Model of Meta-Emotion and Emotion Regulation. *Media Psychol.* 2008;11(1):7–27.
49. Bourgonjon J, Valcke M, Soetaert R, De Wever B, Schellens T. Parental acceptance of digital game-based learning. *Comput Educ.* 2011;57(1):1434–44.
50. Charlton JP, Danforth IDW. Distinguishing addiction and high engagement in the context of online game playing. *Comput Human Behav.* 2007;23(3):1531–48.