

Expert appraisal of criteria for assessing gaming disorder: an international Delphi study

Jesús Castro-Calvo¹, Daniel L. King² , Dan J. Stein³, Matthias Brand⁴, Lior Carmi⁵, Samuel R. Chamberlain^{6,7}, Zsolt Demetrovics⁸, Naomi A. Fineberg^{9,10}, Hans-Jürgen Rumpf¹¹, Murat Yücel¹² , Sophia Achab^{13,14}, Atul Ambekar¹⁵ , Norharlina Bahar¹⁶, Alexander Blaszczynski¹⁷, Henrietta Bowden-Jones^{18,19}, Xavier Carbonell²⁰ , Elda Mei Lo Chan²¹, Chih-Hung Ko²², Philippe deTimary²³, Magali Dufour²⁴, Marie Grall-Bronnec^{25,26} , Hae Kook Lee²⁷, Susumu Higuchi²⁸ , Susana Jimenez-Murcia^{29,30} , Orsolya Király⁸, Daria J. Kuss³¹, Jiang Long^{32,33}, Astrid Müller³⁴, Stefano Pallanti³⁵, Marc N. Potenza³⁶, Afarin Rahimi-Movaghar³⁷, John B. Saunders³⁸, Adriano Schimmenti³⁹, Seung-Yup Lee⁴⁰ , Kristiana Siste^{41,42}, Daniel T. Spritzer⁴³, Vladan Starcevic⁴⁴, Aviv M. Weinstein⁴⁵ , Klaus Wölfling⁴⁶ & Joël Billieux^{47,48} 

Department of Personality, Assessment, and Psychological Treatments, University of Valencia, Spain,¹ College of Education, Psychology, and Social Work, Flinders University, Australia,² SAMRC Unit on Risk and Resilience in Mental Disorders, Department of Psychiatry and Neuroscience Institute, University of Cape Town, Cape Town, South Africa,³ General Psychology: Cognition and Center for Behavioral Addiction Research (CeBAR), University Duisburg-Essen, Germany,⁴ The Data Science Institute, Inter-disciplinary Center, Herzliya, Israel,⁵ Department of Psychiatry, Faculty of Medicine, University of Southampton, Southampton, UK,⁶ Southern Health NHS Foundation Trust, Southampton, UK,⁷ Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary,⁸ University of Hertfordshire, Hatfield, UK, Hertfordshire Partnership University NHS Foundation Trust, Welwyn Garden City, UK,⁹ University of Cambridge School of Clinical Medicine, Cambridge, UK,¹⁰ Department of Psychiatry and Psychotherapy, University of Luebeck, Luebeck, Germany,¹¹ BrainPark, School of Psychological Sciences, Turner Institute for Brain and Mental Health and Monash Biomedical Imaging Facility, Monash University, Melbourne, Victoria, Australia,¹² Specialized Facility In Behavioral Addictions, ReConnecte, Department of Psychiatry, University Hospitals of Geneva, Geneva, Switzerland,¹³ Faculty of Medicine, Geneva University, Geneva, Switzerland,¹⁴ National Drug Dependence Treatment Centre, All India Institute of Medical Sciences, New Delhi, India,¹⁵ Department of Psychiatry and Mental Health, Hospital Selayang, Ministry of Health, Malaysia,¹⁶ Faculty of Science, Brain and Mind Centre, School of Psychology, University of Sydney, Sydney, Australia,¹⁷ National Centre for Gaming Disorders, London, UK,¹⁸ University College London, London, UK,¹⁹ Faculty of Psychology, Education and Sports Sciences Blanquerna, Ramon Llull University, Barcelona, Spain,²⁰ St John's Cathedral Counselling Service, and Division on Addiction, Hong Kong,²¹ Department of Psychiatry, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung City, Taiwan,²² Department of Adult Psychiatry, Institute of Neuroscience, UCLouvain and Cliniques Universitaires Saint-Luc, Brussels, Belgium,²³ Université du Québec à Montréal, Montréal, Québec, Canada,²⁴ CHU Nantes, Department of Addictology and Psychiatry, Nantes, France,²⁵ Universités de Nantes et Tours, UMR 1246, Nantes, France,²⁶ Department of Psychiatry, College of Medicine, The Catholic University of Korea, Seoul, South Korea,²⁷ National Hospital Organization, Kurhama Medical and Addiction Center, Japan,²⁸ Department of Psychiatry, Bellvitge University Hospital-IDIBELL, Barcelona, Spain,²⁹ Ciber Fisiopatología Obesidad y Nutrición (CIBEROBn), Instituto de Salud Carlos III, Madrid, Spain,³⁰ International Gaming Research Unit, Psychology Department, Nottingham Trent University, Nottingham, UK,³¹ Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China,³² Laboratory for Experimental Psychopathology, Psychological Science Research Institute, Université Catholique de Louvain, Louvain, Belgium,³³ Department of Psychosomatic Medicine and Psychotherapy, Hannover Medical School, Hanover, Germany,³⁴ Neuroscience Institute, University of Florence, Italy,³⁵ Departments of Psychiatry and Neuroscience and the Child Study Center, Yale School of Medicine and Connecticut Mental Health Center, New Haven, CT, USA,³⁶ Iranian National Center for Addiction Studies, Tehran University of Medical Sciences, Tehran, Iran,³⁷ Centre for Youth Substance Abuse Research, University of Queensland, Brisbane, Queensland, Australia,³⁸ Faculty of Human and Social Sciences, Kore University of Enna, Enna, Italy,³⁹ Department of Psychiatry, Eunpyeong St Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, South Korea,⁴⁰ Department of Psychiatry, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia,⁴¹ Cipto Mangunkusumo Hospital, Jakarta, Indonesia,⁴² Postgraduate Program in Psychiatry and Behavioral Sciences, Federal University of Rio Grande do Sul, Brazil,⁴³ Faculty of Medicine and Health, Sydney Medical School, Nepean Clinical School, University of Sydney, Sydney, Australia,⁴⁴ Department of Behavioral Science, Ariel University, Israel,⁴⁵ Outpatient Clinic for Behavioral Addictions, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center of the Johannes Gutenberg-University Mainz, Germany,⁴⁶ Institute of Psychology, University of Lausanne, Lausanne, Switzerland,⁴⁷ and Health and Behaviour Institute, University of Luxembourg, Esch-sur-Alzette, Luxembourg,⁴⁸

ABSTRACT

Background and aims Following the recognition of 'internet gaming disorder' (IGD) as a condition requiring further study by the DSM-5, 'gaming disorder' (GD) was officially included as a diagnostic entity by the World Health Organization (WHO) in the 11th revision of the International Classification of Diseases (ICD-11). However, the proposed diagnostic criteria for gaming disorder remain the subject of debate, and there has been no systematic attempt to integrate the views of different groups of experts. To achieve a more systematic agreement on this new disorder, this study employed the Delphi expert consensus method to obtain expert agreement on the diagnostic validity, clinical utility and prognostic value of the DSM-5 criteria and ICD-11 clinical guidelines for GD. **Methods** A total of 29 international experts with clinical and/or research experience in GD completed three iterative rounds of a Delphi survey. Experts rated proposed criteria in progressive rounds until a pre-determined level of agreement was achieved. **Results** For DSM-5 IGD criteria, there was an agreement

both that a subset had high diagnostic validity, clinical utility and prognostic value and that some (e.g. tolerance, deception) had low diagnostic validity, clinical utility and prognostic value. Crucially, some DSM-5 criteria (e.g. escapism/mood regulation, tolerance) were regarded as incapable of distinguishing between problematic and non-problematic gaming. In contrast, ICD-11 diagnostic guidelines for GD (except for the criterion relating to diminished non-gaming interests) were judged as presenting high diagnostic validity, clinical utility and prognostic value. **Conclusions** This Delphi survey provides a foundation for identifying the most diagnostically valid and clinically useful criteria for GD. There was expert agreement that some DSM-5 criteria were not clinically relevant and may pathologize non-problematic patterns of gaming, whereas ICD-11 diagnostic guidelines are likely to diagnose GD adequately and avoid pathologizing.

Keywords Delphi, diagnosis, DSM, gaming disorder, ICD, internet gaming disorder.

Correspondence to: Joël Billieux, Institute of Psychology, University of Lausanne, Quartier UNIL-Mouline – Bâtiment Géopolis, CH-1015 Lausanne, Switzerland. E-mail: joel.billieux@unil.ch

Submitted 26 May 2020; initial review completed 23 September 2020; final version accepted 6 January 2021

INTRODUCTION

Video games are one of the most popular leisure activities world-wide. Alongside the many technological developments that have made gaming increasingly accessible, portable and immersive, there has also been a massive world-wide growth in the popularity of e-sports and on-line gaming-related activities such as streaming services. Although the majority of players enjoy gaming as a recreational activity, some individuals report poorly controlled and excessive gaming that has negative psychosocial consequences [1–3]. Following the inclusion of ‘internet gaming disorder’ (IGD) as a disorder requiring further study in the DSM-5 [4], ‘gaming disorder’ (GD) was officially included by the World Health Organization (WHO) as a diagnostic entity in the 11th revision of the International Classification of Diseases (ICD-11) [5]. Despite the growing acceptance of gaming-related harms as an important public health issue [2,6,7], the precise clinical formulation of gaming as the foundation of an addictive disorder, including conceptual overlap with substance-based addictions, continues to be debated [8,9]. Some claims have also been made that recognizing GD as a mental condition may pathologize healthy gaming patterns [8,10]. Thus, the aim of this study was to systematically develop an international, expert-based agreement regarding the core diagnostic features of GD.

Gaming disorder in the DSM-5 and the ICD-11

In preparation for DSM-5, the American Psychiatric Association’s (APA) work group on substance use and related disorders evaluated available research literature on the relationship between excessive video gaming and a wide range of problems (e.g. psychopathological symptoms, addictive usage patterns, health issues, reduced school grades). This review led to the inclusion of IGD in Section III of the DSM-5 [4]. Section III includes tentative disorders

for which the evidence base is not yet deemed sufficient for formal recognition. Based on accumulated scientific evidence (e.g. epidemiological, psychometric, psychological, neurobiological), in addition to substantial data obtained from treatment providers, GD and ‘hazardous gaming’ were included in the ICD-11 [5]. However, despite this formal recognition, the proposed diagnostic criteria for GD remain the subject of debate. A key criticism of including GD in the nosology has been that such a diagnosis may pathologize healthy gaming and may lead to unnecessary policies and interventions [8,9]. Another criticism has been that the criteria generally used to screen for problematic gaming have sometimes failed to distinguish between patterns of highly engaged and problematic gaming patterns [11–13]. An additional general concern has been that IGD criteria were based on those for substance-use and gambling disorders [11,14]. More precisely, the addictive aetiology of GD was disputed and some authors have indicated that alternative models (e.g. compensatory mechanisms) may have been overlooked [15–17].

There are some important differences between the DSM-5 IGD diagnostic criteria and the ICD-11 diagnostic guidelines for GD that further indicate a lack of expert agreement. In the DSM-5, IGD refers to a gaming pattern that results in significant impairment or distress and is characterized by meeting at least five of nine criteria; namely, pre-occupation, withdrawal, tolerance, impaired control, diminished non-gaming interests, gaming despite harms, deception about gaming, gaming for escape or regulating mood and conflict/interference due to gaming [18] (see Table 2). The inclusion of IGD in the DSM-5 paved the way for promoting research on gaming as a disorder, including the first epidemiological studies using nationally or state-representative samples [19,20]. However, IGD criteria have not been well investigated in treatment-seeking settings. The few available studies of a structured interview approach to assess IGD criteria in clinical samples indicate satisfactory diagnostic validity [21,22], but these studies

also report that the IGD criteria of escapism/mood regulation and deception of others may have lower diagnostic validity (diagnostic accuracy < 80%). Another issue related to IGD criteria pertains to the validity of tolerance and withdrawal constructs: it has been argued that these addiction symptoms (based on the criteria from substance-use or gambling disorders) may fail to distinguish between high but unproblematic involvement (i.e. a 'gaming passion') and problematic gaming [23–26].

The ICD-11 has adopted a more concise set of guidelines for GD compared to the DSM-5. GD is defined as a pattern of gaming involving (1) impaired control; (2) increasing priority given to gaming to the extent that gaming takes precedence over other life interests and daily activities; and (3) continuation or escalation of gaming despite the occurrence of negative consequences (see Table 3). In addition, the gaming pattern must be associated with distress or significant impairment in personal, family, social and/or other important areas of functioning. Compared to the DSM-5 IGD approach, the ICD-11 framework applies a monothetic approach (i.e. all criteria/diagnostic guidelines must be endorsed) rather than a polythetic one (i.e. the five of nine cut-off proposed for IGD in the DSM-5). Notably, the ICD-11 guidelines have eschewed the tolerance and withdrawal criteria. Research on the clinical utility of the ICD-11 guidelines for GD is beginning to emerge in the literature. A recent multi-centre cohort study by Jo and colleagues [27] conducted interviews of high-risk adolescents in Korea to undertake a comparative analysis of the ICD-11 and DSM-5 criteria. This study reported that 32.4% met the DSM-5 criteria for IGD, whereas only 6.4% met the more stringent ICD-11 criteria. Although empirical evidence regarding the diagnostic validity of ICD-11 criteria is not yet available, these preliminary data suggest that the conservative approach of the ICD-11 may better avoid false positives in screening [1,26].

The lack of expert agreement on GD diagnostic criteria may have negative implications for clinical assessment and for policy development. Various techniques can be employed to reach agreement about a given debated topic, but among the most rigorous is the Delphi technique [28], in which a panel of experts rate proposed criteria in progressive (iterative) rounds until a pre-determined level of agreement is achieved. The Delphi technique is a method for systematically gathering data from expert respondents, for assessing the extent of agreement and for resolving disagreement [29,30]. This technique has been widely used in the field of mental health [31]. In a review of 176 studies [31], four types of expert agreement can be reached using the Delphi technique: (a) making estimations where there is incomplete evidence to suggest more accurate answers; (b) making predictions about a topic of particular interest; (c) determining collective values among experts in a field; or (d) defining basic concepts in research and clinical

practice. In the field of addictive disorders the Delphi technique has, for example, been used in order to identify the 'primary' research domain criteria (RDoC) constructs more relevant to substance and behavioural addictions [32], or the central features of addictive involvement in physical exercise [33].

There are several specific advantages of the Delphi method. First, it allows for anonymity between participants and controlled feedback provided in a structured manner through a succession of iterative rounds of opinion collection [29,34]. Secondly, as the process is undertaken anonymously and all experts' opinions are equally weighted, such an approach avoids the dominance of a few experts or a reduced pressure group [35]. Other advantages of anonymity include the fact that: (a) panellists do not have socio-psychological pressure to select a particular answer or rate an item in a certain way, (b) it avoids the unwillingness to abandon publicly expressed opinions and (c) it leads to higher response rates [36]. Furthermore, the fact that controlled and individualized feedback is provided through a series of successive rounds facilitates the process of change of mind usually required to achieve an agreement in groups with divided opinions [34]. As a result, the Delphi technique facilitates collective intelligence or 'wisdom of crowds' (i.e. the combined ability of a group of experts to jointly produce better results than those produced by each expert on his or her own) [37–39]. For all these reasons, the Delphi technique constitutes an ideal method in the context of a study that aims to approach expert agreement regarding the most appropriate diagnostic criteria for GD.

The present study

The aim of this study was to use the Delphi technique to clarify and reach an expert agreement on which criteria could be used to define and diagnose problematic video gaming. This study involved a large international panel of recognized experts on gaming as a disorder in order to approach agreement concerning the diagnostic validity (defined here as the extent to which a specific criterion is a feature of the condition), clinical utility (defined here as the extent to which a specific criterion is able to distinguish normal from problematic behaviour) and prognostic value (defined here as the extent to which a specific criterion is crucial in predicting chronicity, persistence and relapse) of the nine DSM-5 criteria proposed to define IGD and the four ICD-11 clinical guidelines proposed to define GD. For the sake of parsimony, this study employed the term GD to encompass both the DSM-5 and ICD-11 classifications, and to refer generally to gaming as a disorder. Our approach constitutes an innovative contribution to the field as it is the first attempt, to our knowledge, to reach expert agreement through a systematic procedure (the Delphi technique) in

a field where opinions are often polarized, and where previous unstructured expert appraisal of GD criteria [40] have not been able to achieve expert agreement [12,41,42].

METHODS

Implementation of the Delphi technique

In brief, a Delphi study is managed by a facilitator (J.C.C. in the current Delphi study) responsible for the methodology planning, recruitment of experts and compilation of statements that the experts rate to assess expert agreement. At the beginning of the Delphi study, this facilitator gathers responses from the expert panel using a pre-designed questionnaire including the target statements (here, the DSM-5 and ICD-11 proposed diagnostic criteria for GD) and allowing, when relevant, the experts to add new statements to be rated by the group at the next round (here, potentially including relevant criteria not included in the DSM-5 or ICD-11). The Delphi technique capitalizes on multiple iterations to reach agreement between experts. Different versions of the initial questionnaire are rated in subsequent rounds in order to achieve a pre-established level of expert agreement. With each new round, the facilitator provides individualized anonymous feedback to the experts about how their answers align with the answers of the rest of the group. This feedback typically allows for the experts to engage in personal reflection and to potentially re-assess their initial view in light of the whole panel opinion. On the basis of this feedback, experts have the opportunity to revise their response (to align with other experts) or maintain their response (and explaining their rationale for doing so). The Delphi study is completed when a pre-established level of expert agreement is achieved and/or when responses between rounds are stable (i.e. when further significant changes in experts' opinion are not expected) [36]. Delphi studies are usually managed by email and implemented through on-line survey software [43], ensuring anonymity between participants and allowing involvement of international experts with different expertise [37].

Panel formation

The generalizability of the findings derived from a Delphi study depends primarily on the expert selection process [30] and the size and representativeness of the expert panel [44]. In the field of mental health research, members of expert panels are often professionals with extensive clinical and/or research experience concerning the target construct. The particular profile of the panellists (e.g. years of experience, field of expertise, etc.) is guided by the aims of the study [45]. For example, a previous mental-health-related Delphi study by Yap *et al.* [46] defined as expert those professionals with at least 5 years of experience

in research or clinical treatment on adolescent depression and anxiety, whereas Addington *et al.* [43] required that panellists had published at least one relevant publication as first or leading author in a peer-reviewed journal. In other studies, criteria were more restrictive: Yücel *et al.* [32], for example, required a minimum of 5 years of clinical experience and more than 50 articles authored in peer-reviewed journals to be included as expert in their Delphi. In a review of existing Delphi studies performed by Diamond *et al.* [29], 54% of the panels were made up of fewer than 25 experts and these produced stable results [47], similar to those obtained with more experts [44].

In the present Delphi study, experts were defined as professionals with clinical and/or research experience in problematic gaming. Clinical experience was assessed through years of reported experience assessing and treating patients with GD, whereas research experience was assessed through the number of GD-related papers published in peer-reviewed journals as first or last author. We prioritized the recruitment of experts with both clinical and research experience (a minimum of 1 year of clinical experience and at least two articles published in peer-reviewed journals). However, we also considered experts with experience in only one setting when they reported substantial clinical experience only (≥ 5 years) or high research achievement only (≥ 20 papers). The aim of these criteria was threefold: (1) to include different kinds of expertise, (2) to include different opinions (as the criteria used retained experts independently of their adherence or reluctance in including GD in nomenclature systems) and (3) to ensure geographic representativeness. In terms of sample size, we planned to recruit between 25 and 40 experts (which was judged by the research team as an ideal balance between efforts needed for panel management and stability of results). Geographical representativeness of experts may help to ensure generalizability of results; thus, we attempted to have a panel that was as internationally representative as possible.

The experts' selection process is depicted in Fig. 1. Experts were identified based on: (a) authorship of articles on GD, especially those reporting the assessment and treatment of patients qualifying for this clinical condition [27] or including a large number of researchers in the field [48]; (b) membership of editorial boards of relevant academic journals (e.g. *Addictive Behaviors* or *Journal of Behavioral Addictions*); (c) having been involved in ICD or DSM working groups related to addictive behaviours and/or obsessive-compulsive disorders; or (d) having participated in academic papers opposed to the recognition of problematic/pathological gaming as the foundation of a mental disorder [8]. Through these methods, a list of 200 world-wide experts on GD was generated. These 200 potential experts were then categorized and ranked according to their country of residence, research impact and clinical experience.

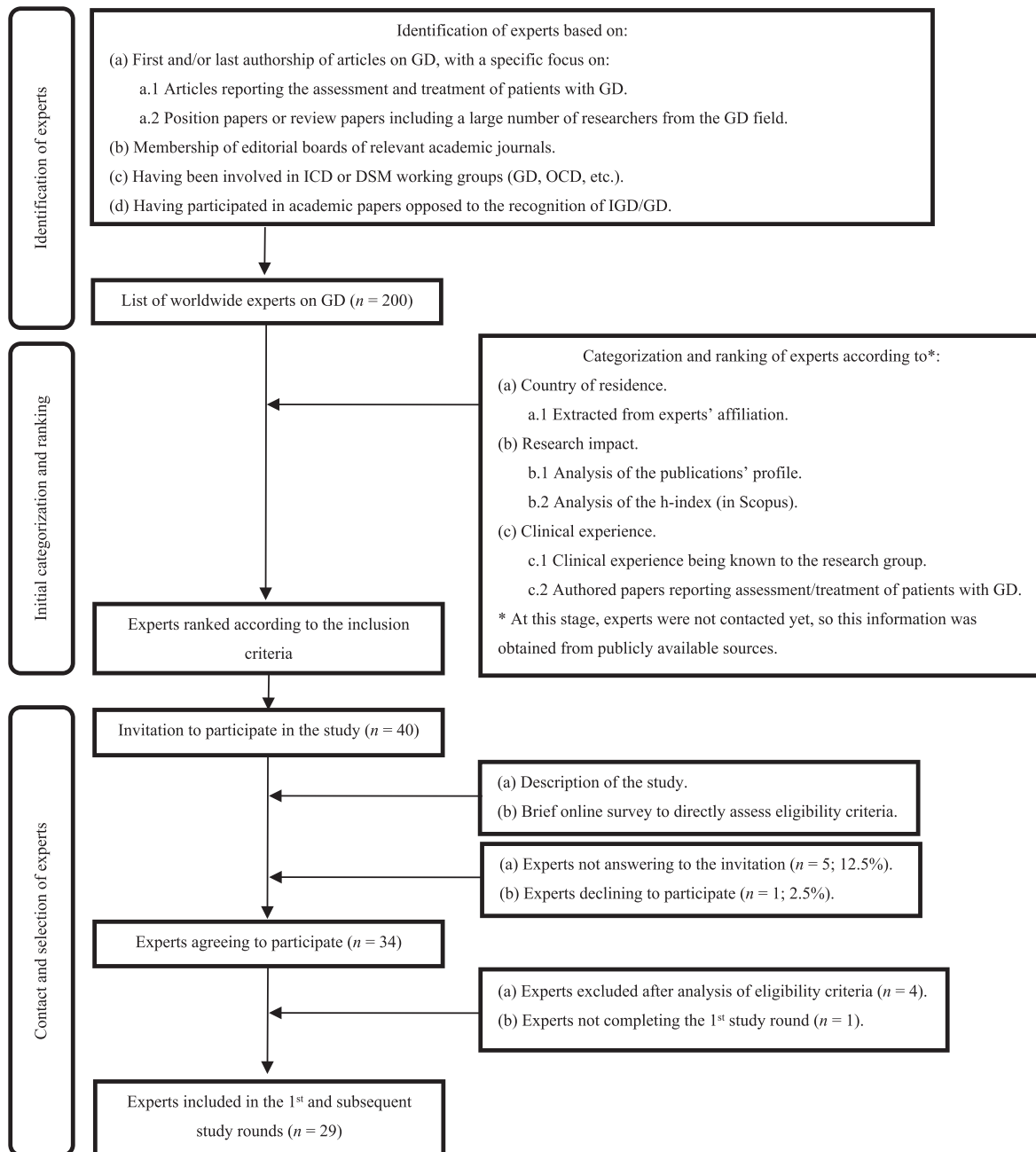


Figure 1 Experts' selection process.

Data were extracted from public sources: country of residence was extracted from experts' affiliation, research impact was assessed by analyzing their publications' profile and h-index (in Scopus) and clinical experience was assessed based on available information (clinical experience reported in publicly available CV, belonging to a specialized clinical programme treating GD patients) or having authored papers reporting the assessment or treatment of patients with GD. This preliminary classification was used to identify the experts that would be personally contacted by the research team. In order to ensure international representativeness, less stringent criteria for

research impact (i.e. < 20 research papers) were used to permit the inclusion of experts from under-represented countries. Eventually, 40 experts were selected and asked to participate in the study through personalized e-mail invitations. This contact e-mail included a description of the study and a link to a brief on-line survey. This survey included questions about clinical and research experience on GD in order to assess experts' eligibility. Of the initial 40 invitations, five experts (12.5%) did not answer the e-mail, one (2.5%) declined to participate and the remaining 34 (85%) agreed to participate and completed the eligibility survey. After analyzing eligibility criteria, four experts

were excluded because they did not match the requested criteria (they reported no clinical experience and published between two and five scientific papers each). Of those who met inclusion criteria, one did not complete the first study round and was excluded from the research. Thus, 29 experts met the criteria and were included in the expert panel.

Data collection and analysis

This Delphi study started with a closed-ended questionnaire incorporating the DSM-5 criteria for IGD and the ICD-11 clinical guidelines for GD. For each criterion, experts were asked to respond to three distinct questions: (a) to what extent do you feel that this criterion is important in the manifestation of treatment-seeking gaming/pathological gaming (diagnostic validity); (b) to what extent do you consider this criterion important as being able to distinguish normal from pathological videogame use (clinical utility); and (c) to what extent do you consider this criterion important to predict the chronicity—persistence and relapse—of pathological video game use (prognostic value)? Experts rated the importance of each criterion on a five-point Likert scale ranging from 1 (unimportant) to 5 (extremely important). As suggested by Yücel *et al.* [32], we avoided the use of a neutral mid-point in order to force panellists to give a deliberate response (3 was coded as ‘moderately important’). An additional response category (do not know) was included for those panellists who considered themselves as not having the required knowledge to respond to a particular statement. In the first study round, an open-ended question was included to give the experts the opportunity to propose additional diagnostic criteria not already included in the DSM-5 or the ICD-11 but which they considered relevant for the diagnosis of GD. Proposed criteria were reviewed by some of the authors who designed the study (J.C.C., J.B., D.L.K., M.B., S.R.C., Z.D., N.F., H.J.R. and M.Y.) in order to ascertain that their content was not already covered by the DSM-5 or the ICD-11 criteria; if so, the criteria were drafted and included in the second round of the Delphi.

After each round, panellists’ responses were screened to determine which criteria reached agreement for diagnostic validity, clinical utility and prognostic value. In accordance with current guidelines in Delphi research methods [32,49,50], criteria rated by $\geq 80\%$ of the experts as either ‘very important’ or ‘extremely important’ were considered to have reached agreement for inclusion and were not rated in subsequent Delphi rounds, whereas criteria rated by $\leq 20\%$ of the experts as ‘very important’ or ‘extremely important’ were determined to have reached agreement for exclusion and were not rated in the subsequent Delphi rounds. Remaining criteria were re-rated in a subsequent round where each expert was presented with feedback

Table 1 Experts’ characteristics.

	Expert panel (n = 29)
Socio-demographic data	
Sex (male)	58.6% (n = 17)
Sex (female)	41.4% (n = 12)
Age (range between 32–69 years)	49.68 (10.19)
Geographical distribution	
North America	6.9% (n = 2)
USA	3.4% (n = 1)
Canada	3.4% (n = 1)
South America	3.4% (n = 1)
Brazil	3.4% (n = 1)
Asia	37.9% (n = 11)
China	6.9% (n = 2)
South Korea	6.9% (n = 2)
India	3.4% (n = 1)
Indonesia	3.4% (n = 1)
Iran	3.4% (n = 1)
Israel	3.4% (n = 1)
Japan	3.4% (n = 1)
Malaysia	3.4% (n = 1)
Taiwan	3.4% (n = 1)
Europe	41.4% (n = 12)
UK	6.9% (n = 2)
Italy	6.9% (n = 2)
Germany	6.9% (n = 2)
Spain	6.9% (n = 2)
Belgium	3.4% (n = 1)
France	3.4% (n = 1)
Hungary	3.4% (n = 1)
Switzerland	3.4% (n = 1)
Oceania	10.3% (n = 3)
Australia	10.3% (n = 3)
Academic degree	
MD	65.5% (n = 19)
MSc	13.8% (n = 4)
PhD	72.4% (n = 21)
Experience on gaming disorder	
Both clinical and research experience	82.8% (n = 24)
Only clinical experience	13.8% (n = 4)
Only research experience	3.4% (n = 1)

about his or her previous answer and the answers of the other experts, thus encouraging experts to reconsider their initial answer based on feedback provided. Therefore, round 2 of the Delphi consisted of: (a) new criteria suggested by the experts, to be rated for the first time and (b) the DSM and ICD criteria that did not achieve agreement for inclusion or exclusion in the initial study round.

For those criteria which failed to reach agreement for inclusion or exclusion in a new rating round, the percentage and statistical significance of opinion movements between rounds was determined [51]. The McNemar χ^2 test was employed to assess stability of responses between rounds [36,52,53]. When the percentage of opinion movements for a particular criterion was $< 15\%$ and the

Table 2 Diagnostic validity, utility, and prognostic value of the DSM-5 criteria for IGD.

Criterion	Diagnostic validity	Clinical utility	Prognostic value
1. Preoccupation with internet games	75.9% in round 2	31% in round 2	31% in round 2
2. Withdrawal symptoms when internet gaming is taken away	31% in round 2	31% in round 2	31% in round 3
3. Tolerance—the need to spend increasing amounts of time engaged in internet games	6.9% in round 2	10.3% in round 1	3.4% in round 2
4. Unsuccessful attempts to control the participation in internet games	93.1% in round 2	86.2% in round 1	82.8% in round 1
5. Loss of interests in previous hobbies and entertainment as a result of, and with the exception of, internet games	82.8% in round 2	55.2% in round 2	58.6% in round 2
6. Continued excessive use of internet games despite knowledge of psychosocial problems	89.7% in round 1	82.8% in round 1	82.8% in round 1
7. Has deceived family members, therapists or others regarding the amount of internet gaming	10.3% in round 2	17.2% in round 1	13.8% in round 1
8. Use of internet games to escape or relieve a negative mood (e.g. feelings of helplessness, guilt, anxiety)	24.1% in round 2	10.3% in round 2	31% in round 2
9. Has jeopardized or lost a significant relationship, job or educational or career opportunity because of participation in internet games	93.1% in round 1	86.2% in round 1	86.2% in round 2

Cells marked in green indicate that the criterion reached agreement for inclusion (i.e. $\geq 80\%$ of experts rated the criterion as 'very important' or 'extremely important'); cells marked in red indicate that the criterion reached agreement for exclusion (i.e. $\leq 20\%$ of experts rated the criterion as 'very important' or 'extremely important'); cells marked in yellow indicate that the criterion did not reach agreement either for inclusion or for exclusion (i.e. $> 20\%$ of experts but $< 80\%$ rated the criterion as 'very important' or 'extremely important'); figures within each cell represent the percentage of experts scoring the criterion as 'very important' or 'extremely important' in the last round that the criterion was rated. IGD = internet gaming disorder. [Colour table can be viewed at wileyonlinelibrary.com]

Table 3 Diagnostic validity, utility and prognostic value of the ICD-11 clinical guidelines for GD.

Criterion	Diagnostic validity	Clinical utility	Prognostic value
1. Impaired control over gaming (e.g. onset, frequency, intensity, duration, termination, context)	93.1% in round 1	93.1% in round 1	82.8% in round 1
2. Increasing priority given to gaming to the extent that gaming takes precedence over other life interests and daily activities	82.8% in round 1	79.3% in round 2	75.9% in round 2
3. Continuation or escalation of gaming despite the occurrence of negative consequences	93.1% in round 1	96.6% in round 1	86.2% in round 1
4. The behaviour pattern is of sufficient severity to result in significant impairment in personal, family, social, educational, occupational or other important areas of functioning	100% in round 1	89.7% in round 1	89.7% in round 1

Cells marked in green indicate that the criterion reached agreement for inclusion (i.e. $\geq 80\%$ of experts rated the criterion as 'very important' or 'extremely important'); cells marked in red indicate that the criterion reached agreement for exclusion (i.e. $\leq 20\%$ of experts rated the criterion as 'very important' or 'extremely important'); cells marked in yellow indicate that the criterion did not reach agreement either for inclusion or for exclusion (i.e. $> 20\%$ of experts but $< 80\%$ rated the criterion as 'very important' or 'extremely important'); figures within each cell represent the percentage of experts scoring the criterion as 'very important' or 'extremely important' in the last round that the criterion was rated. GD = gaming disorder. [Colour table can be viewed at wileyonlinelibrary.com]

McNemar χ^2 test was non-significant, the criterion was not re-rated (i.e. answers were considered stable, with no expectation that they would significantly change in subsequent rounds). Similarly, criteria with a percentage of opinion movements $< 15\%$ and a significant McNemar χ^2 test, or with a percentage of opinion movements $> 15\%$ but a non-significant McNemar χ^2 test, were not re-rated in subsequent study rounds. When the percentage of opinion movements was $> 15\%$ and the McNemar χ^2 test was significant, change in subsequent rounds could be expected and the criterion was thus kept for subsequent rounds. The present Delphi required three rounds to be completed.

This study protocol was approved by the Ethics Review Panel of the University of Luxembourg (ERP 18-047a). The present Delphi study was not pre-registered and is thus exploratory in nature.

RESULTS

Expert panel characteristics and retention rates

A total of 29 international experts were involved in this Delphi survey. Retention rate between rounds was 100% (i.e. all experts who completed round 1 also completed

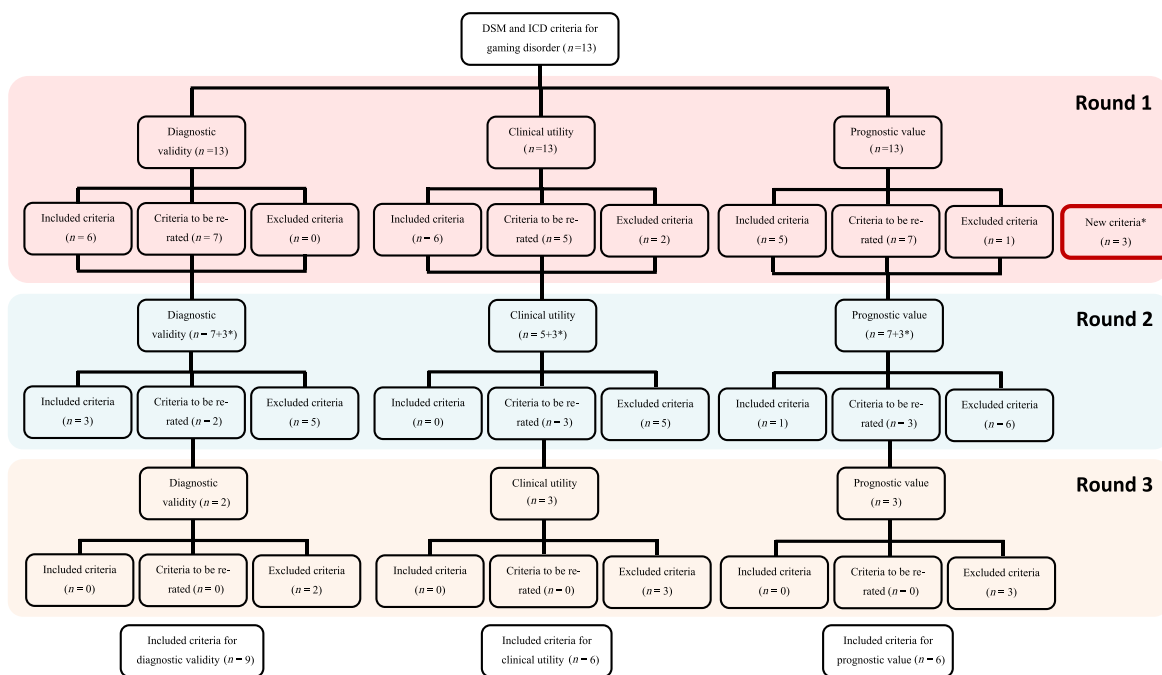


Figure 2 Flow-chart of the criteria inclusion, exclusion or re-rating over the study rounds. [Colour figure can be viewed at wileyonlinelibrary.com]

rounds 2 and 3). Table 1 provides an overview of experts' characteristics.

Among those who reported clinical experience, the mean history of treating patients with GD was 7.9 years [standard deviation (SD) = 5.4; range = 1–25]. Among those reporting research experience, the average number of papers on GD published in peer-reviewed journals was 13.12 (SD = 12.23; range = 2–45). As an indicator of research performance (productivity and impact), we consulted the experts' h-index in Scopus [54]. The average h-index was 23.53 (SD = 17.0; range = 3–76).

Criteria inclusion, exclusion and re-rating

Figure 2 provides a summary of the results. In the first round, experts rated the DSM-5 criteria (nine items) and the ICD-11 clinical guidelines (four items) (n = 13). In round 1, six criteria were rated as 'very' or 'extremely important' by ≥ 80% of the experts in regard to their diagnostic validity and clinical utility and five for their prognostic value. Additionally, three new criteria suggested by the experts were added and rated in the next study round (i.e. dissociation, health consequences and craving). In round 2,

Table 4 Diagnostic validity, utility and prognostic value of the new GD criteria proposed by the expert panel.

Criterion	Diagnostic validity	Clinical utility	Prognostic value
1. Health consequences resulting from gaming activity (e.g. significant sleep deprivation or changes in sleep patterns, significant weight changes due to a reduction of food intake or due to extended periods of physical inactivity, back or wrist pain, etc.)	86.2% in round 2	69% in round 3	34.5% in round 3
2. Craving or a strong desire or urge to play video games	69% in round 3	37.9% in round 3	44.8% in round 3
3. Dissociative-like symptoms while or after playing videogames (deep and absorbed state of consciousness while gaming, loss of time perception, deep immersion, inattention to events happening around during gaming sessions or difficulty distinguishing games and real life)	24.1% in round 3	6.9% in round 3	13.8% in round 2

Cells marked in green indicate that the criterion reached agreement for inclusion (i.e. ≥ 80% of experts rated the criterion as 'very important' or 'extremely important'); cells marked in red indicate that the criterion reached agreement for exclusion (i.e. ≤ 20% of experts rated the criterion as 'very important' or 'extremely important'); cells marked in yellow indicate that the criterion did not reach agreement either for inclusion or for exclusion (i.e. > 20% of experts but < 80% rated the criterion as 'very important' or 'extremely important'); figures within each cell represent the percentage of experts scoring the criterion as 'very important' or 'extremely important' in the last round that the criterion was rated. GD = gaming disorder.

three additional criteria reached expert agreement for inclusion regarding diagnostic validity, one regarding prognostic value and none regarding clinical utility. In round 3, no additional expert agreement was reached. Experts' opinions in round 3 were stable (i.e. percentage of opinion movements between rounds 2 and 3 for the remaining criteria was < 15% and McNemar χ^2 test was non-significant), thus concluding the Delphi survey iterations.

Expert agreement on DSM-5 IGD criteria

Of the nine DSM-5 criteria, only four reached agreement for inclusion as 'very important' or 'extremely important' by > 80% (Table 2): (a) jeopardizing relationship and/or career opportunity; (b) impaired control; (c) continued use; and (d) diminished interests (expert agreement only for diagnostic validity). For three criteria there was expert agreement for exclusion, with ratings 'very important' or 'extremely important' by \leq 20% of the panellists: (a) tolerance; (b) deception (family members, therapists or others); and (c) mood regulation. The remaining criteria did not achieve expert agreement and so were neither retained nor rejected (agreement ratings between 24.1 and 75.9%).

Expert agreement on ICD-11 diagnostic guidelines

As displayed in Table 3, the four items in the ICD-11 clinical guidelines for GD each reached agreement for diagnostic validity, clinical utility and/or prognostic value: (a) functional impairment; (b) continuation or escalation of gaming (expert agreement only for diagnostic validity); (c) impaired control; and (d) increasing priority given to gaming.

Expert agreement on additional diagnostic criteria

Table 4 shows the three criteria not included in the ICD-11 and DSM-5 definitions and considered as relevant criteria by some of the experts in round 1. From these criteria, only one achieved expert agreement for inclusion: the presence of health consequences resulting from gaming (expert agreement only for diagnostic validity). In contrast, there was expert agreement for exclusion of dissociative-like symptoms while or after playing videogames. Opinions regarding the relevance of craving for GD diagnosis were more divided, in particular regarding its diagnostic validity (69% of experts rated the criterion as 'very important' or 'extremely important').

DISCUSSION

The present study employed a Delphi method in an international sample of scholars and clinicians in the field of GD.

The outcomes of this study: (1) provided an expert appraisal of GD criteria perceived to have the highest diagnostic validity, clinical utility and prognostic value; (2) obtained expert agreement that several proposed GD criteria were perceived to have low diagnostic validity, clinical utility and prognostic value; and (3) failed to reach an agreement on the validity of several existing GD criteria. These findings inform continuing discussions of the theoretical conceptualization and clinical diagnosis of GD.

This study successfully identifies a subset of diagnostic criteria that reached expert agreement regarding their high clinical relevance (i.e. diagnostic validity, clinical utility and prognostic value). These criteria include loss of control (DSM-5 and ICD-11), gaming despite harms (DSM-5 and ICD-11), conflict/interference due to gaming (DSM-5) and functional impairment (ICD-11). Crucially, the DSM-5 criteria that performed well in the current Delphi were those that were included in the ICD-11 definition.¹

There was also expert agreement that several proposed criteria (all included in the DSM-5 but not retained in the ICD-11) had low clinical relevance based on the indicators assessed (i.e. diagnostic validity, clinical utility and prognostic value). It appears first that two proposed IGD criteria (tolerance and deception of others about gaming) reached expert agreement regarding their inadequacy in the context of GD (> 80% of agreement among experts for the three indicators assessed), whereas the escape/regulating mood was judged not relevant in terms of clinical utility, but views were mixed regarding diagnostic validity and prognostic value. These findings suggest that these three criteria should not be considered in the definition and diagnosis of GD, which is consistent with findings of previous studies on the diagnostic accuracy of IGD criteria [21,22,55] as well as with work using an Open Science framework to encourage transparent and collaborative operational definition of behavioural addictions [56]. The latter stipulated that high involvement in gaming should not be considered as problematic when it constitutes 'a temporary coping strategy as an expected response to common stressors or losses' [56]. It is worth noting that the additional criterion 'dissociation-like symptoms' (not included in the DSM-5 and ICD-11 definition but proposed by several authors in the first round and assessed in the second round) did not achieve expert agreement, perhaps reflecting the relatively limited empirical evidence available to date [57].

Another important result is that for multiple proposed criteria, expert agreement regarding their clinical relevance in the diagnosis of GD was not reached. This includes preoccupation (DSM-5), withdrawal (DSM-5), diminished interest in non-gaming pursuits (DSM-5 and

¹We consider here that the DSM-5 'conflict/interference due to gaming' overlaps with the ICD-11 'functional impairment' criterion, even if the ICD-11 criterion probably reflects more severe impairments/consequences [27].

ICD-11), health consequences (new criterion proposed by the panel) and craving (new criterion proposed by the panel). Among these, two criteria were considered as clinically valid and thus represent characteristic features of GD (diminished non-gaming interests, health consequences), while no agreement was reached regarding their clinical utility and prognostic value. The other three criteria (preoccupation, withdrawal and craving) failed to reach expert agreement regarding all indicators assessed (diagnostic validity, clinical utility and prognostic value). The lack of agreement regarding these criteria might be explained by different (not mutually exclusive) factors. First, some experts may have followed a conservative approach and were thus reluctant to accept new potential criteria in a context where existing criteria are under debate. This was potentially the case for a criterion such as craving, which demonstrated excellent diagnostic accuracy in a sample of treatment-seeking gamers [21]. Secondly, some experts might have based decisions on availability of conclusive or sufficient evidence (see [58] for the withdrawal criteria) or on theoretical coherence (preoccupation and craving have, for example, been linked to distinct constructs such as cue reactivity, attention bias or irrational beliefs [59,60]). These experts may have considered these aspects as core psychological processes underlying GD (as proposed by recent theoretical models [61]), but not necessarily useful as diagnostic criteria.

Several limitations of the study warrant mention. First, it is not possible to ascertain whether the current panel is sufficiently representative, given the selection criteria. However, in contrast to previous attempts to provide expert agreement for GD criteria [12,18], we used a structured and transparent approach to select experts. Our aim was to select an international group of experts based on their clinical experience and research achievement in the specific field of GD. This resulted in the inclusion of a range of different experts (e.g. experts having taken part in various ICD-11 and DSM-5 working groups related to substance-use and addictive disorders or obsessive-compulsive disorders, experienced clinicians working with patients presenting with gaming-related problems or scholars having opposed the recognition of GD in the ICD-11). In this regard, it is worth mentioning that 11 of the 29 experts in the study were members of the WHO advisory group on GD, which is actually not surprising, because members of this WHO group have in common a specific expertise on gaming disorder research and treatment (such specific expertise was not necessarily present among the experts who created the criteria included in the DSM-5). Moreover, one might argue that experts from other disciplines should have been included in the panel (e.g. sociology, social psychology, game studies, communication sciences, or anthropology). Although these disciplines significantly contributed to the conceptualization

and understanding of GD (see [62] for an anthropological understanding of problematic gaming), we did not include experts from these disciplines as we reasoned that their capacity to judge the diagnostic validity, clinical utility and prognostic value of diagnostic criteria might be more limited, and they are not expected to be familiar with the practical use of diagnostic manuals, nor directly involved in the treatment of GD patients. Secondly, although we took particular care to make the panel as international as possible, most experts came from Europe (42%) or Asia (36.3%), which may be explained by the fact research on GD is more active in these continents and that we limited the maximum number of experts per country (e.g. the maximum was three for Australia). Thirdly, there are different kinds of expertise (other than clinical and research) that were not included in the study. Some Delphi studies have utilized small groups of patients as 'experts on their own condition' [50]. Given their lived experience, patients with GD and/or in remission from this condition may provide valuable input for assessing the performance of GD diagnostic criteria.

CONCLUSIONS

The present Delphi study was the first one, to our knowledge, to use a structured and transparent approach to assess the clinical relevance of the DSM-5 and ICD-11 diagnostic criteria for GD with an international expert panel. The central finding is that there was expert agreement that some of the DSM-5 criteria were not clinically relevant, and that ICD-11 items (except for the criterion relating to diminished non-gaming interests) showed clinical relevance. Our findings, which align with previous critiques of IGD criteria [12,15,26], have important implications given the current widespread use of the DSM-5 IGD criteria for epidemiological, psychometric, clinical and neurobiological research. Furthermore, the data here provide systematic support for the previously expressed view [26] that some criteria for substance-use and gambling disorders adapted for use in the GD context may not sufficiently distinguish between high (but non-problematic involvement) and problematic involvement in video gaming. This is concerning, as gaming is a mainstream hobby in which people all around the world engage very regularly and the risks of over-diagnosis is real [8,10,26]. Specifically, there was an agreement among experts that tolerance and mood regulation should not be used to diagnose GD. These findings should be considered during future revision of the DSM-5. Importantly, there was strong expert agreement that ICD-11 GD diagnostic guidelines are likely to allow the diagnosis of GD without pathologizing healthy gaming. Importantly, the current study by no means validated a subset of diagnostic criteria to diagnose GD. Indeed, additional empirical research with clinical samples is needed to determine the precise diagnostic accuracy and

prognostic value of existing GD diagnostic criteria. Finally, and as suggested elsewhere [11,14,63,64], we believe that our approach should be complemented with phenomenological and qualitative work conducted in treatment-seeking gamers, in order to identify potentially unique features of GD not considered in the current work.

Declaration of interests

D.L.K., D.J.S., Z.D., H.J.R., S.A., A.A., H.B.J., E.M.L.C., H.K.L., S.H., D.J.K., J.L., M.N.P., A.R.M., J.B.S., D.T.S. and J.B. are members of a WHO Advisory Group on Gaming Disorder. D.J.S., N.A.F., H.J.R., S.A., M.N.P. and A.R.M. have been involved in other groups (e.g. related to substance use and addictive disorders or obsessive and compulsive disorders) in the context of ICD-11 and DSM-5 development. J.C.C., M.B., L.C., M.Y., N.B., A.B., X.C., C.H.K., M.D., H.K.L., S. H., S.J.M., O.K., D.J.K., A.M., S.P., A.S., S.Y.L., K.S., V.S., A. M.W. and K.W. have no conflicts of interest to declare. D.J.S., S.R.C., Z.D., N.A.F., P.D.T., M.G.B. and M.N.P. have other conflicts of interest not relevant in the study design, management, data analysis/interpretation or write-up of the data that are therefore not reported here.

Acknowledgements

S.R.C.'s research is funded by a Wellcome Trust Clinical Fellowship (110 049/Z/15/Z). Z.D. was supported by the Hungarian National Research, Development and Innovation Office (grant numbers: KKP126835, NKFIH-1157-8/2019-DT). O.K. was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the ÚNKP-20-5 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund. M.N.P. received support from the Connecticut Department of Mental Health and Addiction Services, the National Center for Responsible Gaming and the Connecticut Council on Problem Gambling. M.Y. has received funding from Monash University, and Australian Government funding bodies such as the National Health and Medical Research Council (NHMRC; including Fellowship no. APP1117188), the Australian Research Council (ARC), Australian Defence Science and Technology (DST) and the Department of Industry, Innovation and Science (DIIS). He has also received philanthropic donations from the David Winston Turner Endowment Fund and the Wilson Foundation. This article is based upon work from COST Action CA16207 'European Network for Problematic Usage of the Internet', supported by COST (European Cooperation in Science and Technology); <http://www.cost.eu>

Author contributions

Jesús Castro-Calvo: Conceptualization; data curation; formal analysis; investigation; methodology; project

administration; supervision. **Daniel King:** Conceptualization; methodology. **Dan Stein:** Conceptualization; methodology. **Matthias Brand:** Conceptualization; methodology. **Lior Carmi:** Conceptualization; methodology. **Samuel Chamberlain:** Conceptualization; methodology. **Zsolt Demetrovics:** Conceptualization; methodology. **Naomi Fineberg:** Conceptualization; methodology. **Hans-Juergen Rumpf:** Conceptualization; methodology. **Murat Yücel:** Conceptualization; methodology. **Sophia Achab:** Investigation. **Atul Ambekar:** Investigation. **Norharlina Bahar:** Investigation. **Alex Blaszczynski:** Investigation. **Henrietta Bowden-Jones:** Investigation. **Xavier Carbonell:** Investigation. **Elda Chan:** Investigation. **Chih-Hung Ko:** Investigation. **Philippe de Timary:** Investigation. **Magali Dufour:** Investigation. **Marie Grall Bronnec:** Investigation. **Hae Kook Lee:** Investigation. **Susumu Higuchi:** Investigation. **Susana Jiménez-Murcia:** Investigation. **Orsolya Király:** Investigation. **Daria Kuss:** Investigation. **Jiang Long:** Investigation. **Astrid Mueller:** Investigation. **Stefano Pallanti:** Investigation. **Marc Potenza:** Investigation. **Afarin Rahimi-Movaghar:** Investigation. **John B. Saunders:** Investigation. **Adriano Schimmenti:** Investigation. **Seung-Yup Lee:** Investigation. **Kristiana Siste:** Investigation. **Daniel Spritzer:** Investigation. **Vladan Starcevic:** Investigation. **Aviv Weinstein:** Investigation. **Klaus Wölfling:** Investigation. **Joël Billieux:** Conceptualization; data curation; formal analysis; investigation; methodology; project administration; supervision; validation.

References

1. Billieux J., King D. L., Higuchi S., Achab S., Bowden-Jones H., Hao W., *et al.* Functional impairment matters in the screening and diagnosis of gaming disorder. *J Behav Addict* 2017; 6: 285–9.
2. Rumpf H. J., Achab S., Billieux J., Bowden-Jones H., Carragher N., Demetrovics Z., *et al.* Including gaming disorder in the ICD-11: the need to do so from a clinical and public health perspective. Commentary on: a weak scientific basis for gaming disorder: let us err on the side of caution (van Rooij *et al.*, 2018). *J Behav Addict* 2018; 7: 556–61.
3. Saunders J. B., Hao W., Long J., King D. L., Mann K., Fauth-Bühler M., *et al.* Gaming disorder: its delineation as an important condition for diagnosis, management, and prevention. *J Behav Addict* 2017; 6: 271–9.
4. American Psychiatric Association *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*, 5th edn. Washington, DC: American Psychiatric Association; 2013.
5. World Health Organization. International Classification of Diseases: ICD-11 for Mortality and Morbidity Statistics [internet]. 2019 [cited 2019 Dec 10]. Available at: <https://icd.who.int/dev11/l-m/en>, (accessed 10 December 2019).
6. King D. L., Koster E., Billieux J. Study what makes games addictive. *Nature* 2019; 573: 346.
7. Stein D. J., Billieux J., Bowden-Jones H., Grant J. E., Fineberg N., Higuchi S., *et al.* Balancing validity, utility and public health considerations in disorders due to addictive behaviours. *World Psychiatry* 2018; 17: 363–4.

8. Aarseth E., Bean A. M., Boonen H., Colder Carras M., Coulson M., Das D., et al. Scholars' open debate paper on the world health organization ICD-11 gaming disorder proposal. *J Behav Addict* 2017; **6**: 267–70.
9. Van Rooij A. J., Ferguson C. J., Carras M. C., Kardefelt-Winther D., Shi J., Aarseth E., et al. A weak scientific basis for gaming disorder: let us err on the side of caution. *J Behav Addict* 2018; **7**: 1–9.
10. Bean A. M., Nielsen R. K. L., van Rooij A. J., Ferguson C. J. Video game addiction: the push to pathologize video games. *Prof Psychol Res Pract* 2017; **48**: 378–89.
11. 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**: 119–23.
12. Griffiths M. D., van Rooij A. J., 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**: 167–75.
13. King D. L., Billieux J., Carragher N., Delfabbro P. H. Face validity evaluation of screening tools for gaming disorder: scope, language, and overpathologizing issues. *J Behav Addict* 2020; **9**: 1–13.
14. Kardefelt-Winther D., Heeren A., Schimmenti A., van Rooij A., Maurage P., Carras M., et al. How can we conceptualize behavioural addiction without pathologizing common behaviours? *Addiction* 2017; **112**: 1709–15.
15. Kardefelt-Winther D. A conceptual and methodological critique of internet addiction research: towards a model of compensatory internet use. *Comput Hum Behav* 2014; **31**: 351–4.
16. Schimmenti A., Caretti V. Psychic retreats or psychic pits?: unbearable states of mind and technological addiction. *Psychoanal Psychol* 2010; **27**: 115–32.
17. Starcevic V., Aboujaoude E. Internet gaming disorder, obsessive-compulsive disorder, and addiction. *Curr Addict Rep* 2017; **4**: 317–22.
18. Petry N. M., Rehbein F., Ko C. H., O'Brien C. P. Internet gaming disorder in the DSM-5. *Curr Psychiatry Rep* 2015; **17**: 72.
19. Rehbein F., Kliem S., Baier D., Mößle T., Petry N. M. Prevalence of internet gaming disorder in German adolescents: diagnostic contribution of the nine DSM-5 criteria in a state-wide representative sample. *Addiction* 2015; **110**: 842–51.
20. Przybylski A. K., Weinstein N., Murayama K. Internet gaming disorder: investigating the clinical relevance of a new phenomenon. *Am J Psychiatry* 2017; **174**: 230–5.
21. Ko C.-H., Yen J.-Y., Chen S.-H., Wang P.-W., Chen C.-S., Yen C.-F. Evaluation of the diagnostic criteria of internet gaming disorder in the DSM-5 among young adults in Taiwan. *J Psychiatr Res* 2014; **53**: 103–10.
22. Müller K. W., Beutel M. E., Dreier M., Wölfling K. A clinical evaluation of the DSM-5 criteria for internet gaming disorder and a pilot study on their applicability to further internet-related disorders. *J Behav Addict* 2019; **8**: 16–24.
23. Brunborg G. S., Mentzoni R. A., Melkevik O. R., Torsheim T., Samdal O., Hetland J., et al. Gaming addiction, gaming engagement, and psychological health complaints among Norwegian adolescents. *Media Psychol* 2013; **16**: 115–28.
24. Charlton J. P., Danforth I. D. W. Distinguishing addiction and high engagement in the context of online game playing. *Comput Hum Behav* 2007; **23**: 1531–48.
25. Deleuze J., Long J., Liu T., Maurage P., Billieux J. Passion or addiction? Correlates of healthy versus problematic use of videogames in a sample of French-speaking regular players. *Addict Behav* 2018; **82**: 114–21.
26. Billieux J., Flayelle M., Rumpf H.-J., Stein D. J. High involvement versus pathological involvement in video games: a crucial distinction for ensuring the validity and utility of gaming disorder. *Curr Addict Rep* 2019; **6**: 323–30.
27. Jo Y. S., Bhang S. Y., Choi J. S., Lee H. K., Lee S. Y., Kweon Y.-S. Clinical characteristics of diagnosis for internet gaming disorder: comparison of DSM-5 IGD and ICD-11 GD diagnosis. *J Clin Med* 2019; **8**: 945.
28. Jones J., Hunter D. Qualitative research: consensus methods for medical and health services research. *BMJ* 1995; **311**: 376–80.
29. Diamond I. R., Grant R. C., Feldman B. M., Pencharz P. B., Ling S. C., Moore A. M., et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol* 2014; **67**: 401–9.
30. Hsu C., Sandford B. The Delphi technique: making sense of consensus. *Pract Assess Res Eval* 2007; **12**: 1–8.
31. Jorm A. F. Using the Delphi expert consensus method in mental health research. *Aust NZ J Psychiatry* 2015; **49**: 887–97.
32. Yücel M., Oldenhof E., Ahmed S. H., Belin D., Billieux J., Bowden-Jones H., et al. A transdiagnostic dimensional approach towards a neuropsychological assessment for addiction: an international Delphi consensus study. *Addiction* 2018; **114**: 1095–109.
33. MacFarlane L., Owens G., Del Pozo Cruz B. Identifying the features of an exercise addiction: a Delphi study. *J Behav Addict* 2016; **5**: 474–84.
34. Bolger F., Wright G. Improving the Delphi process: lessons from social psychological research. *Technol Forecast Soc Change* 2011; **78**: 1500–13.
35. Jairath N., Weinstein J. The Delphi methodology (part one): a useful administrative approach. *Can J Nurs Adm* 1994; **7**: 29–42.
36. von der Gracht H. A. Consensus measurement in Delphi studies. Review and implications for future quality assurance. *Technol Forecast Soc Change* 2012; **79**: 1525–36.
37. Surowiecki J. *The Wisdom of Crowds: Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. New York, NY: Doubleday; 2004.
38. Jönsson M. L., Hahn U., Olsson E. J. The kind of group you want to belong to: effects of group structure on group accuracy. *Cognition* 2015; **142**: 191–204.
39. Davis-Stober C. P., Budescu D. V., Dana J., Broomell S. B. When is a crowd wise? *Decision* 2014; **1**: 79–101.
40. Petry N. M., Rehbein F., Gentile D. A., Lemmens J. S., Rumpf H. J., Mößle T., et al. An international consensus for assessing internet gaming disorder using the new DSM-5 approach. *Addiction* 2014; **109**: 1399–406.
41. Kuss D. J., Griffiths M. D., Pontes H. M. Chaos and confusion in DSM-5 diagnosis of internet gaming disorder: issues, concerns, and recommendations for clarity in the field. *J Behav Addict* 2017; **6**: 103–9.
42. Petry N. M., Rehbein F., Gentile D. A., Lemmens J. S., Rumpf H. J., Mößle T., et al. Griffiths et al.'s comments on the international consensus statement of internet gaming disorder: furthering consensus or hindering progress? *Addiction* 2016; **111**: 175–8.
43. Addington D. E., McKenzie E., Norman R., Wang J., Bond G. R. Essential evidence-based components of first-episode psychosis services. *Psychiatr Serv* 2013; **64**: 452–7.

44. Ross A. M., Kelly C. M., Jorm A. F. Re-development of mental health first aid guidelines for suicidal ideation and behaviour: a Delphi study. *BMC Psychiatry* 2014; **14**: 1–11.
45. Okoli C., Pawlowski S. D. The Delphi method as a research tool: an example, design considerations and applications. *Inf Manage* 2004; **42**: 15–29.
46. Yap M. B. H., Pilkington P. D., Ryan S. M., Kelly C. M., Jorm A. F. Parenting strategies for reducing the risk of adolescent depression and anxiety disorders: a Delphi consensus study. *J Affect Disord* 2014; **156**: 67–75.
47. Akins R. B., Tolson H., Cole B. R. Stability of response characteristics of a Delphi panel: application of bootstrap data expansion. *BMC Med Res Methodol* 2005; **5**: 1–12.
48. King D. L., Gaming Industry Response Consortium Comment on the global gaming industry's statement on ICD-11 gaming disorder: a corporate strategy to disregard harm and deflect social responsibility? *Addiction* 2018; **113**: 2145–6.
49. Langlands R. L., Jorm A. F., Kelly C. M., Kitchener B. A. First aid recommendations for psychosis: using the Delphi method to gain consensus between mental health consumers, carers, and clinicians. *Schizophr Bull* 2007; **34**: 435–43.
50. Law H., Morrison A. P. Recovery in psychosis: a Delphi study with experts by experience. *Schizophr Bull* 2014; **40**: 1347–55.
51. Scheibe M., Skutsch M., Schofer J. Experiments in Delphi methodology. In: *The Delphi Method: Techniques and Applications*. Boston, MA: Addison-Wesley Publishing; 1975, pp. 262–87.
52. Rayens M. K., Hahn E. J. Building consensus using the policy Delphi method. *Policy Polit Nurs Pract* 2000; **1**: 308–15.
53. Weir C. R., Hicken B. L., Rappaport H., Nebeker J. R. Crossing the quality chasm: the role of information technology departments. *Am J Med Qual* 2006; **21**: 382–93.
54. Hirsch J. E. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA* 2005; **102**: 16569–72.
55. Besser B., Loerbroks L., Bischof G., Bischof A., Rumpf H. J. Performance of the DSM-5-based criteria for internet addiction: a factor analytical examination of three samples. *J Behav Addict* 2019; **8**: 288–94.
56. Billieux J., van Rooij A. J., Heeren A., Schimmenti A., Maurage P., Edman J., et al. Behavioural addiction open definition 2.0—using the Open Science framework for collaborative and transparent theoretical development. *Addiction* 2017; **112**: 1723–4.
57. Guglielmucci F., Monti M., Franzoi I. G., Santoro G., Granieri A., Billieux J., et al. Dissociation in problematic gaming: a systematic review. *Curr Addict Rep* 2019; **6**: 1–14.
58. Kaptsis D., King D. L., Delfabbro P., Gradisar M. Withdrawal symptoms in internet gaming disorder: a systematic review. *Clin Psychol Rev* 2016; **43**: 58–66.
59. Billieux J., Potenza M. N., Maurage P., Brevers D., Brand M., King D. L. Cognitive factors associated with gaming disorder. In: Verdejo-García A., editor. *Cognition and Addiction: A Researcher's Guide From Mechanisms Towards Interventions*. Cambridge, MA: Elsevier Academic Press; 2020, pp. 221–30.
60. King D. L., Delfabbro P. H. The cognitive psychology of internet gaming disorder. *Clin Psychol Rev* 2014; **34**: 298–308.
61. Brand M., Wegmann E., Stark R., Müller A., Wölfling K., Robbins T. W., et al. The interaction of person-affect-cognition-execution (I-PACE) model for addictive behaviors: update, generalization to addictive behaviors beyond internet-use disorders, and specification of the process character of addictive behaviors. *Neurosci Biobehav Rev* 2019; **104**: 1–10.
62. Snodgrass J. G., Dengah H. J. F., Polzer E., Else R. Intensive online videogame involvement: a new global idiom of wellness and distress. *Transcult Psychiatry* 2019; **56**: 748–74.
63. van Rooij A. J., Van Looy J., Billieux J. Internet gaming disorder as a formative construct: implications for conceptualization and measurement. *Psychiatry Clin Neurosci* 2017; **71**: 445–58.
64. Brand M., Rumpf H. J., King D. L., Potenza M. N., Wegmann E. Clarifying terminologies in research on gaming disorder and other addictive behaviors: distinctions between core symptoms and underlying psychological processes. *Curr Opin Psychol* 2020; **36**: 49–54.