



AKADÉMIAI KIADÓ

Journal of Behavioral Addictions

10 (2021) 2, 263–280

DOI:


10.1556/2006.2021.00041

© 2021 The Author(s)

FULL-LENGTH REPORT



Development and validation of a nine-item short screening test for ICD-11 gaming disorder (GAMES test) and estimation of the prevalence in the general young population

SUSUMU HIGUCHI^{1*} , YONEATSU OSAKI², AYA KINJO², SATOKO MIHARA¹, MASAKI MAEZONO¹, TAKASHI KITAYUGUCHI¹, TAKANOBU MATSUZAKI¹, HIDEKI NAKAYAMA³, HANS-JÜRGEN RUMPF⁴ and JOHN B. SAUNDERS⁵

¹ National Hospital Organization Kurihama Medical and Addiction Center, Yokosuka, Kanagawa, Japan

² Tottori University, Faculty of Medicine, Yonago, Tottori, Japan

³ Asahiyama Hospital, Sapporo, Hokkaido, Japan

⁴ University of Lübeck, Lübeck, Germany

⁵ National Centre for Youth Substance Use Research, The University of Queensland, Brisbane, Queensland, Australia

Received: December 6, 2020 • Revised manuscript received: April 18, 2021; May 30, 2021 • Accepted: June 6, 2021
Published online: July 6, 2021

ABSTRACT

Background and Aims: A definition of gaming disorder (GD) was introduced in ICD-11. The purpose of this study was to develop a short screening test for GD, utilizing a reference GD group. It also sought to estimate the prevalence of GD among individuals, representative of the general young population in Japan. *Methods:* Two hundred eighty one men and women selected from the general population, aged between 10 and 29 years, and 44 treatment seekers at our center completed a self-reported questionnaire comprising candidate questions for the screening test. The reference group with ICD-11 GD was established, based on face-to-face interviews with behavioral addiction experts, using a diagnostic interview instrument. The questions in the screening test were selected to best differentiate those who had GD from those who did not, and the cutoff value was determined using the Youden index. *Results:* A nine-item screening test (GAMES test) was developed. The sensitivity and specificity of the test were both 98% and the positive predictive value in the study sample was 91%. The GAMES test comprised two factors, showed high internal consistency and was highly reproducible. The estimated prevalence of GD among the general young population was 7.6% (95% confidence interval; 6.6–8.7%) for males and 2.5% (1.9–3.2%) for females, with a combined prevalence of 5.1% (4.5–5.8%). *Discussion and Conclusion:* The GAMES test shows high validity and reliability for screening of ICD-11 GD. The estimated prevalence of 5.1% among the general young population was comparable to the pooled estimates of young people globally.

KEYWORDS

ICD-11, gaming disorder, internet gaming disorder, screening test, validity, prevalence

*Corresponding author.

E-mail: h-susumu@db3.so-net.ne.jp

INTRODUCTION

Digital gaming has become increasingly popular around the globe (King & Delfabbro, 2019). However, gaming when unrestricted, can be highly absorbing, time-consuming, and

potentially addictive for vulnerable individuals. The World Health Organization (WHO) included gaming disorder (GD) as a disorder due to addictive behaviors in the eleventh revision of the International Classification of Diseases (ICD-11) in May 2019 (WHO, 2019). The decision to do so was based on the accumulated evidence and followed intensive discussion among experts from all over the world (WHO, 2016; Saunders et al., 2017). According to the definition of GD in ICD-11, three clinical manifestation criteria and one functional impairment criterion need to be met to make a diagnosis of GD (WHO, 2019). In addition, these behavior patterns and impairments are required to have persisted for at least 12 months, except in cases where severe symptoms are present.

Prior to the inclusion in ICD-11, preliminary diagnostic criteria of internet gaming disorder (IGD) were published by the American Psychiatric Association in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) in 2013 (American Psychiatric Association, 2013). These comprised nine criteria with at least five criteria needing to be met for a diagnosis of IGD to be made.

In a recent review of studies, the prevalence of problematic gaming ranged from 0.7% to 27.5%, depending on study design, measurement, and study population (Mihara & Higuchi, 2017). The prevalence was higher among males than females and tended to be higher among younger rather than older people in some studies. In a meta-analysis of 16 studies published prior to 2017, the pooled prevalence of problematic gaming was 4.6% among adolescents, confirming the extent of the disorder (Fam, 2018). More recently, a comprehensive review revealed that the prevalence of IGD ranged from 0.21% to 57.5%. However, no study has reported the estimated prevalence of ICD-11 GD (Darvesh et al., 2020).

The development of a screening instrument for a newly conceptualized disorder is of great importance. Firstly, it can serve as an instrument to identify individuals with the disorder for the purposes of preventive and treatment interventions, and it may be used as an adjunctive tool with which to make a diagnosis. Secondly, it can be used to make a provisional estimate of the prevalence of the disorder in a general or specific population. In the case of DSM-5 IGD, several assessment tools and screening tests based on the diagnostic criteria of IGD have been developed (Pontes, Király, Demetrovics, & Griffiths, 2014; Lemmens, Valkenburg, & Gentile, 2015; Pontes & Griffiths, 2015; Percy, Roberts, & McEvoy, 2016; Király et al., 2017; van Rooij, Schoenmakers, & van de Mheen, 2017; Jo et al., 2018). Some of these have been extensively validated and translated into other languages (Pontes & Griffiths, 2015; Király et al., 2017). One such test, the Ten-Item Internet Gaming Disorder Test (IGDT-10) has been translated into 11 languages with cross-cultural validation conducted for seven of these (Király et al., 2019).

In a recent comprehensive and systematic review of screening and assessment tools for IGD and GD, the authors identified all available English-language tools ($N = 32$), and evaluated these in relation to 1) conceptual and practical considerations, 2) alignment with DSM-5 and ICD-11

criteria, 3) type and quantity of studies and samples, and 4) psychometric properties (King et al., 2020). The result showed that the coverage of IGD and GD criteria is inconsistent and tools converge on the importance of screening for control over gaming and functional impairment. It also highlighted how no single instrument was clearly superior to any other. Currently, a gold standard tool that allows for the identification of individuals with GD correctly and accurately has yet to be created, which is unsurprising, as the definition of GD was first published by WHO only in 2019. The authors suggest that the development of a gold standard instrument would be invaluable in enabling professionals to correctly identify gaming-related harms, and develop more effective intervention strategies for those in need (King et al., 2020).

With regard to ICD-11 GD, two assessment tools have been developed. The first tool, called the Gaming Disorder Test (GDT), consists of 4 items representing the three diagnostic guidelines and the functional impairment (Pontes et al., 2019). This test was validated using the results of online surveys given to subjects in the UK and China. However, the purpose of this instrument was not to diagnose GD but to assess its severity and consequences, and the suggested cutoff score was not validated (Pontes et al., 2019). The second instrument - the Gaming Disorder Scale for Adolescents (GADIS-A) - comprised 10 items and was created based on the consensus of clinical and scientific experts (Paschke, Austermann, & Thomasius, 2020). It was designed to exclusively be used with adolescents. The classification of the presence of GD was determined by latent class analysis and the 4 criteria of IGD, similar in content to the key elements of GD, were used as a reference standard (Ko, Lin, Lin, & Yen, 2019).

These two instruments employed polytomous scales using the Likert approach, whereas the screening test this study has developed utilized a dichotomous (yes/no) approach. A study comparing the psychometric properties of the polytomous and dichotomous approaches has been conducted, using the same 9-item and 27-item instruments for IGD (Lemmens et al., 2015). It revealed that both types of assessment were reliable, showed good criteria-related validity, and additionally, the dichotomous 9-item IGD scale showed solid psychometric properties and was the most practical scale for diagnostic purposes.

In this context we have developed a screening instrument for ICD-11 GD using a novel but standardized approach. We conducted a self-reported survey using a questionnaire containing a set of questions relating to the four key elements of the ICD-11 GD definition of gaming behavior among randomly selected community sample groups and treatment seekers. Due to the absence of a gold standard instrument for GD, we developed a new interview tool for establishing a reference group in order to develop a screening test. Survey participants with GD were identified using this instrument based on face-to-face interviews conducted by behavioral addiction experts followed by a panel discussion among the interviewers. Question items for the instrument and the cutoff score to screen for GD were

statistically determined to best differentiate participants with GD from those without GD. The psychometric properties were also explored to ascertain if it could be a valid and reliable screening instrument for assessing GD. The new instrument comprising 9 items was named the GAMing Engagement Screener test (GAMES test) (Appendix). In this study we also estimated the prevalence of GD using the GAMES test among subjects representing the general population, aged between 10 and 29 years, in Japan.

METHODS

Study participants

The subjects of this study were drawn from a representative cohort of the Japanese population residing in Japan who responded to a self-reported questionnaire survey about internet game use and daily habits from January through March 2019. For that, 300 national census spots were randomly identified and 9,000 individuals aged 10–29 years randomly selected from the inhabitant register of the local communities containing these census spots. Professional interviewers visited each household containing a respondent identified using the aforementioned process, requesting that the questionnaire be completed and arranging for its subsequent collection. Of the survey participants, 5,096 individuals gave valid responses to the questionnaire (response rate: 56.6%). Despite the relatively low response rate, the age, gender, and geographical distributions of these respondents were similar to those of initial survey invitees ($N = 9,000$). However, the response rate tended to be lower for participants who were 20 years old or older compared to those under 20 years old. With regard to geographical distribution, the response rates in the Kanto and Kinki areas tended to be lower than for those in other areas. The questionnaire that was administered to the 5,096 participants included all questions that were used for developing the GAMES test. Of these respondents, 2,953 gave consent for the additional survey and 766 of those living in the broader Tokyo metropolitan area were invited to participate in the study to develop a screening test. The interview for this investigation was carried out in July and August 2019, and a total of 281 individuals who had consented to participate in the interview subsequently attended and were enrolled in this study. Of the 766 invitees and 281 participants, Tokyo was over-represented compared to those from adjacent prefectures, but no difference was found in relation to age and gender. Because the number of subjects diagnosed as having GD was too small for detailed statistical analysis, the same self-reported questionnaire survey was additionally conducted among the 44 outpatients who had been definitively diagnosed as having ICD-11 GD by addiction psychiatrists at the Specialized Internet Addiction Clinic of the Center (Fig. 1).

Sources of questions in the screening test

The candidate questions for the screening test were prepared from the following sources: 1) Ten questions adapted for

gaming on the basis of the questions related to 3 clinical manifestation criteria of GD, selected from the Internet Related Disorder-Clinical Assessment Tool (I-CAT) (Brandt et al., 2018; Besser, Loerbroks, Bischof, Bischof, & Rumpf, 2019); 2) six original questions related to the GD criteria, prepared utilizing the day-to-day clinical experience of psychiatrists and clinical psychologists in charge of GD treatment at the Center; 3) ten questions constituting the Japanese version of the IGDT-10 (Király et al., 2017); 4) three questions for the evaluation of functional impairment, designed to assess the influence of GD on social activities using the modified 11-grade visual analog scale (VAS) of Sheehan (Sheehan, Harnett-Sheehan, & Raj, 1996; Hodgins, 2013); and 5) fifteen independent candidate questions for the evaluation of functional impairment derived from clinical observation of GD patients at the Center.

Of these 44 questions, 5 questions not found in the ICD-11 definition of GD (related to preoccupation, tolerance, withdrawal symptoms, deception, and escapism in the IGDT-10) were excluded, and the validity of the remaining 39 potential candidate questions were analyzed. Each candidate question fell under one of the following categories found in the definition of ICD-11 GD: A) loss of control over gaming, B) increasing prioritization of gaming in life, C) functional impairment due to gaming, and D) continuation of gaming despite negative consequences.

Diagnosis of GD

GD in the study participants was definitively diagnosed based on face-to-face interviews by professionals at the Center with expertise in the diagnosis and treatment of GD. Because no validated tool for ICD-11 GD diagnosis had been published at the time of the interviews we developed the interview tool for diagnosis used in the present study (Supplementary material). To minimize differences in symptom evaluation arising between interviewers, the latter had been trained in advance using model cases, so that evaluations could be performed in a standardized manner. The interviews for the diagnosis were carried out individually and 18 individuals were identified with suspected GD. The results obtained in the questionnaire survey were totally blind to interviewers. All experts involved in the interviews later met and of the 18 individuals, seven were diagnosed to have GD. When combined with 44 outpatients diagnosed with GD by addiction psychiatrists and confirmed with the same diagnostic instrument, there was a total of 51 individuals with confirmed GD.

Screening test development

Reference group. The gaming disorder group comprised 51 persons definitively diagnosed, as above. The control group comprised 274 subjects, after the exclusion of the 7 individuals who had been diagnosed with GD.

Statistical evaluation of questions. First, the relationship of individual candidate questions to the presence/absence of

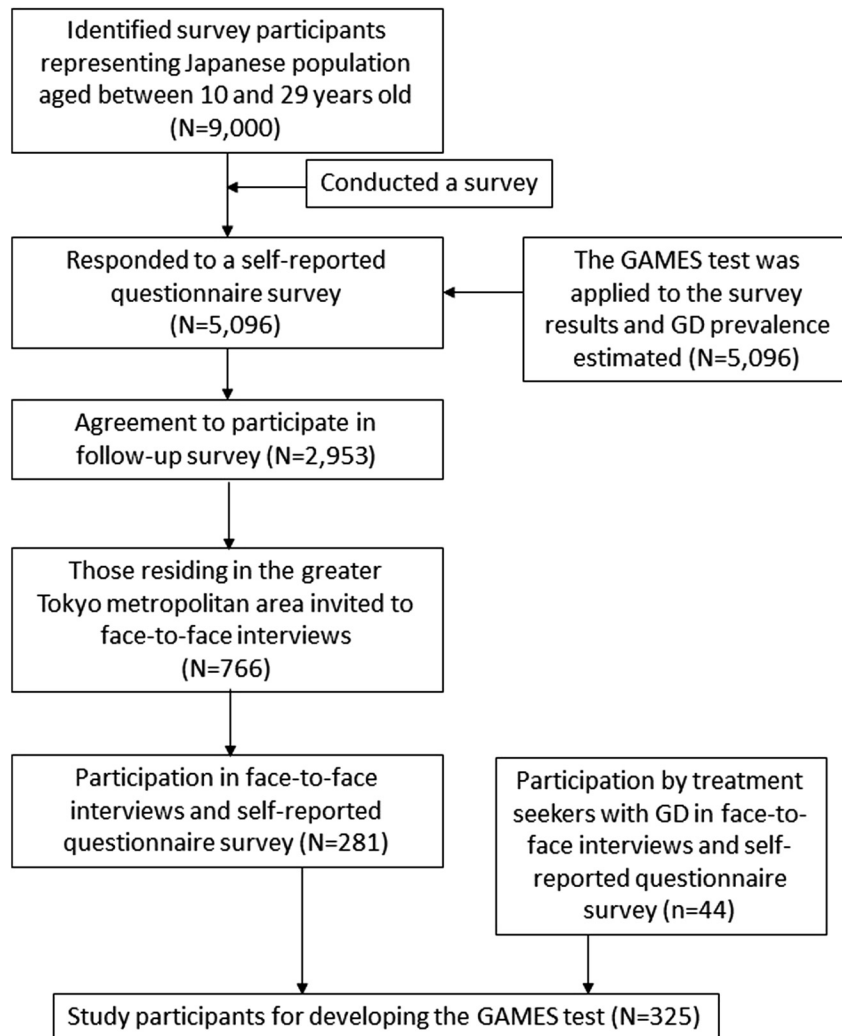


Fig. 1. Flow diagram for identifying study participants in the development of the GAMES test and estimating the prevalence of GD. GAMES = Gaming Engagement Screener. GD = gaming disorder

GD was judged using WAIC (widely applicable information criterion) and WBIC (widely applicable Bayesian information criterion), which are evaluation indicators of the degree of fitness of a statistical model (Watanabe, 2010, 2013). Both WAIC and WBIC can be used, regardless of the data distribution, and the lower the values, the greater the fitness of the model.

To avoid the problem of multicollinearity, the variance inflation factor (VIF) was calculated. No combination of questions yielded VIF values exceeding 10, allowing us to rule out multicollinearity. When the Spearman correlation coefficient between two items was used for the evaluation of multicollinearity, a high correlation coefficient was noted sporadically. When the correlation coefficient between two items was 0.6 or higher, only one of the two was adopted in the multivariate statistical model.

Selection of variables for screening and setting their cutoff level. Next, the variables were selected in accordance with the definition of GD. For each of the 4 categories (A through D) mentioned above, 2 variables with a relatively low WAIC

or WBIC were selected. Then, multiple logistic regression analysis was conducted, using the presence/absence of GD as the dependent variable with the gender, age, and presence/absence of candidate items as covariates for adjustment to evaluate the degree of fitness of the statistical model. The Hosmer-Lemeshow test was used to evaluate the applicability for logistic regression models. We determined in advance that an item which was met by at least two out of seven diagnosed individuals among the general young population would be included in the test, because an item positive for only one diagnosed participant was deemed inappropriate.

As the time spent gaming on weekdays has been shown to be a variable useful for the identification of GD, this was added to the 8 variables selected by the above-mentioned process to yield the Gaming Engagement Screener test (GAMES test). The sensitivity and specificity of this screening test were evaluated by separately conducting an analysis for optimal identification of the combined 51 subjects with GD. The optimum cutoff level was determined using the Youden's Index (Sensitivity + Specificity - 1) (Bantis, Nakas, & Reiser, 2014).

Evaluation of the reliability and validity of the items constituting the GAMES test

Evaluation by factor analysis. An exploratory factor analysis was conducted to examine whether or not the items constituting the screening test possessed a factor structure conforming to the framework of the GD definition. Principal factor analysis for 325 subjects was used as the method for this factor analysis. The generalized least squares method was employed for factor extraction. Promax rotation was conducted for interpretation of the results of factor analysis. The Bartlett sphericity test and Kaiser-Meyer-Olkin measure of sampling adequacy were used to check the prerequisites for the factor analysis (Malhotra, 1999; Hair, Black, Babin, & Anderson, 2010).

Relationship to the IGDT-10 score. The degree of agreement between the total GAMES test score and the total IGDT-10 score and between the presence of GD (score 5 or higher) and the presence of IGD were examined. Because the Japanese version of the IGDT-10 has not been validated, the presence of IGD was determined using the originally suggested cutoff point (Király et al., 2017). The weighted κ coefficient and the intraclass correlation coefficient (ICC) served as the degree of agreement.

Evaluation of internal consistency and test-retest reliability. Internal consistency was evaluated using Cronbach's alpha and the omega coefficients. For evaluation of the test-retest reproducibility, a second self-reported survey using the same questionnaire was conducted two weeks later involving a randomly selected half of the participants ($N = 146$) and the 44 outpatients with GD. Responses were collected from 109 survey participants and from all 44 outpatients. Excluding the 9 survey participants who failed to give valid answers to the primary questions, the responses of the remaining 144 subjects were included in the analysis. Reproducibility, an indicator of reliability, was evaluated with the weighted κ coefficient and ICC.

Estimation of the prevalence of GD among young Japanese

The GAMES test was applied to 5,096 respondents of the initial survey, representing young Japanese aged between 10 and 29 years. The data were adjusted by every one-year age group and gender for the estimated population in Japan as of October 1, 2018 (Statistics Bureau of Japan, 2018).

All statistical analyses of the data were conducted using SPSS Version 25.0 (IBM Corp, 2017).

Ethics

This study was approved by the Ethics Committee of the Center (Approval No. 332). Informed consent to participate in the study was obtained from all participants. If participants were younger than 18 years old informed consent was obtained from not only the participants but also their

guardians. The study procedures were performed in accordance with the Declaration of Helsinki.

RESULTS

Development of the GAMES test

Selection of 8 items representing 4 criteria of GD. Of the 39 items analyzed, 11 items received negative responses from all 7 individuals identified as having GD in the survey of the general population. Excluding these 11 items, the WAIC and WBIC values of the remaining 28 items are shown in Table 1. The items with low WAIC and WBIC values that received an affirmative answer from 2 or more of these 7 individuals were questions 4, 5, and 6 in Group A, questions 7, 8, 9, and 11 in Group B, questions 13, 14, 15, 16, 20, and 21 in Group C, and questions 24 and 26 in Group D.

Two items were selected from each of the 4 groups and multiple logistic regression analysis was repeatedly conducted. The highest R^2 (Cox-Snell $R^2 = 0.548$, Nagelkerke $R^2 = 0.943$) was recorded by the model including questions 4 and 5 from Group A, questions 7 and 9 from Group B, questions 16 and 21 from Group C, and questions 24 and 26 from Group D. All of these items were met for at least 3 out of 7 diagnosed individuals from among the young population. When the presence/absence of GD was predicted using these 8 variables with a cutoff score of 0.5 (absence = 0/ presence = 1), the accuracy of the prediction was 96.1%. The test of Hosmer and Lemeshow yielded a P value of 0.999. Thus, the null hypothesis was not rejected and the model fitness was judged as high.

The results of the multiple logistic regression analysis are shown in Table 2. Because the data from outpatients with GD were added to this analysis to enable multivariate analysis, the odds ratio changed as the number of added outpatients increased or decreased. The frequency of subjects providing affirmative answers for each item differed between the GD and control groups; the odds ratios were quite high, and confidence intervals were also wide. Therefore, we judged that the odds ratio value may not be important in deciding whether to include/exclude a given item from the screening test.

Cutoff score and sensitivity and specificity. Table 3 shows the data on the sensitivity and specificity of the 8-item screening test without the ninth item and the 9-item GAMES test. The table additionally gives the results of the analysis of the cutoff level for the screening test under two settings; combined 51 subjects with GD and only 7 subjects from the general population. Based on the Youden index values, the optimal cutoff score was ≥ 4 (full score = 8, score point of 1 for each item that elicited an affirmative response and 0 for each item that elicited a negative response among the 8 items) for the combined data. In this case, a sensitivity of 94% and specificity of 97% were obtained.

When the average gaming time on a typical weekday was added to the items analyzed, both the sensitivity and



Table 1. WAIC and WBIC values and odds ratios of candidate items for the presence/absence of gaming disorder and the number of items affirmed by 7 individuals from the general population with gaming disorder^{a,b)}

No.	Source ^{c)}	Diagnostic category	Question items	WAIC	WBIC	WAIC+WBIC	Rank by WAIC	Rank by WBIC	Rank by WAIC+WBIC	Yes by person without disorder	No by person without disorder	Yes by person with disorder	No by person with disorder	Odds ratio	Yes by general population subject with disorder (n = 7)
1	I-CAT	A	Have you felt the continuing desire to stop or reduce gaming because you noticed to use games too much?	279.37	140.94	420.31	26	25	26	70	204	29	22	3.84	4
2	I-CAT	A	Have you repeatedly tried to stop or reduce gaming because you thought to use games too much and have been unsuccessful with this attempt?	293.10	147.89	440.99	28	28	28	65	209	17	34	1.61	1
3	Original	A	Have you often played games under circumstances where you were not allowed to (e.g., during school or work hours)?	278.62	141.66	420.28	25	26	25	11	263	13	38	8.18	1
4	Original	A	Have you often been unable to stop gaming when you should have?	244.33	124.60	368.92	18	19	19	69	205	42	9	13.86	4
5	Original	A	Have you often played games for longer than intended before you started gaming?	273.56	139.11	412.67	24	24	24	155	119	46	5	7.06	6
6	IGDT-10	A	Have you ever in the past 12 months unsuccessfully tried to reduce the time spent on games?	239.11	120.88	359.99	16	15	15	15	259	27	24	19.43	3

(continued)

Table 1. Continued

No.	Source ^{c)}	Diagnostic category	Question items	WAIC	WBIC	WAIC+WBIC	Rank by WAIC	Rank by WBIC	Rank by WAIC+WBIC	Yes by person without disorder	No by person without disorder	Yes by person with disorder	No by person with disorder	Odds ratio	Yes by general population subject with disorder ($n = 7$)
7	I-CAT	B	Have you noticed that you have significantly lost interest in important activities, such as sports, hobbies, or meetings with friends or relatives because of gaming?	209.24	106.75	316.00	5	5	5	9	265	36	15	70.67	3
8	I-CAT	B	Have you limited or given up important activities, such as sports, hobbies, or meetings with friends or relatives because of gaming?	222.37	113.66	336.03	11	10	11	8	266	32	19	56.00	3
9	Original	B	Is gaming the most important part of your daily life?	234.13	118.11	352.24	14	13	14	3	271	26	25	93.95	4
10	Original	B	Do you usually give first priority to gaming when making a plan or schedule?	238.96	121.22	360.17	15	16	16	6	268	26	25	46.45	1
11	Original	B	Is gaming more important than study, work, fellowship with your friends, or family events?	245.14	125.60	370.73	20	20	20	4	270	23	28	55.45	3
12	IGDT-10	B	Have you selected gaming over meeting friends or engaging in hobbies or sports activities that you had previously enjoyed during the last 12 months?	216.82	111.80	328.62	7	9	7	0	274	16	35	∞	1

(continued)



**Table 1. Continued**

No.	Source ^{c)}	Diagnostic category	Question items	WAIC	WBIC	WAIC+WBIC	Rank by WAIC	Rank by WBIC	Rank by WAIC+WBIC	Yes by person without disorder	No by person without disorder	Yes by person with disorder	No by person with disorder	Odds ratio	Yes by general population subject with disorder (n = 7)
13	Sheehan	C	To what extent has gaming cast an unfavorable impact on your school or job performance during the past 12 months. Please encircle the matching point along the line given below.	183.51	93.38	276.89	2	1	2	20	254	42	9	59.27	4
14	Sheehan	C	To what extent has gaming cast an unfavorable impact on your social activities (meeting friends, hobbies, etc.) during the past 12 months. Please encircle the matching point along the line given below.	220.22	115.49	335.70	10	11	10	6	268	32	19	75.23	7
15	Sheehan	C	To what extent has gaming cast an unfavorable impact on your family life and roles within the family during the past 12 months. Please encircle the matching point along the line given below.	211.72	108.82	320.54	6	6	6	18	256	35	16	31.11	6
16	Original	C	Reduction in school or job performance	241.03	122.37	363.40	17	17	17	37	237	36	15	15.37	4
17	Original	C	Change of school (e.g., from ordinary senior high school to correspondence high school)	283.36	144.26	427.61	27	27	27	0	274	5	46	∞	1

(continued)

Table 1. Continued

No.	Source ^{c)}	Diagnostic category	Question items	WAIC	WBIC	WAIC+WBIC	Rank by WAIC	Rank by WBIC	Rank by WAIC+WBIC	Yes by person without disorder	No by person without disorder	Yes by person with disorder	No by person with disorder	Odds ratio	Yes by general population subject with disorder ($n = 7$)
18	Original	C	Decrease in the number of real-world friends	262.44	133.22	395.67	22	22	22	0	274	14	37	∞	1
19	Original	C	Worsened relationships with family members	229.67	116.77	346.44	12	12	12	4	270	28	23	82.17	1
20	Original	C	Difficulty in getting up in the morning (30 days or more during the past 12 months)	219.52	111.16	330.68	9	8	9	23	251	38	13	31.90	3
21	Original	C	Day/night reversal or a tendency towards it (30 days or more during past 12 months)	207.95	105.56	313.52	4	4	4	11	263	37	14	63.19	3
22	Original	C	Social withdrawal (6 months or more in total during the past 12 months)	232.43	118.70	351.13	13	14	13	2	272	26	25	141.44	1
23	Original	C	Failure to eat regularly	251.82	130.51	382.33	21	21	21	11	263	24	27	21.25	3
24	I-CAT	D	Have you continued gaming although you endangered your education or risked or lost your job because of gaming?	181.29	95.17	276.46	1	2	1	8	266	42	9	155.17	3

(continued)



**Table 1. Continued**

No.	Source ^{c)}	Diagnostic category	Question items	WAIC	WBIC	WAIC+WBIC	Rank by WAIC	Rank by WBIC	Rank by WAIC+WBIC	Yes by person without disorder	No by person without disorder	Yes by person with disorder	No by person with disorder	Odds ratio	Yes by general population subject with disorder (n = 7)
25	I-CAT	D	Have you continued gaming although it has caused or aggravated significant physical problems or illnesses (e.g. back or eye pain, headache, joint problems, fainting, significant overweight or underweight)?	266.13	135.26	401.39	23	23	23	36	238	27	24	7.44	2
26	I-CAT	D	Have you continued gaming despite experiencing mental health problems caused by gaming, e.g. becoming depressed or anxious, or experiencing problems with sleeping?	218.75	111.12	329.87	8	7	8	19	255	37	14	35.47	3
27	I-CAT	D	Have you continued gaming although your gaming costs so much money that it has caused you serious problems (e.g. through the purchase of games, add-ons, games hardware, game currency, game items, subscriptions, apps or other things)?	244.85	123.94	368.79	19	18	18	4	270	23	28	55.45	3

(continued)

Table 1. Continued

No.	Source ^{c)}	Diagnostic category	Question items	WAIC	WBIC	WAIC+	WBIC	Rank by WAIC	Rank by WBIC	Rank by WAIC+WBIC	Yes by person without disorder	No by person without disorder	Yes by person with disorder	No by person with disorder	Odds ratio	Yes by general population subject with disorder (n = 7)
28	IGDT-10	D	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)?	202.74	105.37	308.11	3	3	3	3	3	271	29	22	119.08	1

a) Eleven items affirmed (answered “Yes”) by no subject from the general population were deleted.

b) WAIC = widely applicable information criterion, WBIC = widely applicable Bayesian information criterion.

c) I-CAT = the Internet Related Disorder-Clinical Assessment Tool, Original = original questions, IGDT-10 = the Ten-Item Internet Gaming Disorder Test, Sheehan = the modified 11-grade visual analog scale of Sheehan.

specificity increased, accompanied by further improvement of the screening efficiency. When the cut-off score for a “positive” case was set at ≥ 5 (full score = 10, 9 items in total comprising the above-mentioned 8 items in addition to average gaming time, with the latter assigned a score of 0 if less than 2 h, 1 if more than 2 h but less than 6 h, and 2 if 6 h or more), the combined analysis on the general population and treatment seeking groups showed a sensitivity of 98%, specificity of 98%, and the area under the curve (AUC) was 0.981. When the analysis was confined to the general population it also showed significant screening accuracy (sensitivity 100%, specificity 98% and AUC 0.991). The positive predictive value was 91% for the combined data and 58% for the analysis confined to the general population.

Reliability and validity of the GAMES test

Evaluation of the constituent concept validity by factor analysis. When an exploratory factor analysis was conducted on the 9-item screening test, the Kaiser-Meyer-Olkin measure of sampling adequacy was as high as 0.91. The Bartlett spherical test endorsed the statistical significance ($\chi^2(36) = 1,122.1, P < 0.01$). Two factors were extracted based on the Kaiser-Guttman Rule on the number of factors. Table 4 shows the factor pattern after Promax rotation ($\kappa = 4$) following extraction of factors using the generalized least squares method.

Regarding the first of the two extracted factors, high factor loads were obtained for Items 3, 4, 6, 7, 8, and 9, suggesting that these were items associated with “prioritization of gaming and negative consequences.” In regard to the second factor, high factor loads were obtained for Items 1, 2, and 5, suggesting that these were items associated with “loss of control.” Because the proportion variance of factor 1 was far higher than that of factor 2, the former factor dominates the factor structure of this instrument.

Relationship to the IGDT-10 data. The weighted κ coefficient for the degree of agreement between the GAMES test total score and the IGDT-10 total score was 0.40, with an ICC of 0.54. The κ coefficient for the degree of agreement between GAMES test scores of ≥ 5 and IGDT-10 scores of ≥ 5 was 0.25, with an ICC of 0.25, indicating a lower degree of agreement.

Internal consistency and test-retest reliability. The alpha and omega coefficients for the GAMES test were 0.98 and 0.86, respectively. In regard to the test-retest reproducibility, both the weighed κ coefficient for GAMES test score and ICC were quite high (0.86 and 0.96, respectively).

Estimated prevalence of ICD-11 GD in the general young population. Table 5 shows the estimated prevalence of ICD-11 GD among the young population in Japan by 5-year age group and gender, based on the result of the GAMES test. The prevalence tended to be higher among younger respondents and to decrease with the advancement of age for both genders. The estimated prevalence was 7.6% (95%



Table 2. Results of multiple logistic regression analysis using presence/absence of gaming disorder serving as a dependent variable^{a)}

Item No.	ID ^{b)}	Question items	Yes by person without disorder	No by person without disorder	Yes by person with disorder	No by person with disorder	Adjusted odds ratio	95% confidence interval
1	4-A	Have you often been unable to stop gaming when you should have?	69	205	42	9	28.2	(0.75–1,063.9)
2	5-A	Have you often played games for longer than intended before you started gaming?	155	119	46	5	0.02	(0.00–0.95)
3	7-B	Have you noticed that you have significantly lost interest in important activities, such as sports, hobbies, or meetings with friends or relatives because of gaming?	9	265	36	15	17.0	(1.00–287.4)
4	9-B	Is gaming the most important part of your daily life?	3	271	26	25	667.2	(7.71–57,767)
5	16-C	Reduction in school or job performance	37	237	36	15	47.3	(1.46–1,533)
6	21-C	Day/night reversal or a tendency towards day/night reversal (30 days or more during past 12 months)	11	263	37	14	111.0	(2.28–5,402)
7	24-D	Have you continued gaming although you endangered your education or risked or lost your job because of gaming?	8	266	42	9	45.5	(1.17–1771)
8	26-D	Have you continued gaming despite experiencing mental health problems caused by gaming, e.g. becoming depressed or anxious, or experiencing problems with sleeping?	19	255	37	14	49.4	(1.51–1,621)

a) Adjusted for gender, age and presence/absence of job.

b) ID is the combination of the original number and the diagnostic category in Table 1.

Table 3. Analysis for setting the cutoff level of the draft screening test

		Gaming disorder (+)		Gaming disorder (-)		Sensitivity	Specificity	Positive predictive value	Positive Likelihood ratio	Youden Index
		Yes	No	Yes	No					
8-item screening test ^{a)}										
Analysis on combined subjects ^{c)}										
	≥2 points	51	0	98	176	1.00	0.64	0.34	2.80	0.64
	≥3 points	50	1	34	240	0.98	0.88	0.60	7.90	0.86
	≥4 points	48	3	9	265	0.94	0.97	0.84	28.65	0.91
	≥5 points	42	9	2	272	0.82	0.99	0.95	112.82	0.82
	≥6 points	33	18	0	274	0.65	1.00	1.00	∞	0.65
Analysis confined to the general population										
	≥2 points	7	0	98	176	1.00	0.64	0.07	2.80	0.64
	≥3 points	7	0	34	240	1.00	0.88	0.17	8.06	0.88
	≥4 points	6	1	9	265	0.86	0.97	0.40	26.10	0.82
	≥5 points	2	5	2	272	0.29	0.99	0.50	39.14	0.28
	≥6 points	0	7	0	274	0.00	1.00	0.00	0.00	0.00
9-item screening test (GAMES test) ^{b)}										
Analysis on combined subjects ^{c)}										
	≥3 points	51	0	61	213	1.00	0.78	0.46	4.49	0.78
	≥4 points	51	0	22	252	1.00	0.92	0.70	12.45	0.92
	≥5 points	50	1	5	269	0.98	0.98	0.91	53.73	0.96
	≥6 points	46	5	2	272	0.90	0.99	0.96	123.57	0.89
	≥7 points	39	12	0	274	0.76	1.00	1.00	∞	0.76
Analysis confined to the general population										
	≥3 points	7	0	61	213	1.00	0.78	0.10	4.49	0.78
	≥4 points	7	0	22	252	1.00	0.92	0.24	12.45	0.92
	≥5 points	7	0	5	269	1.00	0.98	0.58	54.80	0.98
	≥6 points	4	3	2	272	0.57	0.99	0.67	78.29	0.56
	≥7 points	1	6	0	274	0.14	1.00	1.00	∞	0.14

a) A screening test reflecting 4 key elements in the definition of gaming disorder.

b) A screening test including the time spent gaming on weekdays.

c) Combined subjects of the general population and treatment seekers.



Table 4. Result of explanatory factor analysis of the 9 items of GDST-9^{a)}

Item No.	Factor 1 loadings	Factor 2 loadings	Communalities
1	0.105	0.544	0.39
2	-0.157	0.739	0.43
3	0.793	-0.053	0.58
4	0.699	-0.102	0.44
5	0.172	0.485	0.42
6	0.759	-0.049	0.53
7	0.73	0.159	0.73
8	0.665	0.042	0.52
9	0.678	0.116	0.59
Proportion variance	0.44	0.06	

a) The inter-factorial correlation between factor 1 and 2 was 0.68.

Table 5. The estimated prevalence of ICD-11 gaming disorder by age class and gender^{a)}

Age class (years old)	Males % (95% CI)	Females % (95%CI)	Total % (95%CI)
10–14	9.2 (8.3–10.1)	3.1 (2.6–3.7)	6.2 (5.7–6.8)
15–19	12.0 (10.9–13.1)	3.0 (2.4–3.6)	7.6 (7.0–8.3)
20–24	6.2 (4.9–7.5)	2.3 (1.5–3.0)	4.1 (3.4–4.8)
25–29	4.0 (2.9–5.0)	1.5 (0.9–2.2)	2.7 (2.1–3.3)
Total	7.6 (6.5–8.7)	2.5 (1.9–3.2)	5.1 (4.5–5.8)

a) CI = confidence interval.

confidence interval, 6.5–8.7%) for males, 2.5% (1.9–3.2%) for females and 5.1% (4.5–5.8%) in total.

DISCUSSION

We developed a new screening test for ICD-11 GD, the GAMES test, using a reference group with GD comprising study participants from both the general population and treatment seekers. A diagnosis of GD was made based on face-to-face interviews with experts in the diagnosis and treatment of behavioral addictions using a diagnostic interview instrument newly developed for this study. The interviews and self-reported questionnaire surveys were conducted on the same day to avoid changes in target conditions (Reitsma, Rutjes, Khan, Coomarasamy, &). The question items were selected based on appropriate statistical methods from a pool of questions derived from multiple sources (Besser et al., 2019; Brandt et al., 2018; Hodgins, 2013; Király et al., 2017; Sheehan et al., 1996).

In previous studies on the development of screening and assessment tools for both IGD and GD, other approaches have been employed (Bertens et al., 2013). Several studies used a latent class analysis to assess diagnostic accuracy because a reference group assessment of IGD or GD was unavailable (Pontes et al., 2014; Lemmens et al., 2015; Király et al., 2017; Paschke et al., 2020). However, this analysis may not have captured the real target group of subjects and if the

model is incorrect, it is unclear whether the resulting estimates are meaningful (Albert & Dodd, 2004; Pepe & Janes, 2007). Other studies did not validate the cutoff scores, perhaps because question items in the instruments were matched to the diagnostic criteria of IGD (Pearcy et al., 2016; van Rooij, Schoenmakers, & van de Mheen, 2017). Other instruments, have been developed, primarily to assess the severity of the disorder and its consequences, rather than for the purpose of making a diagnosis (Pontes et al., 2019; Pontes & Griffiths, 2015).

The GAMES test showed extremely high sensitivity and specificity for subjects, who were representative of the general young population in Japan. Positive predictive values were also high, not only for the general population but also for the combined subject groups. The results suggest that if this screening test were applied to the general population, about 60% of the individuals judged as positive would have GD. The GAMES test showed high internal consistency with the alpha and omega coefficients standing at well over 0.8. In addition, the test-retest reliability of the test was high, demonstrating that the GAMES test evaluation was highly reproducible.

Exploratory factor analysis revealed a two-factor solution to best model the nine items of the GAMES test. The first factor was a major element, explaining approximately 44% of the portion variance. This factor represents the prioritization of gaming and negative consequences or functional impairment. The prioritization of gaming, which can be described as “salience” is one of the most important characteristics of addictions and is the core feature of the clinical manifestations of GD. Two out of four key elements of the GD definition are related to negative consequences. One of the unique features of the GD definition is the inclusion of a functional impairment item which must be met in order to make a diagnosis. It seeks to avoid the overdiagnosis of GD, which is common to polythetic approaches such as the IGD criteria (Billieux, Schimmenti, Khazaal, Maurage, & Heeren, 2015; Billieux et al., 2017). For these reasons, it is appropriate that these items constitute the main factor in the GAMES test. The second factor comprises three items mainly related to loss of control over gaming.

A recently-developed instrument, the GADIS-A, comprising ten items, also has a two-factor solution (Paschke et al., 2020). The main factor of GADIS-A is “impending or manifest consequences due to gaming behavior”, which may be construed as reflecting priority of gaming and impairment, and the second, is “pathological gaming behavior”. The values of portion variances for the two factors are almost equal in the case of GADIS-A (Paschke et al., 2020). The structure of another screening tool for GD - GDT - showed a one-factor solution with almost the same variance value as factor 1 of the GAMES test (Pontes et al., 2019).

Concurrent validity was examined using the Japanese version of IGDT-10 (Király et al., 2019). The reasons for using this instrument were two-fold: 1) assessment or screening instruments for ICD-11 GD were unavailable when this study was conducted, and 2) IGDT-10 has been widely

used internationally as a screening instrument for IGD (Király et al., 2019). The IGDT-10 was translated into Japanese and the accuracy of translation confirmed by comparing the original English version and the English version that was retranslated from the Japanese version, but validation of the instrument has not been conducted using Japanese samples. Nonetheless, the Japanese version has already been used in previous studies (Kinjo & Osaki, 2019; Nakayama, Matsuzaki, Mihara, Kitayuguchi, & Higuchi, 2020). The results of this study showed that correlation between the GAMES test and IGDT-10 was too low. This finding should not be interpreted as a low concurrent validity of the GAMES test, rather that the diagnoses of GD and IGD may have different constructs. Another reason may be related to the use of the unvalidated Japanese version of IGDT-10 in the study. Low concordance of the two diagnoses in the same clinical samples has been reported. For example, the rates of subjects with IGD who were also diagnosed as having GD were 64% in a study in Taiwan and as low as 16% in a study in Korea (Jo et al., 2019; Ko et al., 2019).

To the authors' knowledge, this study showed for the first time the estimated prevalence of ICD-11 GD among the general young population (Darvesh et al., 2020). As previously mentioned, the prevalence of IGD has been estimated in a number of studies in different countries. Findings were quite diverse, depending on the characteristics of study samples, screening instruments used, and method of data collection (Chia et al., 2020; Darvesh et al., 2020; Fam, 2018; Mihara & Higuchi, 2017; Saunders et al., 2017;). The estimated prevalence of GD among representative Japanese young people in this study was 5.1%, which is comparable to a pooled estimate of problematic gaming among young people, derived from a meta-analysis of 16 studies conducted globally (Fam, 2018). The estimated prevalence of GD also showed a male preponderance and a higher tendency among adolescents compared to older generations, which is consistent with previous studies on the prevalence of IGD (Darvesh et al., 2020; Mihara & Higuchi, 2017). This study provides a reliable tool to estimate the prevalence of GD, which is expected to serve as a basis on which measures against the disorder will be implemented in different jurisdictions.

Methodological limitations

Finally, the methodological limitations of this study should be summarized. Firstly, the sample size of interviewees and the comparatively low prevalence of GD in the general population resulted in a small participant group diagnosed as having GD. To increase the statistical strength, we added treatment seekers with GD to the reference group. We separately analyzed the data of the seven participants that had GD and those of the combined 51 participants with GD in the process of developing the GAMES test. Secondly, because validated interview schedules for the diagnosis of GD are not available, we devised a schedule and administered it to interviewees. In addition, the interview was conducted by one interviewer, although each interviewer had clinically seen numerous patients with GD. Recently, there

has been a debate in the field around whether GD/IGD is a formative construct (van Rooij, Looy, & Billieux, 2017). GD has a formative construct feature in which the four categories (A through D) are essential in order to define GD. On the contrary, the GAMES test is a reflective construct, not requiring all four categories to be met for screening purposes. Fourthly, with regard to the evaluation of concurrent validity, an additional study using appropriate instruments may be necessary in the future. Finally, the GAMES test was developed solely utilizing Japanese participants. In this context and considering the small sample size of the reference group, this study can be regarded as a pilot study, and the validity and reliability of this instrument needs to be examined in future studies using larger samples and other ethnic groups with different gaming cultures.

CONCLUSION

In this study, a new screening test for ICD-11 GD, consisting of 9 items (GAMES test) was developed utilizing a reference group with GD and a non-GD general young population. A diagnosis of GD was made based on face-to-face interviews with experts in behavioral addictions using a diagnostic interview instrument newly developed for this study. The question items for the GAMES test were selected employing appropriate statistical procedures from a pool of questions derived from multiple sources contained in a self-reported questionnaire conducted on the same day as the face-to-face interviews. The test showed high sensitivity, specificity, positive predictive values, and internal consistency, and is highly reproducible. It consists of one main factor representing the salience and negative consequences of GD and a highly correlated second factor related to loss of control.

The GAMES test allowed us, for the first time, to estimate the prevalence of GD among subjects representing the general young population in Japan. The estimated prevalence was 5.1% which is comparable to the pooled estimate of problematic gaming derived from a meta-analysis of studies conducted globally. Consistent with previous studies, it showed a male preponderance and higher tendency among adolescents compared to older generations.

Funding sources: This study was funded by a research grant on addiction given to the National Center for Addiction Services Administration and subsidized by the Ministry of Health, Labour and Welfare of Japan.

Authors' contributions: SH, YO, AK, HJR and JBS developed the conception and design of the study. All authors contributed to the collection, analysis, and interpretation of data. SH and YO drafted the article. All authors contributed to the revision and approved the final version of the article. All authors had full access to the data in this study, and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of interest disclosure: All authors report no financial conflict or other relationship relevant to the subject of this article. SH, SM, MM, TK, TM, HN and JBS have been engaged in the treatment of disorders due to substance use and addictive behaviors including gaming disorder.

SH, SM, HJR and JBS are members of the World Health Organization (WHO) Work Group on “Behavioural Disorders Associated with Excessive Use of the Internet, Computers, Smartphones and Similar Electronic Devices”, which was established in 2014. They have been involved in developing and reviewing the diagnostic guidelines of gaming disorder in the ICD-11. The preliminary findings of the present study were presented to the Work Group in 2019. The views expressed in this publication are those of the authors and do not necessarily represent the decisions or policies of the WHO.

Acknowledgments: This study was funded by a research grant on addiction given to Kurihama Medical and Addiction Center and subsidized by the Ministry of Health, Labour and Welfare of Japan. The authors wish to thank Ms. Kiyoko Hidaka, Mr. Satoshi Furuno, Ms. Naoko Endo, Ms. Haruka Miyakoshi and Ms. Moemi Shibasaki for contributing to the data collection for this study.

SUPPLEMENTARY MATERIAL

Supplementary data to this article can be found online at <https://doi.org/10.1556/2006.2021.00041>.

REFERENCES

- Albert, P. S., & Dodd, L. E. (2004). A cautionary note on robustness of latent class models for estimating diagnostic error without a gold standard. *Biometrics*, *60*, 427–435. <http://doi.org/10.1111/j.0006-341X.2004.00187.x>.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*, 5th ed. (DSM-5). Arlington, VA: American Psychiatric Association.
- Bantis, L. E., Nakas, C. T., & Reiser, B. (2014). Construction of confidence regions in the ROC space after the estimation of the optimal Youden index-based cut-off point. *Biometrics*, *70*, 212–223. <http://doi.org/10.1111/biom.12107>.
- Bertens, L. C. M., Broekhuizen, B. D. L., Naaktgeboren, A., Rutten, F. H., Hoes, A. W., Van Mourik, Y., . . . Restsma, J. B. (2013). Use of expert panels to define the reference standard in diagnostic research: A systematic review of published methods and reporting. *Plos Medicine*, *10*, e1011531. <http://doi.org/10.1371/journal.pmed.1001531>.
- Besser, B., Loerbroks, L., Bischof, G., Bischof, A., & Rumpf, H. J. (2019). Performance of the DSM-5-based criteria for Internet addiction: A factor analytical examination of three samples. *Journal of Behavioral Addiction*, *8*, 288–294. <http://doi.org/10.1556/2006.8.2019.19>.
- Billieux, J., King, D., Higuchi, S., Achab, S., Bowden-Jones, H., Hao, W., . . . Poznyak, V. (2017). Functional impairment matters in the screening and diagnosis of gaming disorder. *Journal of Behavioral Addictions*, *6*, 285–289. <http://doi.org/10.1556/2006.6.2017.036>.
- Billieux, J., Schimmenti, A., Khazaal, Y., Maurage, P., & Heeren, A. (2015). Are we overpathologizing everyday life? A tenable blueprint for behavioral addiction research. *Journal of Behavioral Addictions*, *4*, 119–123. <http://doi.org/10.1556/2006.4.2015.009>.
- Brandt, D., Glanert, S., Bischof, G., Besser, B., Bischof, A., & Rumpf, H. J. (2018). Determination of the test-retest reliability of a computerized diagnostic interview for internet-related disorders. *Journal of Behavioral Addiction*, *7*(Suppl. 1), 47.
- Chia, D. X. Y., Ng, C. W. L., Kandasami, G., Seow, M. Y. L., Choo, C. C., Chew, P. K. H., . . . Zhang, M. W. B. (2020). Prevalence of internet addiction and gaming disorder in Southeast Asia: A meta-analysis. *International Journal of Environmental Research and Public Health*, *17*, 2582. <http://doi.org/10.3390/ijerph17072582>.
- Darvesh, N., Radhakrishnan, A., Lachance, C. C., Ninic, V., Sharpe, J. P., Ghassemi, M., . . . Tricco, A. C. (2020). Exploring the prevalence of gaming disorder and internet gaming disorder: A rapid scoping review. *Systematic Reviews*, *9*, 68. <http://doi.org/10.1186/s13643-020-01329-2>.
- Fam, J. Y. (2018). Prevalence of Internet gaming disorder in adolescents: A meta-analysis across three decades. *Scandinavian Journal of Psychology*, *59*, 524–531. <http://doi.org/10.1111/sjop.12459>.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective*. New Jersey: Pearson Prentice Hall.
- Hodgins, D. C. (2013). Reliability and validity of the Sheehan disability scale modified for pathological gambling. *BMC Psychiatry*, *13*, 177. <http://doi.org/10.1186/1471-244X-13-177>.
- IBM Corp. (2017). *IBM SPSS statistics for windows, version 25.0*. NY: IBM Corp.
- Jo, Y. S., Bhang, S. Y., Choi, J. S., Lee, H. K., Lee, S. Y., & Kweon, Y. S. (2019). Clinical characteristics of diagnosis of internet gaming disorder: Comparison of DSM-5 IGD and ICD-11 GD diagnosis. *Journal of Clinical Medicine*, *8*, 945. <http://doi.org/10.3390/jcm8070945>.
- Jo, S.-J., Yim, H. W., Lee, H.-K., Lee, H.-C., Choi, J.-S., & Baek, K.-Y. (2018). The internet gaming use-elicited symptom screen proved to be a valid tool for adolescents aged 10–19 years. *Acta Paediatrica*, *107*, 511–516. <http://doi.org/10.1111/apa.14087>.
- King, D., Chamberlain, S. R., Carragher, N., Billieux, J., Stein, D., Muller, K., . . . Delfabbro, P. H. (2020). Screening and assessment tools for gaming disorder: A comprehensive systematic review. *Clinical Psychology Review*, *77*, 101831. <http://doi.org/10.1016/j.cpr.2020.101831>.
- King, D., & Delfabbro, P. (2019). *Internet gaming disorder: Theory, assessment, treatment, and prevention*. London: Academic Press.
- Kinjo, A., & Osaki, Y. (2019). Epidemiology related to gaming disorder. *Journal of Clinical and Experimental Medicine (IGAKU NO AYUMI)*, *271*, 567–571. (In Japanese).
- Király, O., Bőthe, B., Ramos-Diaz, J., Rahimi-Movaghar, A., Lukavska, K., Hrabec, O., . . . Demetrovics, Z. (2019). Ten-Item

- Internet gaming disorder test (IGDT-10): Measurement invariance and cross-cultural validation across seven language-based samples. *Psychology of Addictive Behaviors*, 33, 91–103. <http://doi.org/10.1037/adb0000433>.
- Király, O., Slezcka, P., Pontes, H. M., Urbán, R., Griffiths, M. D., & Demetrovics, Z. (2017). Validation of the ten-item internet gaming disorder test (IGDT-10) and evaluation of the nine DSM-5 internet gaming disorder criteria. *Addictive Behaviors*, 64, 253–260. <http://doi.org/10.1016/j.addbeh.2015.11.005>.
- Ko, C. H., Lin, H. C., Lin, P. C., & Yen, J. Y. (2019). Validity, functional impairment and complications related to Internet gaming disorder in the DSM-5 and gaming disorder in the ICD-11. *Australian & New Zealand Journal of Psychiatry*, 48(67419881499). <http://doi.org/10.1177/0004867419881499>.
- Lemmens, J. S., Valkenburg, P. M., & Gentile, D. A. (2015). The internet gaming disorder scale. *Psychological Assessment*, 27, 567–582. <http://doi.org/10.1037/pas0000062>.
- Malhotra, N. K. (1999). *Marketing research: An applied orientation*. New Jersey: Prentice Hall.
- Mihara, S., & Higuchi, S. (2017). Cross-sectional and longitudinal epidemiological studies of Internet gaming disorder: A systematic review of the literature. *Psychiatry and Clinical Neurosciences*, 71, 425–444. <http://doi.org/10.1111/pcn.12532>.
- Nakayama, H., Matsuzaki, T., Mihara, S., Kitayuguchi, T., & Higuchi, S. (2020). Relationship between problematic gaming and age at the onset of habitual gaming. *Pediatrics International*, 65, 1275–1281. <http://doi.org/10.1111/ped.14290>.
- Paschke, K., Austermann, M. S., & Thomasius, R. (2020). Assessing ICD-11 gaming disorder in adolescent gamers: Development and validation of the gaming disorder scale for adolescents (GADIS-A). *Journal of Clinical Medicine*, 9, 993. <http://doi.org/10.3390/jcm9040993>.
- Pearcy, B. T. D., Roberts, L. D., & McEvoy, P. M. (2016). Psychometric testing of the personal internet gaming disorder evaluation-9: A new measure designed to assess internet gaming disorder. *Cyberpsychology, Behavior, and Social Networking*, 19, 335–341. <http://doi.org/10.1089/cyber.2015.0534>.
- Pepe, M. S., & Janes, H. (2007). Insights into latent class analysis of diagnostic test performance. *Biostatistics*, 8, 474–484. <http://doi.org/10.1093/biostatistics/kxl038>.
- Pontes, H. M., & Griffiths, M. D. (2015). Measuring DSM-5 internet gaming disorder: Development and validation of a short psychometric scale. *Computers in Human Behavior*, 45, 137–143. <http://dx.doi.org/10.1016/j.chb.2014.12.006>.
- Pontes, H. M., Schivinski, B., Sindermann, C., Li, M., Becker, B., Zhou, M., & Montag, C. (2019). Measurement and conceptualization of gaming disorder according to the world Health Organization framework: The development of the gaming disorder test. *International Journal of Mental Health Addiction*. <http://dx.doi.org/10.1007/s11469>.
- Reitsma, J. B., Rutjes, A. W., Khan, K. S., Coomarasamy, A., & Bossuyt, P. M. (2009). A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *Journal of Clinical Epidemiology*, 62, 797–806. <http://doi.org/10.1016/j.jclinepi.2009.02.005>.
- Saunders, J. B., Hao, W., Long, J., King, D. L., Mann, K., Fauth-Büler, M., ... Poznyak, V. (2017). Gaming disorder: Its delin- eation as an important condition for diagnosis, management, and prevention. *Journal of Behavioral Addictions*, 6, 271–279. <https://doi.org/10.1556/2006.6.2017.039>.
- Sheehan, D. V., Harnett-Sheehan, K., & Raj, B. A. (1996). The measurement of disability. *International Clinical Psychopharmacology*, 11(Suppl. 3), 89–95.
- Statistics Bureau of Japan. *Current population estimates as of October 1st, 2018*. Tokyo: Statistics Bureau of Japan. Retrieved from <https://www.stat.go.jp/english/data/jinsui/2.html>.
- van Rooij, A. J., Looy, J. V., & Billieux, J. (2017). Internet gaming disorder as a formative construct: Implications for conceptualization and measurement. *Psychiatry and Clinical Neurosciences*, 71, 445–458. <https://doi.org/10.1111/pcn.12404>.
- van Rooij, A. J., Schoenmakers, T. M., & van de Mheen, D. (2017). 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’. *Addictive Behaviors*, 64, 269–274. <https://doi.org/10.1016/j.addbeh.2015.10.018>.
- Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 3571–3594.
- Watanabe, S. (2013). A widely applicable Bayesian information criterion. *Journal of Machine Learning Research*, 14, 867–897.
- World Health Organization. (2016). *Behavioural disorders associated with excessive use of the Internet, computers, smartphones and similar electronic devices: Clinical descriptions, diagnostic guidelines and priorities for international research: Meeting report*. Geneva, Switzerland: World Health Organization.
- World Health Organization. (2019). *ICD-11 for mortality and morbidity statistics 2019*. Retrieved from <https://icd.who.int/browse11/l-m/en>.

APPENDIX

The Gaming Engagement Screener (GAMES) test

Please read questions 1 to 8 and select an appropriate answer (Yes or No) for each question in relation to your gaming

behavior over the last 12 months. For question 9, please select the most appropriate answer. Gaming means games played on smartphones, gaming devices, personal computers, or other similar devices.

No.	Question items	Answer	
		Yes	No
1	Have you often been unable to stop gaming when you should have?	1	0
2	Have you often played games for longer than intended before you started gaming?	1	0
3	Have you noticed that you have significantly lost interest in important activities, such as sports, hobbies, or meeting with friends or relatives because of gaming?	1	0
4	Is gaming the most important part of your daily life?	1	0
5	Has your school or job performance deteriorated because of gaming?	1	0
6	Have you experienced day/night reversal or a tendency towards day/night reversal because of gaming (30 days or more during the past 12 months)?	1	0
7	Have you continued gaming although you endangered your education or risked or lost your job because of gaming?	1	0
8	Have you continued gaming despite experiencing mental health problems caused by gaming e.g. becoming depressed or anxious, or experiencing problems with sleeping?	1	0
9	Approximately how many hours do you spend gaming on a typical weekday?	0: Less than 2 h 1: At least 2 h but less than 6 h 2: 6 h or longer	

Evaluation method: A total score of 5 or more obtained by totaling the scores for all the items suggests the presence of the ICD-11 gaming disorder.

