Psychometric validation of the Internet Gaming Disorder-20 Test among Chinese middle school and university students

SHU M. YU, IVAN JACOB AGALOOS PESIGAN, MENG XUAN ZHANG and ANISE M. S. WU*
Department of Psychology, Faculty of Social Sciences (FSS), University of Macau, Avenida da Universidade, Macao, China

(Received: May 24, 2018; revised manuscript received: October 24, 2018; second revised manuscript received: February 15, 2019; third revised manuscript received: March 19, 2019; accepted: April 6, 2019)

Background and aims: Internet gaming disorder (IGD) was proposed in Diagnostic and Statistical Manual of Mental Disorders of American Psychiatric Association as an area warranting more research attention. High prevalence of excessive Internet game use and related addictions has been reported in China, especially among youth; however, there is a lack of psychometrically and theoretically sound instruments for assessing IGD in the Chinese language. Methods: This study aimed to examine the psychometric properties of a Chinese version of the Internet Gaming Disorder Test (IGD-20 Test) among Chinese middle school and university students. Results: Confirmatory factor analysis results showed that the Chinese version of the IGD-20 Test had five factors (i.e., salience-tolerance, mood modification, withdrawal, conflict, and relapse). Measurement invariance was confirmed across the two samples. The test score was positively associated with the modified Young’s Internet Addiction Test for gaming addiction. Concurrent validation was further demonstrated by the IGD-20 Test’s positive correlation with weekly gameplay and depression symptoms. The latent profile analysis showed four different gamer classes (i.e., regular gamers, low-risk engaged gamers, high-risk engaged gamers, and probable disordered gamers), with the estimated prevalence of 2.1% of the last group. Discussion and conclusion: The IGD-20 Test was applicable to Chinese youth and its Chinese version generally demonstrated good psychometric properties.

Keywords: IGD-20 Test, online game, addiction, assessment, youth

INTRODUCTION

In the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association [APA], 2013), Internet gaming disorder (IGD) has been included in the “Emerging Measures and Models” section as an area warranting more research attention. World Health Organization (2019) also included gaming disorder as an addictive disorder in the 11th revision of International Classification of Diseases, with major symptoms including “impaired control over gaming,” “increasing priority given to gaming,” and “continued gaming despite of negative consequences.” These symptoms are generally consistent with the DSM-5 criteria [including preoccupation, withdrawal symptoms, tolerance, relapse (failure of cutting off), loss of interest in non-gaming activities, conflicts, and adverse consequences], but DSM-5 specifically states deception and mood modification as possible symptoms. Although the inclusion of gaming disorder as a formal mental disorder remains controversial due to potential stigmatization of a pastime entertainment activity (van Rooij et al., 2018), such inclusion draws both clinical and public attention to problematic gaming and facilitates further research on assessment tools, treatment, and prevention (Griffiths, Kuss, Lopez-Fernandez, & Pontes, 2017; Rumpf et al., 2018; van den Brink, 2017).

Accumulating evidence suggests that this behavioral addiction has become an increasingly prolific public health issue in Chinese communities, especially among youth (APA, 2013). A high prevalence estimate was reported in not only middle-school students [e.g., 15.5%; assessed by Huang’s Online Gaming Addiction Scale (HOGAS); Yu et al., 2010], but also in university students [e.g., 27.2% dependents and 14.3% addicts; assessed by Online Game Addiction Index (OGAI); Zhou & Li, 2009]. However, the inventories used did not consider all DSM-5 criteria for IGD; therefore, the estimated prevalence might not accurately reflect the actual situation.

Common inventories used for assessing IGD and similar disorders (e.g., problematic video gaming) in previous literature have included: (a) a symptom list adapted from the DSM criteria for other addictions (e.g., substance dependence and pathological gambling; Lemmens,
To demonstrate the test (Byrne, Shavelson, & Muthén, 1989) and thus establishing factorial invariance (at least partial scalar invariance) holds factors (e.g., latent means) can only be carried out when equation modeling. Comparison of attributes of the common university student sample). It is the samples (one middle-school student sample and one mainland China have stricter regulations for the school activities for students at the graduating grade who are preparing for high-school entrance examination. The middle-school sample was composed of 569 participants who were 11- to 16-year-old (M = 13.34, SD = 0.68; female = 46.2%). The university sample was composed of 523 participants who were 17- to 28-year-old (M = 20.12, SD = 1.73; female = 48.4%). All of them reported that they had played online game before (lifetime experience). University students spent significantly more time (M = 11.11 hr), but not expense, on online gaming than the middle-school students (M = 4.47; t = −8.49; p < .001). All the descriptive statistics are presented in Table 1.

Measures

Participants were asked to state their gender, age, year of study, and weekly online gameplay time [i.e., “number of hours on a weekly basis?” (fill-in-the-blank) and expense [i.e., “amount (in CNY) on a weekly basis?” (fill-in-the-blank)]. Participants also responded to the following scales, in which a higher score indicated a higher level of corresponding construct measured.
Chinese IGD-20 Test validation

Table 1. Descriptive statistics for the measures

|                      | Middle school | University | Total     | Group difference
|----------------------|---------------|------------|-----------|------------------|
|                      | n  | %   | n  | %   | n  | %   | χ²  |
| Total                | 302/263 | 53.1/46.2 | 270/253 | 51.6/48.4 | 572/516 | 52.4/47.3 | 0.363 |
| Gender               | Male/female | M  | SD  | M  | SD  | M  | SD  | t  |
| Age                  | 13.34 | 0.68  | 20.12 | 1.73  | 16.59 | 3.63  | -83.63*** |
| Weekly expense       | 28.46 | 426.90 | 60.44 | 505.00 | 43.74 | 465.90 | -1.12  |
| Weekly game time     | 4.47  | 9.65  | 11.11 | 14.99  | 7.65  | 12.90  | -8.49*** |
| IGD-20 Test          | 39.76 | 12.57 | 43.26 | 13.72  | 41.44 | 13.25  | -4.37*** |
| YIAT                 | 40.22 | 15.46 | 42.90 | 16.03  | 41.53 | 15.79  | -2.72** |
| Depression           | 9.31  | 9.40  | 9.86  | 9.10   | 9.61  | 9.23   | -0.90   |

Note: SD: standard deviation; IGD-20 Test: 20-item Internet Gaming Disorder Test; YIAT: Young Internet Addiction Test.

*Four missing cases. **In Chinese Yuan (CNY); 1 CNY ≈ 0.16 USD.

**p < .01. ***p < .001 for group differences across middle school and university groups.

Chinese version of the IGD-20 Test. The Chinese version of the IGD-20 Test (see Appendix) was self-administered by the participants to assess their IGD symptom severity. The test has a 5-point Likert response scale (ranging from 1 = strongly disagree to 5 = strongly agree). After recoding the two reverse items (Items 2 and 19), a higher total score represents a higher level of addictive tendency toward gaming.

Modified YIAT. The 20-item YIAT (Young, 1996) was developed for assessing Internet addiction, but was modified to assess different Internet-related addictions (e.g., social networking addiction; Yu, Wu, & Pesigan, 2016). The modified YIAT has various language versions including English, French, Arabic, and Chinese (King et al., 2013). The YIAT, with a 5-point Likert response scale ranging from 1 = not at all to 5 = always, showed good reliability in both middle school and university groups (α = .93 and .95, respectively).

Depression Anxiety Stress Scales-21 (DASS-21). The 7-item depression subscale of the Chinese version of the DASS-21 (Moussa, Lovibond, & Laube, 2001) was used to measure depression symptoms (ranging from 0 = did not apply to me at all to 3 = applied to me very much). This scale showed good reliability in both middle school and university groups (α = .89 and .90, respectively).

Statistical analyses

Confirmatory factor analysis (CFA) was conducted using Lavaan package in R (Rosseel, 2012) using maximum likelihood estimation to test the structural model. Before running the CFA, preliminary exploratory data analysis showed that missing data patterns did not appear to be systematic, and the majority of the respondents completed the entire survey (84.2%). Missing data were handled using maximum likelihood in the measurement models (i.e., full information maximum likelihood). Multisample modeling analysis was conducted with both samples simultaneously to investigate the measurement invariance of the IGD-20 Test. The goodness of fit of the model was considered acceptable and preferable if the comparative fit index (CFI) and the Tucker–Lewis fit index (TLI) were greater than 0.90 (Hu & Bentler, 1999), and both the root mean square error of approximation [RMSEA with its 90% confidence interval (90% CI)] and the standardized root mean square residual (SRMR) were less than 0.08 (Browne & Cudeck, 1993; Hu & Bentler, 1999).

To examine the model invariance across the middle school and university samples, we conducted further tests on metric (factor loading) invariance and scalar (measurement intercept) invariance in the model with both samples. The baseline configural invariance model was fitted first, which constrained the factor structure to be equal across the two samples, while allowing the factor loadings and measurement intercepts to be freely estimated. Subsequently, the two more restrictive invariance models were fitted to evaluate measurement invariance across samples, and model comparisons were performed with the likelihood ratio test, adjusting for scaling (Satorra & Bentler, 2001) and the change in CFI. Two out of the following criteria being met indicated that the more restrictive version of invariance does not hold (Chen, 2007; Cheung & Rensvold, 2002; Vandenberg & Lance, 2000): (a) the χ² difference test is significant (p < .05), (b) the change in CFI ≥ 0.01, and (c) the change in TLI ≥ 0.02.

Concurrent validity was tested by correlational analysis, whereas Cronbach’s αs were computed to ensure the internal consistency of the (sub)scale.

Since different types of gamers are expected to exhibit different configurations of IGD symptoms (Pontes et al., 2014), we used latent profile analysis in Mplus 7.11 (Muthen & Muthen, 1998–2012) to test models with one to six latent classes. Multiple imputations (100 iterations) using multivariate imputation by chained equations (van Buuren & Groothuis-Oudshoorn, 2011) were used to deal with missing data prior to creating aggregate factor scores. The YIAT score, weekly gameplay time and amount, age, and gender were used as auxiliary variables in the process. To determine the correct number of latent classes, we considered: (a) the models with lower values on the fit indices (i.e., lower values on Akaike’s information criterion, Bayesian information criterion (BIC), and sample-size adjusted BIC refer to better fit), (b) entropy or the standardized summary measure of the classification accuracy of placing participants into classes (i.e., higher entropy values reflect a
better classification of individuals), and (c) the Lo–Mendell–Rubin adjusted likelihood ratio test (i.e., non-significant test value suggests that the model with one fewer class is a better explanation of the data; Lo, Mendell, & Rubin, 2001). These guidelines were used in conjunction with previous theory to identify which of the models fit the data and substantive theory adequately.

**Ethics**

All participants provided informed consent for their participation, while parents’ permission was also obtained for those less than 18 years of age. The procedures, including data collection and analyses, were carried out in accordance with the Declaration of Helsinki. The ethical approval for this study (ref. no.: MYRG2016-00162-FSS) was also obtained from the Ethics Committee of University of Macau.

**RESULTS**

**Confirmatory factor analysis**

We first tested the one-factor model \((N = 1092)\), but the model of fit was unsatisfactory, \(\chi^2(170) = 1,249.58, p < .001; \text{CFI} = 0.887; \text{TLI} = 0.874; \text{RMSEA} = 0.077, 90\% \text{CI} [0.073, 0.081]; \text{SRMR} = 0.047.\) We then tried to replicate the original six-factor model (i.e., salience, mood modification, tolerance, withdrawal symptoms, conflict, and relapse; Griffiths, 2005) but, the Psi matrix, which is the variance–covariance matrix of the latent variables, was not positive definite. It resulted in model mis-specification and further test could not be conducted on this specified model. Hierarchical factor model and bi-factor model testing also showed the same problem. The model mis-specification was probably a result of very high correlations between the latent variables.

Given the consistently high correlations between “salience” and “tolerance” factors in previous validation studies \((r \geq .94; \text{Fuster et al., 2016; Pontes et al., 2014})\), we tested another six-factor model in which these two factors reflected a higher-order factor of “engagement.” However, “salience” showed negative variance. We therefore tested a five-factor model combining the items of these two factors. This five-factor model showed an acceptable model fit, \(\chi^2(160) = 778.507, p < .001; \text{CFI} = 0.935; \text{TLI} = 0.923; \text{RMSEA} = 0.060, 90\% \text{CI} [0.056, 0.064]; \text{SRMR} = 0.036.\) All factor loadings, except for Items 2 (0.238) and 19 (0.013), were satisfactorily high (0.392–0.818; Figure 1).

**Factorial invariance across the middle school and university samples**

Table 2 shows all the factor loadings of the five-factor model, which were similar across samples as in the whole group.

![Figure 1. Confirmatory factor analysis diagram](image-url)
Chinese IGD-20 Test validation

Table 2. Factor loadings of the IGD-20 Test and Cronbach’s α of the subfactors

| Subfactor                        | Middle school (α = .90) | University (α = .92) | All (α = .90) |
|----------------------------------|------------------------|----------------------|--------------|
| Salience-tolerance (α)           |                        |                      |              |
| IGD1                             | .820                   | .860                 | .840         |
| IGD2                             | .535                   | .651                 | .602         |
| IGD3                             | .601                   | .584                 | .595         |
| IGD4                             | .754                   | .775                 | .741         |
| IGD5                             | .808                   | .767                 | .788         |
| IGD6                             | .571                   | .750                 | .663         |
| IGD7                             | .661                   | .719                 | .697         |
| Mood modification (α)            |                        |                      |              |
| IGD2                             | .570                   | .610                 | .590         |
| IGD8                             | .198                   | .288                 | .238         |
| IGD9                             | .712                   | .645                 | .678         |
| IGD14                            | .851                   | .895                 | .874         |
| Withdrawal (α)                   |                        |                      |              |
| IGD4                             | .820                   | .820                 | .820         |
| IGD10                            | .727                   | .761                 | .760         |
| IGD16                            | .764                   | .813                 | .792         |
| Conflict (α)                     |                        |                      |              |
| IGD5                             | .430                   | .580                 | .520         |
| IGD11                            | .588                   | .782                 | .710         |
| IGD17                            | .576                   | .600                 | .605         |
| IGD19                            | .481                   | .659                 | .593         |
| IGD20                            | .014                   | .005                 | .013         |
| Relapse (α)                      |                        |                      |              |
| IGD6                             | .317                   | .479                 | .392         |
| IGD12                            | .794                   | .770                 | .781         |
| IGD18                            | .395                   | .772                 | .811         |

Note. IGD-20 Test: 20-item Internet Gaming Disorder Test.

As shown in Table 3, the configural invariance model showed a good fit, χ²(320) = 1,011.777, p < .001; RMSEA = 0.063; CFI = 0.929; SRMR = 0.040, indicating the same factor structure found in both samples. The metric invariance model also had a good fit, χ²(335) = 1,061.435; RMSEA = 0.063; CFI = 0.925; SRMR = 0.050. The scaled Δχ²(15) = 49.658, p < .001 and ΔCFI = 0.004. Although the likelihood ratio test was significant at the 0.01 level, there was little change in CFI (i.e., <0.01), supporting that the metric invariance model held across samples. For the scalar invariance model, the scaled Δχ²(15) = 137.422, p < .001. ΔCFI was 0.012. Although these values did not meet the suggested criteria (e.g., ΔCFI = 0.010), other indices met the standards of a good model fit, χ²(350) = 1,198.857; RMSEA = 0.067; CFI = 0.913; SRMR = 0.055. Therefore, we concluded that the full scalar invariance model marginally held. Our results indicated that both factor loadings and measurement intercepts were equal across the middle school and university samples.

Reliability and concurrent validity

Table 2 showed that the reliability of the IGD-20 Test was high. Its subscales also had satisfactory α, except the

Table 3. Tests of invariance of the IGD-20 Test

| Model             | χ²    | df | CFI  | TLI  | RMSEA [90% CI]       | SRMR | Δχ²   | Δdf   | ΔCFI  | ΔTLI  | ΔRMSEA |
|-------------------|-------|----|------|------|----------------------|------|-------|-------|-------|-------|--------|
| Configural        | 1,011.777 | 320 | 0.929| 0.916| 0.063 [0.059, 0.068]  | 0.040|       |       |       |       |        |
| Metric            | 1,061.435 | 335 | 0.925| 0.915| 0.063 [0.059, 0.068]  | 0.050| 49.658***| 15 | 0.004 | 0.001 | 0.000  |
| Scalar            | 1,198.857 | 350 | 0.913| 0.905| 0.067 [0.063, 0.071]  | 0.054| 137.422***| 15 | 0.012 | 0.010 | 0.004  |

Note. A subsequent model was evaluated to have significant change in goodness of fit than the previous model when two of the three criteria are met: the χ² difference test was significant (p < .05), the change in CFI was higher than or equal to 0.01 and the change in the TLI was higher than or equal to 0.02 (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000). CFI: comparative fit index; TLI: Tucker–Lewis index; RMSEA: root mean square error of approximation; SRMR: standardized root mean square residual; CI: confidence interval; IGD-20 Test: 20-item Internet Gaming Disorder Test.

***p < .001.
Conflict subscale (α = .52; due to the low factor loading of Item 19).

Concurrent validity of the IGD-20 Test was shown by its strong positive association with YIAT score, depression symptoms, weekly gameplay hours, and gaming expense. The total IGD-20 Test scores had strong and positive correlations with YIAT in overall, middle school, and university samples (r = .87, .86, and .88, respectively; p < .001). It also moderately and positively correlated with depression symptoms (r = .38, .28, and .45, respectively; p < .001). It was positively correlated with gameplay hours but the correlation strength varied from mild to moderate (r = .40, .35, and .43, respectively; p < .001). In terms of gaming expense, the correlation was mild and statistically significant in only overall and middle-school samples (r = .11; p < .001 and r = .13; p < .01, respectively), but not university sample (r = .09; p = .056).

Latent profile analysis

Table 4 showed that the higher the number of class specified, the lower the fit indices. Therefore, we relied on other criteria to determine the appropriate model. The models of two, three, four, five, and six classes had good entropy values, with the four- and six-class models showing the highest (0.91). The Lo–Mendell–Rubin adjusted likelihood ratio test revealed that the three-class model was an improvement from the two-class model, and the four-class model was an improvement from the three-class model.

The four classes were labeled in Figure 2, with reference to Fuster et al. (2016) and Pontes et al. (2014). Pairwise Wald test of mean equality for latent class predictors in mixture modeling was also conducted (Table 5). The first class represented regular gamers (47.3%), who generally scored below the mean average. The second class represented a low-risk engaged group (32.8%), whereas the third class represented high-risk engaged gamers (17.8%). The latter group scored significantly higher on salience-tolerance, withdrawal, conflict, and relapse than the former group. The fourth class represented probable disordered gamers (2.1%), who scored much higher on all factors (except for Conflict) and YIAT than other groups. Probable disordered gamers and high-risk engaged gamers did not significantly differ on weekly gameplay hours, but reported more hours than regular and low-risk gamers. Group differences on weekly expense in gaming, although seemed to be large, were not statistically significant (p > .05) due to high standard error values.

An 18-item model

Given their low factor loading, a model after removal of Items 2 and 19 was also tested. This 18-item model showed a good model fit, χ²(153) = 668.643, p < .001;
**Table 5. Comparison of the four latent classes**

|                      | Regular gamers (n = 517) | Low-risk gamers (n = 358) | High-risk gamers (n = 194) | Probable disordered gamers (n = 23) |
|----------------------|--------------------------|---------------------------|----------------------------|-----------------------------------|
| Salience-tolerance   | Mean: 1.32<sup>a</sup> 0.02 | Mean: 2.26<sup>b</sup> 0.06 | Mean: 2.93<sup>c</sup> 0.04 | Mean: 3.97<sup>d</sup> 0.24 |
| Mood modification     | Mean: 2.50<sup>a</sup> 0.05 | Mean: 3.22<sup>a</sup> 0.04 | Mean: 3.33<sup>b</sup> 0.05 | Mean: 3.93<sup>c</sup> 0.15 |
| Withdrawal            | Mean: 1.11<sup>a</sup> 0.01 | Mean: 1.87<sup>a</sup> 0.04 | Mean: 2.92<sup>c</sup> 0.09 | Mean: 4.27<sup>d</sup> 0.15 |
| Conflict              | Mean: 1.66<sup>a</sup> 0.02 | Mean: 2.17<sup>b</sup> 0.04 | Mean: 2.79<sup>c</sup> 0.05 | Mean: 3.06<sup>d</sup> 0.24 |
| Relapse               | Mean: 1.20<sup>a</sup> 0.02 | Mean: 2.29<sup>b</sup> 0.07 | Mean: 2.99<sup>c</sup> 0.07 | Mean: 4.30<sup>d</sup> 0.20 |
| IGD-20 Test           | Mean: 30.67<sup>a</sup> 0.25 | Mean: 46.52<sup>a</sup> 0.32 | Mean: 59.21<sup>b</sup> 0.38 | Mean: 76.66<sup>c</sup> 1.79 |
| YIAT                 | Mean: 29.77<sup>a</sup> 0.41 | Mean: 47.59<sup>a</sup> 0.59 | Mean: 59.34<sup>b</sup> 0.70 | Mean: 77.34<sup>c</sup> 3.20 |
| Weekly game time      | Mean: 3.71<sup>a</sup> 0.30 | Mean: 8.64<sup>a</sup> 0.66 | Mean: 13.53<sup>a</sup> 1.25 | Mean: 20.71<sup>c</sup> 4.24 |
| Weekly expense        | Mean: 16.36<sup>a</sup> 8.02 | Mean: 17.71<sup>a</sup> 8.02 | Mean: 105.82<sup>b</sup> 52.81 | Mean: 512.93<sup>c</sup> 435.49 |
| Age                  | Mean: 16.12<sup>a</sup> 0.16 | Mean: 16.84<sup>a</sup> 0.20 | Mean: 17.39<sup>b</sup> 0.27 | Mean: 16.36<sup>c</sup> 0.84 |
| Gender               | Mean: 0.36<sup>a</sup> 0.02 | Mean: 0.60<sup>a</sup> 0.03 | Mean: 0.80<sup>a</sup> 0.03 | Mean: 0.68<sup>c</sup> 0.10 |

Note. Mean values having different superscript letters (a–d) are different on at least p < .05 level according to the pairwise Wald test of mean equality for latent class predictors in mixture modeling. SD: standard deviation; SE: standard error; IGD-20 Test: 20-item Internet Gaming Disorder Test; YIAT: Young Internet Addiction Test.

<sup>a</sup> in Chinese Yuan (CNY); 1 CNY ≈ 0.16 USD.

CFI = 0.943; TLI = 0.930; RMSEA = 0.063, 90% CI [0.059, 0.068]; SRMR = 0.033, which outperformed the one-factor model. It had satisfactory factor loading for all items (0.392–0.864), high factor correlations (0.51–0.96), and high reliabilities (whole scale: α = 92; subscales: α = 68–84). Configural [χ²(250) = 856.434, p < .001; RMSEA = 0.067; CFI = 0.937; SRMR = 0.038], metric [χ²(263) = 903.365; RMSEA = 0.067; CFI = 0.933; SRMR = 0.050], and scalar [χ²(276) = 1,030.169; RMSEA = 0.071; CFI = 0.921; SRMR = 0.054] invariance models held invariance for latent class predictors in mixture modeling: equality for latent class predictors in mixture modeling. Disorder Test; YIAT: Young Internet Addiction Test.

**DISCUSSION AND CONCLUSIONS**

We tested the psychometric properties of a Chinese version of IGD-20 test in two Chinese youth samples. The original six-factor model of the test was not replicated in this study; however, the five-factor model, which modified the six-factor model by combining “salience” and “tolerance” factors into one, showed good model fit. This five-factor model showed model invariance across middle school and university samples. It also yielded satisfactory results in terms of reliability and concurrent validity in the overall and separate student samples.

An extremely high correlation between “salience” and “tolerance” factors has been consistently observed (r ≥ .94; Fuster et al., 2016; Pontes et al., 2014). It is difficult to differentiate these two factors, plausibly because of a high comorbidity of these two symptoms among young people with IGD. Individuals may need to spend increasing amount of time on online gaming (tolerance development) as a result of their increased need and effort to reach their goals of higher achievement in the game (salience; King, Herd, & Delfabbro, 2017, 2018). Meanwhile, the increased amount of time spent on gaming can lead to gaming to be the dominant activity in the gamer’s daily life. In some existing IGD assessment tools, indicative items of salience and tolerance are also grouped under the same factor. For instance, items related to tolerance (e.g., “I tend to spend increasing amounts of time playing online game”) and salience (e.g., “I image to play online game when I am not playing”) were used as the indicators of “failure of self-control” factor in the Problematic Online Game Use Scale (Kim & Kim, 2010). Our findings showed that the factor resulted from combining the “salience” and “tolerance” factors in our Chinese version keeps its practical function because this factor was not only significantly correlated with YIAT score and depression symptoms, but also helpful to differentiate various types of gamers.

This study established concurrent validity of this Chinese version of the IGD-20 Test by demonstrating the measure’s strong and positive correlation with the modified YIAT score, moderate correlations with depression symptoms and weekly gameplay time, and mild correlation with gaming expense, which are consistent with previous studies (Fuster et al., 2016; Kuss, Griffiths, & Pontes, 2017; Pontes et al., 2014; Wu et al., 2018). Although IGD tendencies and time spent on gaming are generally expected, one should note that “highly engaged” gamers are most often not addicted (Griffiths, 2009; Kiraly, Toth, Urban, Demetrovics, & Maraz, 2017). This was also supported by our results of the pairwise Wald test of mean equality for latent class predictors in mixture modeling: compared to high-risk engaged gamers, probable disordered gamers reported significantly higher scores on both the IGD-20 Test and YIAT, but not more time on gaming. Compared to high gaming engagement, gaming addiction imposes greater mental health threat (Loton, Borkoles, Lubman, & Polman, 2015).
As expected, the IGD-20 Test score was consistently found to be associated with depression symptoms reported by the Chinese students. Our finding is consistent with previous ones of moderate positive correlations between IGD (or video game addiction or IA) and depression (Sigerson et al., 2017; Wu et al., 2018). Such association is likely due to IGD symptoms leading to various negative consequences with respect to physical health, relationship quality, and productivity in work/study, which are associated with psychological distress (Kendler, Karkowski, & Prescott, 1999). Previous studies have also found that people with depression are more vulnerable to IGD because they may excessively and inflexibly use online games to cope with or modify their distress (Cheng, Sun, & Mak, 2015; Loton et al., 2015). A vicious cycle may be resulted.

The results of the latent profile analysis showed that the four-class solution, with four parallel lines, is the best among other solutions. Patterns of lines crossing each other in previous studies (e.g., Pontes et al. 2014) were not observed, which indicate that the IGD-20 Test differentiated Chinese gamers well. As expected, probable disordered gamers (i.e., 2.1%) scored significantly higher in all factors (except Conflict) than other groups of gamers. The conflict factor score of the probable disordered group was much lower than the scores on the other factors in the same group, and was not significantly different from that of the high-risk group. A plausible reason for this unexpected observation is that, since students take fewer social responsibilities and roles than (working) adults, our Chinese student gamers would suffer from less social sanctions against their problem gaming and experience fewer role and interpersonal conflicts at the moment but the conflicts would be more likely to be aware in long run.

This study has some limitations. First, the sole use of self-report questionnaires has some inherent disadvantages. Future research may attempt to cross-validate the test in clinical samples with structured interviews, and a cut-off score for screening purpose can then be estimated. Second, although the 20-item model showed a satisfactory model fit, it should be noted that Items 2 and 19 had low factor loadings. It could have been due to the fact that they were (and were the only two) reverse-worded items, which might be less stable than straightforward items in the Chinese language (Wei, Zhang, Li, Xue, & Zhang, 2015). Future research may consider removing or modifying these two items (e.g., changing the items straightforward-worded) and examine if some psychometric properties (e.g., Conflict subscale’s reliability) would further improve. This study recruited only student gamers by convenience and hence future research may examine if the current findings can be replicated in other gamer groups (e.g., working adults).

On the whole, the findings of this study have shown that our Chinese version of the IGD-20 Test is a psychometrically sound measurement tool for assessing IGD in accordance with the diagnostic criteria in DSM-5 among Chinese youth. However, special attention should be paid to Items 2 and 19 when use. Both this research and the previous one did not examine the measurement invariance of the IGD-20 Test across gender. Future research may also further investigate the test performance in two genders across cultures.

**Funding sources:** The research was supported by the research grant of the University of Macau (ref. no.: MYRG2016-00162-FSS). The funding source had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication.

**Authors’ contribution:** Dr. SY, Ms. MZ, and Prof. AMSW were responsible for the research conception, questionnaire design, data collection, finding interpretation, and preparation of the manuscript. Mr. IJAP conducted the statistical analysis and was involved in result interpretation and report. All authors had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. They also contributed to and have approved the final version of the manuscript.

**Conflict of interest:** All authors declare no conflict of interest.

**REFERENCES**

American Psychiatric Association [APA]. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 185–216. doi:10.1177/004912419700100301

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions, 154*, 136–136. doi:10.1177/0049124192021002005

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*(3), 456–466. doi:10.1037/0033-2909.105.3.456

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. doi:10.1080/1070510701301834

Cheng, C., Sun, P. Z., & Mak, K. K. (2015). Internet addiction and psychosocial maladjustment: Avoidant coping and coping inflexibility as psychological mechanisms. *Cyberspsychology, Behavior, and Social Networking, 18*(9), 539–546. doi:10.1089/ cyber.2015.0121

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233–255. doi:10.1207/S15328007SEM0902_5

Fuster, H., Carbonell, X., Pontes, H. M., & Griffiths, M. D. (2016). Spanish validation of the Internet Gaming Disorder-20 (IGD-20) Test. *Computers in Human Behavior, 56*, 215–224. doi:10.1016/j.chb.2015.11.050

Griffiths, M. D. (2005). A ‘components’ model of addiction within a biopsychosocial framework. *Journal of Substance Use, 10*(4), 191–197. doi:10.1080/14659890500114359

Griffiths, M. D. (2009). The role of context in online gaming excess and addiction: Some case study evidence. *International
Journal of Mental Health and Addiction, 8(1), 119–125. doi:10.1007/s11469-009-9229-x
Griffiths, M. D., Kuss, D. J., Lopez-Fernandez, O., & Pontes, H. M. (2017). Problematic gaming exists and is an example of disordered gaming. Journal of Behavioral Addictions, 6(3), 296–301. doi:10.1556/2006.6.2017.037
Hawi, N. S., & Samaha, M. (2017). Validation of the Arabic version of the Internet Gaming Disorder-20 Test. Cyberpsychology, Behavior, and Social Networking, 20(4), 268–272. doi:10.1089/cyber.2016.0493
Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1–55. doi:10.1080/10705519909540118
Kendler, K. S., Karkowski, L. M., & Prescott, C. A. (1999). Causal relationship between stressful life events and the onset of major depression. American Journal of Psychiatry, 156(6), 837–841. doi:10.1176/appi.156.6.837
Kim, M. G., & Kim, J. (2010). Cross-validation of reliability, convergent and discriminant validity for the Problematic Online Game Use Scale. Computers in Human Behavior, 26(3), 389–398. doi:10.1016/j.chb.2009.11.010
Kuss, D. J., Grif, O., & Pontes, H. M. (2014). The conceptualisation and measurement of DSM-5 Internet gaming disorder: The development of the IGD-20 Test. PLoS One, 9(10), e110137. doi:10.1371/journal.pone.0110137
Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. Journal of Statistical Software, 48(2), 1–36. doi:10.18637/jss.v048.i02
Rumpf, H. J., Achab, S., Billieux, J., Bowden-Jones, H., Carragher, N., Demetrovics, Z., Higuchi, S., King, D. L., Mann, K., Potenza, M., Saunders, J. B., Abbott, M., Ambekar, A., Aricak, O. T., Assanangkornchai, S., Bahar, N., Borges, G., Brand, M., Chan, E. M., Chung, T., Derevensky, J., Kashef, A. E., Farrell, M., Fineberg, N. A., Gandel, C., Gentile, D. A., Griffiths, M. D., Goudriaan, A. E., Grill-Bronnec, M., Hao, W., Hodgins, D. C., Ip, P., Király, O., Lee, H. K., Kuss, D., Lemmens, J. S., Long, J., Lopez-Fernandez, O., Mihara, S., Petry, N. M., Pontes, H. M., Rahimi-Movaghar, A., Rehbein, F., Rehm, J., Scafato, E., Sharma, M., Spritzer, D., Stein, D. J., Tam, P., Weinstein, A., Wittchen, H. U., Wölling, K., Zullino, D., & Poznyak, V. (2018). Including gaming disorder in the ICD-11: The need to do so from a clinical and public health perspective. Journal of Behavioral Addictions, 7(3), 556–561. doi:10.1556/2006.7.2018.59
Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. Psychometrika, 66(4), 507–514. doi:10.1007/BF02296192
Sigerson, L., Li, A. Y., Cheung, M. W., Luk, J. W., & Cheng, C. (2017). Psychometric properties of the Chinese Internet Gaming Disorder Scale. Addictive Behaviors, 74, 20–26. doi:10.1016/j.addbeh.2017.05.031
Sperber, A. D. (2004). Translation and validation of study instruments for cross-cultural research. Gastroenterology, 126, S124–S128. doi:10.1053/j.gastro.2003.10.016
van Buuren, S., & Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in R. Journal of Statistical Software, 45(3), 1–67. doi:10.18637/jss.v045.i03
van den Brink, W. (2017). ICD-11 gaming disorder: Needed and just in time or dangerous and much too early? Journal of Behavioral Addictions, 6(3), 290–292. doi:10.1556/2006.6.2017.040
van Rooij, A. J., Ferguson, C. J., Colder Carras, M., Kardfelt-Winther, D., Shi, J., Aarseth, E., Bean, A. M., Bergmark, K. H., Brus, A., Coulsdon, M., Deleuze, J., Dullur, P., Dunkels, Edman, J., Elson, M., Ettchells, P. J., Fiskalaa, A., Granic, I., Ianzj, J., Karlsen, F., Kaye, L. K., Kirsh, B., Lieberoth, A., Markey, P., Mills, K. L., Nielsen, R. K. L., Orben, A., Poulsen, A., Pruse, N., Prax, P., Quandt, T., Schimmert, A., Starcevic, V., Stutman, G., Turner, N. E., van Looy, J., & Przybylski, A. K. (2018). A weak scientific basis for gaming disorder: Let us err on the side of caution. Journal of Behavioral Addictions, 7(1), 1–9. doi:10.1556/2006.7.2018.19
van Rooij, A. J., Van Looy, J., & Billieux, J. (2017). Internet gaming disorder as a formative construct: Implications for conceptualization and measurement. Psychiatry and Clinical Neurosciences, 71(7), 445–458. doi:10.1111/pcn.12404
Vandenbergh, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research.
Wei, J., Zhang, C., Li, Y., Xue, S., & Zhang, J. (2015). Psychometric properties of the Chinese version of the fear of negative evaluation scale-brief (BFNE) and the BFNE-straightforward for middle school students. *PLoS One, 10*(3), e0115948. doi:10.1371/journal.pone.0115948

World Health Organization. (2019). *Gaming disorder*. Retrieved from https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/1448597234

Wu, A. M. S., Chen, J. H., Tong, K. K., Yu, S., & Lau, J. T. F. (2018). Prevalence and associated factors of Internet gaming disorder among community dwelling adults in Macao, China. *Journal of Behavioral Addictions, 7*(1), 62–69. doi:10.1556/2006.7.2018.12

Young, K. S. (1996). *Caught in the net*. New York, NY: John Wiley & Sons.

Yu, S., Wu, A. M. S., & Pesigan, J. J. A. (2016). Cognitive and psychosocial health risk factors of social networking addiction. *International Journal of Mental Health and Addiction, 14*(4), 550–564. doi:10.1007/s11469-015-9612-8

Yu, Z., Sun, P., Zhang, S., Zhu, J., Guo, X., & Zhao, M. (2010). An investigation of the online game addiction of middle school students in Guangzhou. *Journal of Clinical Psychosomatic Diseases, 16*(2), 137–141.

Zhou, Y., & Li, Z. (2009). Online game addiction among Chinese college students measurement and attribution. *Annual Review of Cybertherapy and Telemedicine, 7*(1), 149–154. doi:10.3233/978-1-60750-017-9-149
请根据过去十二个月的情况，对下列陈述选出最符合你的答案。

1. 我经常因为长时间打游戏而睡眠不足。
2. 我从来不会为了让自己感觉好一点而去打游戏。
3. 在过去的一年里我打游戏的时间明显增多了。
4. 我打游戏的时候会感觉比较急躁易怒。
5. 因为打游戏，我对我其他的爱好失去了兴趣。
6. 我想少打点游戏但是很难做到。
7. 不打游戏的时候我总是想下下次玩的时候的事。
8. 打游戏可以帮助我处理我的坏情绪。
9. 我需要越来越多的时间来打游戏。
10. 不能打游戏的时候我会觉得难过。
11. 我曾因为打游戏而对家人说谎。
12. 我觉得我没法不打游戏。
13. 我觉得打游戏已经成为我生活中最消耗时间的活动了。
14. 我打游戏来忘掉烦心事。
15. 我经常觉得一整天时间不够我做完我在游戏里想做的事。
16. 不论我因为什么原因不能打游戏的时候，我会觉得比较焦虑。
17. 我觉得我打游戏这件事已经危害了我跟我另一半的关系。
18. 我经常尝试少打点游戏但发现自己做不到。
19. R* 我认为打游戏对我生活里重要的东西有负面影响。
20. 我经常因为打游戏而忘记了其它事务（比如上班、上学、家务等等）产生负面影响。

Dimensions:
Salience-tolerance: 1, 3, 7, 9, 13, 15
Mood modification: 2, 8, 14
Withdrawal: 4, 10, 16
Conflict: 5, 11, 17, 19, 20
Relapse: 6, 12, 18

Note. A 5-point Likert response scale (1 = 非常不同意, 2 = 不同意, 3 = 中立, 4 = 同意, and 5 = 非常同意) is used.
*R reversed items.