<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Psychometric properties of the Chinese Internet Gaming Disorder Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Sigerson, L; Li, AYL; Cheung, MWL; Luk, JW; Cheng, C</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>Addictive Behaviors, 2017, v. 74, p. 20-26</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2017</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/244734">http://hdl.handle.net/10722/244734</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
Psychometric properties of the Chinese Internet Gaming Disorder Scale

Leif Sigersona, Angel Y.-L. Lia, Mike W.-L. Cheungb, Jeremy W. Lukc, Cecilia Chenga,⁎

a Department of Psychology, The University of Hong Kong, Pokfulam Road, Hong Kong
b Department of Psychology, National University of Singapore, 9 Arts Link, Singapore
c Health Behavior Branch, Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, 6710B Rockledge Drive, Bethesda, MD, USA

ARTICLE INFO

Keywords:
Internet gaming disorder
Problematic gaming
Behavioral addiction
Scale validation
DSM
Measurement invariance

ABSTRACT

To develop a consensus on the definition and measurement of Internet gaming disorder (IGD), several recent studies have used the DSM-5's proposed criteria for IGD as the basis in scale construction. This study contributes to this emerging consensus by developing and validating a new Chinese Internet Gaming Disorder Scale (C-IGDS) based on the DSM-5 criteria. A representative sample of Hong Kong community adults (n = 502, 50% men, mean age = 37.1, age range = 18–60) was recruited for a telephone survey with random digit dialing. Various statistical techniques were used to assess the psychometric properties of the C-IGDS. The C-IGDS had good reliability (Cronbach’s α = 0.91) and structural validity (CFA model fit: RMSEA = 0.027, CFI = 0.991, TLI = 0.988) in our sample. Moderate to moderately strong correlations with depressive symptoms (r = 0.617, p < 0.001), social anxiety symptoms (r = 0.366, p < 0.001), and gaming hours (r = 0.412, p < 0.001) supported the criterion validity of the C-IGDS. In addition, the C-IGDS exhibited strict measurement invariance for sex and at least strong measurement invariance for age. In addition to providing the first Chinese scale for measuring IGD based on the DSM-5's proposed criteria, this study provides empirical support for the validity of these diagnostic criteria as the basis for a universal measure of IGD. Most important, this study is the first to reveal the criteria's measurement invariance, thereby indicating their suitability for use with diverse demographic groups.

1. Introduction

1.1. Problematic gaming as a global public health issue

In the present Cyberage, excessive game playing can lead to what is known as Internet gaming disorder (IGD) that is detrimental to mental health (e.g., Sarda, Bègue, Bry, & Gentile, 2016). The prevalence rates of IGD are estimated to range from 5% to 8% in North America (e.g., Desai, Krishnan-Sarin, Cavallo, & Potenza, 2010; Gentile, 2009), from 0.2% to 12% in Europe (e.g., Festl, Scharkow, & Quandt, 2013; Wittek et al., 2016), and from 8% to 46% in Asia (e.g., Gentile et al., 2011; Wan & Chiou, 2006). Such wide ranges in prevalence rate may be attributable to the varied conceptualizations and assessments for IGD found in the rapidly growing literature. In the Fifth Edition of the Diagnostic and Statistical Manual of Mental Disorder, IGD is included as a “condition for further study” (American Psychiatric Association, 2013, p. 795). Hence, the proposed diagnostic criteria and cutoff are not set in stone and need to be validated.

In 2015, there were almost 400 million Chinese online gamers (China Internet Network Information Center, 2016). Studies have shown positive associations between IGD and a wide range of issues, including substance-related addictions, behavioral addictions, and emotional disorders (e.g., Sigerson, Li, Cheung, & Cheng, 2017; Zhang et al., 2016). However, there is a lack of standardized tools for assessing IGD in Chinese populations. The availability of a reliable and valid Chinese assessment tool is thus of paramount importance for further research on this problem. The present study thus aims to (a) develop and validate a Chinese IGD scale (C-IGDS) based on the DSM-5's proposed diagnostic criteria; and (b) conduct measurement invariance tests to determine the applicability of the C-IGDS for a demographically-diverse Chinese community sample.

1.2. DSM-5 proposed diagnostic criteria for IGD

Since their proposal, the DSM-5 diagnostic criteria for IGD have been transformed and adapted to distinguish between average and problematic gamers in several countries (e.g., Lemmens, Valkenburg, & Gentile, 2015; Pontes & Griffiths, 2015; Rehbein, Kliem, ...
1.3. Evaluation of scale stability across demographic groups

This study also aimed to evaluate whether the C-IGDS assesses IGD consistently across different respondents within the population by testing its measurement invariance across groups with distinct demographic characteristics (i.e., men and women, younger and older adults). Measurement invariance indicates that respondents’ scores on a scale depend on their levels of the variable being assessed by the scale rather than their membership of a particular demographic group (Meredith & Milsap, 1992).

IGD is known to be more prevalent among the young and among males (e.g., Fuster, Carbonell, Pontes, & Griffiths, 2016; Zanetta Dauriat et al., 2011), and these imbalances could introduce systematic bias into the measurement of IGD. As epidemiological research typically involves large samples with assorted demographic profiles, the establishment of measurement invariance is required to facilitate meaningful comparisons across diverse demographic groups. Given that the C-IGDS is designed for use in heterogeneous Chinese samples, it is thus essential to test for its measurement invariance.

2. Methods

2.1. Data collection and participants

Prior research ethics approval has been obtained from the Human Research Ethics Committee of the University of Hong Kong. Data were collected via a population-based telephone survey using random digit dialing, and an adult member from each household was selected according to the most-recent-birthday criterion. Upon completion of the interviews, participants were entered into a lucky draw for ten 500 Hong Kong dollars (approximately 65 US dollars) supermarket vouchers.

Within two months, 1045 participants had successfully completed the survey. Any participants who reported 0 hours per week of gaming were excluded from the analyses, leaving a final sample of 502 gamers. The final sample’s demographics are presented in Table 1. This sample was evenly balanced by gender (49.8% female) and featured a wide age range (mean = 37.1, SD = 13.3, range = 18–60). According to the most recent census data, this sample is roughly representative of the Hong Kong population (Hong Kong Census and Statistics Department, 2016).

2.2. Development of the C-IGDS

The DSM-5 (see Section 3) proposes nine symptoms indicative of IGD. In adapting these symptoms into a scale for our survey, we first converted each symptom into a self-report item, altering the DSM’s original wording as little as possible. For instance, one of the proposed DSM symptoms is “Use of Internet games to escape or relieve a negative mood (e.g., feelings of helplessness, guilt, anxiety)" (American Psychiatric Association, 2013, p. 795). For our survey, we converted it into “Do you play Internet games in order to escape or relieve a negative mood (e.g., feelings of helplessness, guilt, anxiety)?” To emulate the DSM’s original format, we included “yes” or “no” response options for each item. A complete list of the converted items is available in the Appendix.

To develop a Chinese version of the IGD scale, we adopted the back-translation method recommended by Brislin (1986). A bilingual researcher first translated the DSM-5 IGD diagnostic criteria from English into Chinese, and another bilingual researcher then back-translated them into English. In the final stage, one of the authors (Li) reviewed and resolved any discrepancies with both translators before finalizing the scale.

2.3. Additional measures

We administered the Social Interaction Anxiety Scale-Short form (SIAS; Peters, Sunderland, Andrews, Rapee, & Mattick, 2012). The Chinese version has been validated by Yang (1997). The Chinese SIAS displayed good reliability in this study (Cronbach’s α = 0.78).

The Center for Epidemiologic Studies Depression Scale-Short Form (CES-D; Cole, Rabin, Smith, & Kaufman, 2004) was used. The Chinese version has been validated by Cheung, Liu, and Yip (2007). The Chinese CES-D was found to reliably measure depressive symptoms in the present sample (Cronbach’s α = 0.79).

Participants were asked to report their sex and age. In addition, we assessed participants’ amount of weekly gaming with two items asking how many hours per day, on average, they played during the week and on weekends, respectively. The responses were then multiplied and summed together to assess how many hours per week each participant spent on gaming.

3. Results

3.1. Psychometric properties of Chinese IGD scale

To examine whether the C-IGDS is adequate for assessing IGD in Hong Kong, we assessed its three major psychometric properties: reliability, structural validity, and criterion validity. These analyses were conducted in lavaan version 5.20 (Rosseel, 2012), and R version 3.2.2 (R Core Team, 2015).
procedures outlined above, we
Numerous studies have demonstrated IGD to be consistently associated
symptoms, social anxiety symptoms, and gaming hours per week.
C-IGDS and three common criterion measures of IGD: depressive
3.1.2. Criterion validity
and structural validity.
The reliability of the scale assessed via internal consistency was
good, with a Cronbach’s alpha of 0.91
To assess the structural validity of the C-IGDS, we fitted a simple
eight items of the scale. As the responses
to the scale are dichotomous, we used the diagonally weighted
leastsquares estimator. To evaluate model fit, we relied on
the conventional criteria of RMSEA < 0.06, CFI > 0.095, and
TLI > 0.095 (Hu & Bentler, 1999).
The resulting CFA measurement model for IGD had a close fit to
the data: RMSEA = 0.027 (90% CI = 0.000, 0.047), CFI = 0.991,
TLI = 0.988, \( \chi^2(df = 27) = 36.429 \). In addition, no standardized fac-
tor loading was below 0.59 (see Table 2 for all factor loadings). On the
basis of these results, we concluded that the C-IGDS has good reliability and
structural validity.

3.1.2. Criterion validity
To test criterion validity, we examined the correlations between the
C-IGDS and three common criterion measures of IGD: depressive
symptoms, social anxiety symptoms, and gaming hours per week.
Numerous studies have demonstrated IGD to be consistently associated
with these variables (see the review by Russ & Griffiths, 2012). Thus, if
the scale was indeed assessing IGD, it should have been correlated with
all three criteria. To ensure that the C-IGDS was not simply assessing a
general negative mood, or, alternatively, simply high engagement with
games (Charlton & Danforth, 2007), we also needed to ensure that it did
not exhibit any excessive overlap with any one of the three variables. To
assess whether this would be the case, we relied on
Brown’s (2006) recommendation that a factor correlation above 0.80 indicates a lack of
distinction between separate variables (p. 131).
To remove measurement error and obtain a more accurate estimate of
the true correlations between these variables (Bollen & Lennox, 1991), we constructed a CFA model comprising three latent variables
and their corresponding items: IGD, depressive symptoms (measured with the
CES-D), and social anxiety symptoms (measured with the
SIAS). Gaming hours per week was included as a directly measured
variable. This model is depicted in Fig. 1. Following the model fitting
procedures outlined above, we fitted a CFA model to the data, and
obtained a good fit: RMSEA = 0.033 (90% CI = 0.027, 0.040), CFI = 0.953,
TLI = 0.948, \( \chi^2(df = 294) = 443.294 \). As expected, the
C-IGDS had moderate to strong positive correlations with depressive
symptoms, social anxiety symptoms, and gaming hours. However, no
factor correlation approached 0.80, suggesting that the scale differen-
tiates IGD well from all of these three criteria. These correlations,
along with descriptive statistics for each variable, are available in

---

Table 2
Factor loadings of Chinese Internet Gaming Disorder Scale items (derived from confirmatory factor analysis model).

<table>
<thead>
<tr>
<th>Item</th>
<th>Standardized factor loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Do you feel preoccupied with Internet games (think about previous gaming activity or anticipate playing the next game)?</td>
<td>0.804</td>
</tr>
<tr>
<td>2. Do you feel irritable, anxious, or sad when Internet gaming is taken away?</td>
<td>0.751</td>
</tr>
<tr>
<td>3. Do you feel the need to engage in Internet games with increasing amounts of time in order to achieve satisfaction?</td>
<td>0.788</td>
</tr>
<tr>
<td>4. Have you repeatedly made unsuccessful attempts to control your participation in Internet games?</td>
<td>0.684</td>
</tr>
<tr>
<td>5. Have you experienced loss of interests in previous hobbies and entertainment as a result of, and with the exceptions of, Internet games?</td>
<td>0.753</td>
</tr>
<tr>
<td>6. Do you continue to use Internet games excessively despite knowledge of psychosocial problems?</td>
<td>0.928</td>
</tr>
<tr>
<td>7. Have you deceived family members, therapists, or others regarding the amount of Internet gaming?</td>
<td>0.594</td>
</tr>
<tr>
<td>8. Do you use Internet games as a way of escaping or relieving a negative mood (e.g., feeling of helplessness, guilt, anxiety)?</td>
<td>0.725</td>
</tr>
<tr>
<td>9. Have you jeopardized or lost a significant relationship, job, or educational or career opportunity because of participation in Internet games?</td>
<td>0.651</td>
</tr>
</tbody>
</table>

---

In summary, this set of results indicates that the C-IGDS is internally consistent and has good structural and criterion validity in a Chinese sample, suggesting that this new scale is both reliable and valid for use among Chinese adults.

3.2. Measurement invariance testing
The measurement invariance of the C-IGDS was tested for sex and age. Measurement invariance is most commonly tested by fitting and comparing a series of CFA models, with each model constraining an additional set of parameters to be equal across groups. If the constraint does not result in a large decrease in model fit, then the parameters are deemed equivalent across groups. The last model constrains all parameters to be equal across groups.

As the responses to the C-IGDS are dichotomous, its CFA measurement model involves different parameters from those constrained in traditional measurement invariance testing with continuous data. As an alternative, we used the technique proposed by Millsap and Yun-Tein (2004) for testing measurement invariance with dichotomous data (p. 502–510). In the first model (to test configural invariance), the following parameter constraints are needed to achieve model identification (a technical requirement for CFA models): (a) the factor loading of the first item is set to 1; (b) the item thresholds (similar to item intercepts in models with continuous data) are constrained to be equal across groups; (c) the factor mean of the first group is set to 0; (d) the variances of the item residuals are set to 1 in the first group; and (e) the variances of the first item’s residuals are fixed to 1 in the other groups. In the next model (to test strong invariance), the factor loadings are constrained to be equal across groups. In the last model (to test strict invariance), the variances of the item residuals are constrained to be equal across groups. This technique does not include a test for weak invariance, one of the intermediate steps in the traditional technique. However, the final test, which assesses strict invariance, is equivalent to the traditional test of strict measurement invariance.

There is currently no consensus on which fit indices and criteria to use in measurement invariance testing, and so we relied on multiple sources and recommendations. A commonly used criterion, proposed by Cheung and Rensvold (2002), is that if the CFI decreases by > 0.01, then measurement invariance does not hold. Alternatively, reliance on the RMSEA was recommended by both Kim and Yoon (2011) and Chen (2007), with the latter proposing that an increase of > 0.015 in this criterion demonstrates a lack of measurement invariance and also supporting Cheung and Rensvold’s (2002) proposal. Taking all of these recommendations into consideration, we relied on both the RMSEA and CFI in assessing measurement invariance for the C-IGDS. We performed the measurement invariance tests using Mplus version 6.1 (Muthén & Muthén, 2007).

To ensure the robustness of the results, two different types of age grouping were tested, following the procedures of several recent studies (Estabrook, Sadler, & McGuie, 2015; Spaapen, Waters, Brummer,
We first split the sample by the median to create two age groups, and then split it by the 33rd and 67th percentiles to create three groups. Measurement invariance was then tested separately for the two types of age grouping, as well as for the sexes. Demographic data for all sub-samples used in these tests are presented in Table 1, and full results of the measurement invariance tests can be found in Table 4.

The results showed that though the RMSEA indicated that the C-IGDS has strict measurement invariance for age and gender (i.e., no increase greater than the cutoff of 0.015), there were three instances where the CFI indicated a lack of invariance (i.e., a decrease of > 0.010). In resolving this contradiction, we referenced several empirical and simulation studies which showed that the RMSEA is more accurate than the CFI when used with CFA models containing categorical data (Hutchinson & Olmos, 1998; Nye & Drasgow, 2011; Sugawara & MacCallum, 1993). Thus, the RMSEA is likely to be a better indicator of measurement invariance with our categorical data if both indices yielded inconsistent findings.

The results showed that the C-IGDS did have strict measurement invariance for sex, because both the RMSEA and CFI supported the strict invariance model. For the strong invariance model, however, the decrease in CFI was slightly above the cutoff (i.e., 0.011); but we considered that this was overruled by the more trustworthy RMSEA, whose decrease was well below the cutoff (i.e., 0.006). Given this, and the consistent findings accepting the strict invariance model, it is reasonable to conclude that the C-IGDS has strict measurement invariance for sex.

For age, the results revealed that the C-IGDS had strong invariance, because both the RMSEA and CFI accepted the strong invariance models for both types of age groupings (median split and tripartite split). However, the strict invariance model for both age groupings was rejected by the CFI (decreases of 0.016 for each), but accepted by the RMSEA (increases of 0.012 and 0.005).

![Fig. 1. CFA model testing criterion validity of the C-IGDS](image)

*Note: GH = gaming hours, IGD = Internet Gaming Disorder, SA = social anxiety symptoms, Dep = depressive symptoms. Loadings and parameters are standardized.*
Taken together, the present results indicate that the C-IGDS has strict measurement invariance for sex, and at least strong measurement invariance for age. There is currently some debate about whether strict measurement invariance is required for a scale (e.g., Wu, Li, & Zumbo, 2007). However, at the minimum the present results show that the C-IGDS can be used in the same form with different demographic groups, that the C-IGDS can be used to assess relationships between IGD and other variables in heterogeneous samples, and that it can be used to compare IGD among demographic groups (Hirschfeld & von Brachel, 2014).

4. Discussion

The population-based telephone survey of Hong Kong adults carried out for this study demonstrated the C-IGDS to have good reliability, structural validity, and criterion validity. In addition, measurement invariance tests revealed the C-IGDS to exhibit strict measurement invariance for sex and at least strong measurement invariance for age.

4.1. Implications for IGD researchers

Our findings have several implications for future research on IGD. First and foremost, we provide rigorous psychometric evidence to show that the C-IGDS is a reliable, valid tool for assessing IGD. This is the first Chinese IGD scale based on the DSM-5’s proposed criteria, and we have demonstrated that it can be used for future studies on IGD in Chinese societies, where the disorder is particularly prevalent and an issue of considerable public concern. A potential concern for IGD researchers is that age or sex may compromise measurement of this disorder. Despite this concern, no other scale based on the DSM-5’s proposed criteria for IGD has been tested for measurement invariance. This study addresses this unexplored issue by performing measurement invariance tests, with the results showing that there is little or no age or sex-related bias in the C-IGDS. IGD researchers can thus administer it knowing that the scores reflect participants’ perceived level of IGD rather than their age or sex.

In addition to providing a robust scale for measuring IGD in Chinese samples, the results of this study may help to advance the broader study of IGD. A recent editorial (Griffiths et al., 2016) noted that a key limitation in the study of IGD is the lack of consensus concerning its definition and measurement, and this limitation must be overcome if IGD is to be fully recognized as a disorder in the DSM. They also advocated the adoption of the requirements set out by Koronczai et al. (2011) for a suitable measure of problematic Internet use (which would include IGD). Koronczai et al. advocated that any such measure must be (a) comprehensive, short, reliable, and valid for different methods of data collection, (b) reliable and valid for different age groups, (c) cross-culturally reliable and valid, and (d) validated on clinical samples. A number of subsequent studies recently show that the DSM-5’s IGD criteria fulfill some of these requirements (e.g., Ko et al., 2014; Lemmens et al., 2015; Rehbein et al., 2015).

Our study adds to this discussion by providing several new sources of evidence indicating that the DSM-5 criteria meet the aforementioned requirements. First, it is the only study to validate those criteria using a population-based telephone survey, supporting their reliability and validity with different methods of data collection. Second, the good psychometric properties of the C-IGDS suggest that the DSM-5 proposed criteria are well-suited among Chinese adults, supporting the cross-cultural applicability of the proposed criteria. It is also noteworthy that, as the only short-form Chinese measure of IGD, the C-IGDS will facilitate future Chinese IGD research that can be directly compared with studies carried out in other cultural regions, allowing for a broader investigation of cultural variations and similarities. Third, the results of our measurement invariance tests suggest that the DSM-5 criteria are reasonably good indicators of the underlying IGD construct across both age groups and sexes. This is an especially important finding because measurement invariance is often assumed rather than actually tested, and it has never been tested before for the DSM-5’s proposed IGD criteria.

4.2. Limitations and future research directions

This study had some limitations that suggest fruitful directions for future research on IGD. First, although our sample was representative and heterogeneous, it was conducted in a single geographic location. Future studies may help to advance the study of IGD by conducting research in other countries, especially those with diverse cultural and socioeconomic characteristics (e.g., Cheng, Cheung, & Montasem, 2016; Cheng & Li, 2014). Second, it is also noteworthy that our study was cross-sectional in nature, which limits the inferences we can make about the nature and progression of IGD. Future longitudinal studies may build on our findings by examining the predictive validity and test-retest reliability of the C-IGDS.

4.3. Conclusions

This study contributes to the literature by providing a reliable, valid, and demographically stable Chinese IGD scale, the C-IGDS, which can be used by both researchers and clinicians. If subsequently acknowledged as a formal DSM disorder, this scale may be adapted to a brief screening tool used in clinical settings. More generally, this scale provides several new sources of support for the validity of the DSM-5 criteria, and allows for future Chinese IGD research that can be directly compared to findings from other countries.

Role of funding sources

This study was funded by Hong Kong Research Grants Council’s General Research Fund (17400714), and the University of Hong Kong’s Seed Fund for Incubating Group-based Collaborative Research Projects (102009405) and Seed Fund for Basic Research (201411159152) to Cecilia Cheng. Jeremy W. Luk’s effort on this project was supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The funders had no role in the study design and administration, data analysis or interpretation, manuscript writing, or the decision to submit the paper for publication.

Contributors

Authors CC and AYLL designed and implemented data collection. LS and MWLC conducted the data analysis. LS and AYLL conducted literature searches and wrote the first draft of the manuscript. JWL and CC provided manuscript editing and consultation on the data analysis. All authors contributed to subsequent revisions and have approved the final manuscript.

Conflicts of interest

All authors declare that they have no conflicts of interest.
Appendix A

Appendix
The Internet Gaming Disorder Scale.

Instructions: Please answer the following questions about your gaming activity in the past year

Yes No

1. Do you feel preoccupied with Internet games (think about previous gaming activity or anticipate playing the next game)?

2. Do you feel irritable, anxious, or sad when Internet gaming is taken away?

3. Do you feel the need to engage in Internet games with increasing amounts of time in order to achieve satisfaction?

4. Have you repeatedly made unsuccessful attempts to control your participation in Internet games?

5. Have you experienced loss of interests in previous hobbies and entertainment as a result of, and with the exceptions of, Internet games?

6. Do you continue to use Internet games excessively despite knowledge of psychosocial problems?

7. Have you deceived family members, therapists, or others regarding the amount of Internet gaming?

8. Do you use Internet games as a way of escaping or relieving a negative mood (e.g., feeling of helplessness, guilt, anxiety)?

9. Have you jeopardized or lost a significant relationship, job, or educational or career opportunity because of participation in Internet games?

References


Griffiths, M. D., Rooij, A. J., Kardefelt-Winther, D., St推迟cevic, V., Kirily, O., Pallesen, S.,...
information technology addictions and their relationships with non-technology-related addictions. Computers in Human Behavior (resubmitted with minor revisions).


