Journal of Behavioral Addictions
DOI: 10.1556/2006.6.2017.025

Psychometric validation of the Persian nine-item Internet Gaming Disorder Scale – Short Form: Does gender and hours spent online gaming affect the interpretations of item descriptions?

TZU-YI WU[1], CHUNG-YING LIN[2], KRISTOFER ÅRESTEDT[3], MARK D. GRIFFITHS[4], ANDERS BROSTRÖM[6] and AMIR H. PAKPOUR[5,6]*

[1]Institute of Economics, Academia Sinica, Taipei, Taiwan
[2]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
[3]Faculty of Health and Life Sciences, Linnaeus University, Kalmar, Sweden
[4]International Gaming Research Unit, Psychology Department, Nottingham Trent University, Nottingham, UK
[5]Social Determinants of Health Research Center, Qazvin University of Medical Sciences, Qazvin, Iran
[6]Department of Nursing, School of Health and Welfare, Jönköping University, Jönköping, Sweden

(Received: January 11, 2017; revised manuscript received: March 31, 2017; accepted: April 1, 2017)

Background and aims: The nine-item Internet Gaming Disorder Scale – Short Form (IGDS-SF9) is brief and effective to evaluate Internet Gaming Disorder (IGD) severity. Although its scores show promising psychometric properties, less is known about whether different groups of gamers interpret the items similarly. This study aimed to verify the construct validity of the Persian IGDS-SF9 and examine the scores in relation to gender and hours spent online gaming among 2,363 Iranian adolescents. *Methods:* Confirmatory factor analysis (CFA) and Rasch analysis were used to examine the construct validity of the IGDS-SF9. The effects of gender and time spent online gaming per week were investigated by multigroup CFA and Rasch differential item functioning (DIF). *Results:* The unidimensionality of the IGDS-SF9 was supported in both CFA and Rasch. However, Item 4 (*fail to control or cease gaming activities*) displayed DIF (DIF contrast = 0.55) slightly over the recommended cutoff in Rasch but was invariant in multigroup CFA across gender. Items 4 (DIF contrast = −0.67) and 9 (*jeopardize or lose an important thing because of gaming activity*; DIF contrast = 0.61) displayed DIF in Rasch and were non-invariant in multigroup CFA across time spent online gaming. *Conclusions:* Given the Persian IGDS-SF9 was unidimensional, it is concluded that the instrument can be used to assess IGD severity. However, users of the instrument are cautioned concerning the comparisons of the sum scores of the IGDS-SF9 across gender and across adolescents spending different amounts of time online gaming.

Keywords: adolescence, confirmatory factor analysis, differential item functioning, Internet Gaming Disorder, measurement invariance, Rasch analysis

INTRODUCTION

Internet Gaming Disorder (IGD) is defined as the problematic use of online games (American Psychiatric Association [APA], 2013) with prevalence rates of IGD being reported from 0.6% to 5.4% around the world (Király, Griffiths, & Demetrovics, 2015; Rehbein, Kliem, Baier, Mossle, & Petry, 2015). A prominent characteristic of IGD is that individuals spend most of their time playing online games and ignore physical and psychological needs. Consequently, they are very likely to compromise important social relationships, educational achievements, and/or career opportunities (Pontes, Király, Demetrovics, & Griffiths, 2014). Because problematic online gaming shares many similarities with more traditional addictions, the latest (fifth) edition of the *Diagnostic and statistical manual of mental disorders* (DSM-5) placed IGD as a potential behavioral addiction (like gambling disorder) into the Appendix of the DSM-5 as a possible emerging disorder worthy of further empirical study (APA, 2013). Given the opportunities for engaging in online gaming are ever increasing, protecting the physical and psychological well-being of individuals with IGD is of increasing importance. Therefore, there is a growing need to develop an effective instrument for healthcare providers to early detect IGD problems.

To date, there has not been a Persian version of any scale assessing IGD. Of the commonly used scales assessing IGD, the nine-item Internet Gaming Disorder Scale – Short Form (IGDS-SF9) has been used to assess IGD in both adults and adolescents and has a number of advantages. First, the IGDS-SF9 was developed using the nine

* Corresponding author: Amir H. Pakpour; Social Determinants of Health Research Center (SDH), Qazvin University of Medical Sciences, Shahid Bahonar Blvd, Qazvin 3419759811, Iran; Phone: +98 28 33239259; Fax: +98 28 33239259; E-mails: Pakpour_Amir@yahoo.com, apakpour@qums.ac.ir

criteria of IGD defined by DSM-5 (i.e., the scores of the IGDS-SF9 may be useful for clinicians to make a diagnosis). Second, the IGDS-SF9 has robust psychometric properties as evidenced in its internal consistency, construct validity, criteria-related validity, concurrent validity, and minimum floor and ceiling effects (Monacis, Palo, Griffiths, & Sinatra, 2016; Pontes & Griffiths, 2015, 2016). Third, the IGDS-SF9 is quick to administer because it contains only nine items, and has the capability for clinicians to ease their workload in a busy clinical setting. Furthermore, the IGDS-SF9 may have great potential in evaluating IGD severity among teenagers.

Although the IGDS-SF9 has been validated in several studies and different languages [e.g., English, Portuguese, and Italian (Monacis et al., 2016; Pontes & Griffiths, 2015, 2016)], all previous IGDS-SF9 validations were carried out among Western populations. Because online gaming behavior may differ across different cultures, studies examining psychometric properties of the IGDS-SF9 in Eastern countries are needed. Furthermore, no previous study has examined whether different groups similarly (and more specifically, males vs. females; and number of hours spent online gaming per week) respond to the IGDS-SF9 items in a similar pattern. Several studies show that gender affects addictive behaviors (Ko, Yen, Chen, Chen, & Yen, 2005; Struber, Luck, & Roth, 2008). Similarly, adolescents spending longer time playing online games may have different perceptions to those spending shorter time playing online games. However, for comparative purposes, there is an assumption that adolescents with different characteristics (i.e., gender and time spent online gaming per week) share the same interpretations when responding to IGDS-SF9 items. Therefore, studies assessing measurement invariance or differential item functioning (DIF) are required to determine whether the comparisons are valid.

Studies have demonstrated that the adolescent brain is more vulnerable to high-risk behaviors including addiction (Crews, He, & Hodge, 2007) and much of the research in the IGD and video game addiction field has been carried out on adolescents (although in recent years, there has been increasing research using adult samples; Griffiths, Kuss, & Pontes, 2016). In addition, although there has been increasing research examining IGD, most of it comes from Europe, North America, and South East Asia (Griffiths et al., 2016). Therefore, this study investigated IGD in an adolescent population in a country where there has been relatively little research on the topic. To address the aforementioned issues, this study translated the IGDS-SF9 into a version suitable for the use in a Middle East country (i.e., Iran) in high-school students. In addition, the psychometric properties (i.e., internal consistency, composite reliability (CR), test–retest reliability, unidimensionality, construct validity, criteria-related validity, floor and ceiling effects) of the Persian IGDS-SF9 followed by its measurement invariance and DIF were examined. With comprehensive validation of IGDS-SF9 in the target population (i.e., large sample of Iranian high-school students), clinicians can easily understand the psychometric properties of the IGDS-SF9, and use it comfortably in a clinical setting.

METHODS

Participants and procedure

Participants were randomly selected via a two-stage sampling design. According to the data by the Ministry of Education, at the time of study (2015) there were 10,482 students studying at 58 high schools (typically aged 14–16 years although a very small number of very clever children can have an accelerated education and be moved up one or two grades) in Qazvin. In total, 15 schools were randomly selected from 58 high schools in Qazvin (Iran) during August 2015 to December 2015. Twelve of 15 invited schools agreed to participate (80.0% response rate for schools) and 2,389 students agreed to participate from 2,986 students that were approached (80.0% response rate). Twenty-six students did not respond to survey items and were excluded from the subsequent analyses. The age of participants ranged between 12 and 19 years (median = 15; mean $\pm$ SD = 15.6 $\pm$ 1.2). After 2 weeks, the same adolescents were asked to repeat the study measures (n = 2,010, 85.1% response rate).

Instruments

A background information sheet was used to collect data concerning age, gender, parental education, current smoking status (yes vs. no), and hours spent online gaming per week.

Weekly gameplay. Two open-ended questions were used to assess weekly time spent playing by the gamers: "How many hours do you spend video game on specific platforms (i.e., PCs, consoles, handheld gaming devices) on a school day?" and "How many hours do you spend video game on specific platforms (i.e., PCs, consoles, handheld gaming devices) on a weekend day?" To calculate total weekly hours playing video game, the adolescent's responses during the week were multiplied by 5 while the response of weekend were multiplied by 2.

Internet Gaming Disorder Scale – Short Form (IGDS-SF9). The IGDS-SF9 is a nine-item scale assessing severity of IGD based on IGD questions in DSM-5. The scale uses a 5-point Likert-type scale with the following scored responses: *Never* (1), *Rarely* (2), *Sometimes* (3), *Often* (4), and *Very often* (5). A total score is generated by summing up the nine items (range: 9–45), where a higher score represents a higher degree of IGD. Moreover, the one-factor construct of the IGDS-SF9 has been supported by both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Only one factor was extracted in the EFA with total variance explaining 45.4%; the CFA fit indices were satisfactory [comparative fit index (CFI) = 0.964, Tucker–Lewis index (TLI) = 0.952, root mean square error of approximation (RMSEA) = 0.054, and standardized root mean square residual (SRMR) = 0.034; Pontes & Griffiths, 2015].

The translation process in this study was carried out according to the recommendations of Beaton, Bombardier, Guillemin, and Ferraz (2000). The English version of the IGDS-SF9 was translated to Persian/Farsi language by two independent bilingual translators whose mother tongue was Persian (forward translation). Following this, discrepancies

between the translated versions were resolved in a session with the translators and project manager to synthesize the Persian IGDS-SF9. The interim Persian IGDS-SF9 was then back-translated into English by two independent native English speakers who were not aware of the original version. Translated versions were compared with the original version to ensure they shared semantic meaning. Afterward, all translated versions were consolidated in a session with expert committee including a psychiatrist, psychologist, nurse, and the translators. Finally, the Persian IGDS-SF9 was pretested on a sample of 49 adolescents for cognitive interviewing (41.0% female, mean age = 15.2 years).

Depression Anxiety Stress Scale (DASS). The 21-item DASS assesses three psychiatric symptoms (i.e., depression, anxiety, and stress) with each symptom assessed using seven items. The scale uses a 4-point Likert-type scale with 0 meaning *did not apply to me at all* and 3 meaning *applied to me very much, or most of the time* (Lovibond & Lovibond, 1995). The internal consistency was satisfactory for each subscale (α = .86 for depression, .88 for anxiety, and .89 for stress; Pontes & Griffiths, 2016). The three-factor structure of the DASS Persian version was previously confirmed in an Iranian sample using CFA (CFI = 0.91 and RMSEA = 0.068; Asghari, Saed, & Dibajnia, 2008). The DASS was used for examining the criteria-related validity because studies showed the relationship between IGD and mental health problems (Kim et al., 2016). The internal consistency of the DASS was satisfactory in this study (α = .84 for depression, .88 for anxiety, and .90 for stress).

Data analysis

Descriptive analyses were used to demonstrate the characteristics of the participants. Classical test theory (CTT) and Rasch analyses were used to test psychometric properties and measurement invariance of the IGDS-SF9. More specifically, the analysis tested floor and ceiling effects, internal consistency, test–retest reliability, corrected item–total correlation, average variance extracted (AVE), CR, standard error of measurement, construct validity, criteria-related validity in CTT, item validity, item difficulty, item and person separation reliability, and item and person separation index using Rasch models. Moreover, multigroup CFA based on CTT and DIF based on Rasch were used to test the measurement invariance across gender and hours spent online gaming per week. The reasons of using both CTT and Rasch were twofold: first, CTT results were used to compare with previous studies; second, Rasch results were reported to provide additional psychometric information on sample-independent findings.

Floor and ceiling effects were computed using percentages for lowest and highest scores in the IGDS-SF9, with a percentage of <20.0% being acceptable (Jette, Warren, & Wirtalla, 2005). Test–retest reliability was conducted using Pearson's correlation, and an $r > .4$ is considered adequate (Lin et al., 2012). Internal consistency was demonstrated using Cronbach's α with >.7 considered satisfactory (Cheng, Luh, Yang, Su, & Lin, 2016). Corrected item–total correlation was computed with a value of >.4 considered acceptable (Wang, Wang, & Shee, 2007). Standard error of measurement was calculated with smaller values suggesting better reliability (Cheng et al., 2016). Criteria-related validity was examined using four regression models: using the IGDS-SF9 as the dependent variable, each criterion (depression, anxiety, stress, and hours spent online gaming per week) as the independent variable in each regression model, and age and gender as the confounders to test the associations between the criteria and the IGDS-SF9. In addition to using the standardized coefficient (β) to examine the criteria-related validity, the ΔR^2 (i.e., R^2 of the criterion plus confounders minus the R^2 of the confounders) was performed to demonstrate magnitudes of the associations. The association for hours spent online gaming per week was expected to be stronger than other associations because most diagnostic IGD criteria are related to the hours spent online gaming (e.g., preoccupation/obsession with Internet games, more time needs to be spent playing the games, fail to stop playing Internet games, etc.)

CFA with the estimator of diagonally weighted least squares was used to test construct validity, and the following fit indices were used to determine whether the construct validity was supported or not: CFI and TLI >0.9, RMSEA and SRMR <0.08, non-significant χ^2 (Cheng et al., 2016; Lin et al., 2012). AVE and CR were computed using the factor loadings and the uniqueness values in the CFA; AVE > 0.5 and CR > 0.6 suggest excellent construct validity and reliability, respectively, of the IGDS-SF9 (Bagozzi & Yi, 1988; Fornell & Larcker, 1981).

Multigroup CFA included three models: configural model (Model 1; i.e., a first-order IGDS-SF9 framework with all nine items loading on the underlying concept of IGD); metric invariance model that constrained all factor loadings being equal across groups (Model 2); scalar invariance model that constrained all factor loadings and item intercepts being equal across groups (Model 3). The three models were applied to two sets of groups: gender (male vs. female) and hours spent online gaming per week (below median hours vs. above median hours). Measurement invariance in the multigroup CFA was assessed by comparing Model 1 with 2 for metric invariance and Model 2 with 3 for scalar invariance (Bagheri, Jafari, Tashakor, Kouhpayeh, & Riazi, 2014), with the following fit indices suggest invariance: ΔCFI > −0.01, ΔSRMR < 0.01, and ΔRMSEA < 0.015 (Chen, 2007). However, if the metric or scalar invariance was not supported, some factor loadings or item intercepts across groups were relaxed to test for the partial invariance. Analysis by the present authors identified the discrepancies of the factor loadings or item intercepts between the two groups, and then relaxed the greatest discrepancy one at a time. Relaxing any factor loadings or item intercepts were stopped when fit indices were achieved.

A Rasch partial credit model was used to report the item difficulty with an interval unit (i.e., *logit*, a higher *logit* represents a more difficult item) and examine the item validity for the IGDS-SF9 items. Information-weighted fit statistic (infit) mean square (MnSq), and outlier-sensitive fit statistic (outfit) MnSq were used, and both ranges between 0.5 and 1.5 suggested a good fit (Jafari, Bagheri, & Safe, 2012). If an item has MnSq < 0.5, the item could be a redundant item; if an item has MnSq > 1.5, the item could be an out-of-concept item (Khan, Chien, & Brauer, 2013). In

addition, item reliability and person reliability were conducted using Rasch analysis with a value >0.7 suggesting good reproducibility of hierarchical item difficulty and person ordering on respondents' ability (Chang, Wang, Tang, Cheng, & Lin, 2014; Jafari, Bagheri, Ayatollahi, & Soltani, 2012). Item and person indices were also conducted using a value >2.0 suggesting that the IGDS-SF9 can separate individuals (person separation index) or items (item separation index) into more than two distinct groups (Kook & Varni, 2008).

In addition, DIF in the Rasch models was used to corroborate measurement invariance results from multigroup CFA, especially at item level. That is, DIF was tested for each item; respondents in a particular group (i.e., different gender and different hours spent online gaming per week in this study) are easier or harder than the respondents in the other group to fulfill the item description if the item displays DIF (Lin, Yang, Lai, Su, & Wang, 2015). Moreover, DIF contrasts (i.e., the difference of difficulty between two groups) with a value >0.5 were substantial DIF (Shih & Wang, 2009), and this criterion was used to demonstrate DIF items.

All the descriptive and CTT analyses were conducted using IBM SPSS 23.0 (IBM Corp., Armonk, NY), except for the CFAs, which were analyzed using the lavaan package (Rosseel et al., 2015) in R software. In addition, the Rasch models were done using WINSTEPS 3.75.0.

Ethics

All adolescents (and their parents) gave written informed consent to participate. This study was approved by the Ethics Committee of Qazvin University of Medical Sciences and the Organization for Education at Qazvin.

RESULTS

The characteristics of the participants are presented in Table 1. Their mean (*SD*) age was 15.6 (1.2) years with 18 (5) hr on Internet gaming per week. Median hours spent online gaming per week was 19 hr. More specifically, 1,323 (56.0%) spent ≤19 hr online gaming per week and 1,040 (44.0%) spent ≥20 hr online gaming per week. Moreover, nearly two thirds of the participants were male ($n = 1{,}531$, 64.8%).

Table 1. Participants' characteristics ($N = 2{,}363$)

	Mean ± *SD* or *n* (%)
Age (years)	15.6 ± 1.2
Gender (male)	1,531 (64.8)
Fathers' educational year	7.7 ± 3.9
Mothers' educational year	6.3 ± 3.6
Currently smoker (yes)[a]	669 (28.3)
Score in Internet Gaming Disorder Scale – Short Form	24.0 ± 7.1
Score in depression[b]	7.6 ± 4.7
Score in anxiety[b]	8.2 ± 4.8
Score in stress[b]	7.6 ± 5.0
Weekly hours on Internet gaming[c]	18.0 ± 5.0

[a]Fourteen participants did not report their smoking status. [b]Measured using Depression Anxiety Stress Scales. [c]Median weekly hours were 19 hr; 1,323 (56.0%) spent ≤19 hr on Internet gaming; 1,040 (44.0%) spent ≥20 hr on Internet gaming.

Both CFA and Rasch analyses demonstrated promising psychometric properties of the IGDS-SF9. At the scale level, ceiling effects (1.8%) and floor effects (0.8%) were trivial, the internal consistency was excellent ($\alpha = .90$), and the test–retest reliability was adequate ($r = .87$). CFA showed satisfactory fit in the proposed fit statistics: CFI = 0.990, TLI = 0.987, RMSEA = 0.061, and SRMR = 0.049. AVE was 0.47, which is close to the recommended cutoff of 0.5. CR was 0.89, which is substantially over the recommended cutoff of 0.6. Although a cutoff for standard error of measurement was not proposed, the present authors believe that our value (2.27) is adequate to support the reliability of the IGDS-SF9. In terms of the psychometric properties tested using Rasch analyses, all were satisfactory. Item separation reliability was 1.0, person separation reliability was 0.86, item separation index was 28.79, and the person separation index was 2.50 (Table 2).

At the item level of the IGDS-SF9, all factor loadings derived from the CFA were significant with standardized values ranged between 0.56 and 0.80. The corrected item–total correlations were between 0.54 and 0.74, and the test–retest reliability coefficients were between 0.79 and 0.91. Rasch analyses showed all infit and outfit MnSq being acceptable (ranging between 0.79 and 1.37 for infit, and between 0.74 and 1.34 for outfit) with the item difficulties ranging between −1.06 and 1.57 (Table 3).

No substantial DIF was displayed for all IGDS-SF9 items, apart from Item 4 (DIF contrast = −0.55) across gender (i.e., *fail to control or cease gaming activities*). Furthermore, two DIF items were found across time spent online gaming per week (DIF contrast = −0.67 for Item 4, 0.61 for Item 9, i.e., *jeopardize or lose an important thing because of gaming activity*; Table 3). Measurement invariance was fully supported for the IGDS-SF9 across gender in metric invariance (ΔCFI = −0.002, ΔSRMR = 0.004, ΔRMSEA = 0.001) and scalar invariance (ΔCFI = −0.002, ΔSRMR = 0.003, ΔRMSEA = −0.002). However, metric invariance across hours spent online gaming per week was not supported if all factor loadings being equal across two groups were constrained (ΔCFI = −0.032, ΔSRMR = 0.027, ΔRMSEA = 0.035). Therefore, two factor loadings were relaxed one at a time (first for Item 4, second for Item 9), and partial metric invariance was examined again. Partial metric invariance was then supported (ΔCFI = −0.006, ΔSRMR = 0.006, ΔRMSEA = 0.007) and the scalar invariance with all item intercepts constrained being equal across groups was also supported (ΔCFI = −0.009, ΔSRMR = 0.009, ΔRMSEA = 0.005; Table 4).

Criteria-related validity was supported by the DASS and the hours spent online gaming per week as shown by the significant associations ($p < .001$). Also, the magnitudes between the criteria and the IGDS-SF9 were expected: the associations between IGDS-SF9 and depression ($\beta = .139$, $\Delta R^2 = .02$); between IGDS-SF9 and anxiety ($\beta = .148$, $\Delta R^2 = .02$); and between IGDS-SF9 and stress ($\beta = .103$, $\Delta R^2 = .01$) were weaker than that between IGDS-SF9 and hours spent online gaming per week ($\beta = .663$, $\Delta R^2 = .39$; Table 5).

Table 2. Psychometric properties of the Internet Gaming Disorder Scale – Short Form (IGDS-SF) in item level

	Analyses from classical test theory			Analyses from Rasch				
Item no.	Factor loadings[a]	Item–total correlation	Test–retest reliability[b]	Infit MnSq	Outfit MnSq	Difficulty	DIF contrast across gender[c,d]	DIF contrast across time on gaming[c,e]
IGDS-SF1	0.68	0.63	0.80	1.08	1.05	1.57	−0.02	−0.10
IGDS-SF2	0.80	0.74	0.83	0.79	0.74	1.42	0.30	0.14
IGDS-SF3	0.77	0.72	0.87	0.83	0.77	−0.17	0.11	−0.07
IGDS-SF4	0.56	0.54	0.84	1.37	1.34	0.26	−0.55	−0.67
IGDS-SF5	0.79	0.74	0.79	0.79	0.86	−0.46	0.09	0.34
IGDS-SF6	0.69	0.65	0.83	1.01	1.00	−0.29	0.07	−0.15
IGDS-SF7	0.63	0.60	0.91	1.15	1.06	−0.54	−0.22	−0.35
IGDS-SF8	0.77	0.72	0.88	0.84	0.80	−0.73	0.21	0.23
IGDS-SF9	0.66	0.62	0.79	1.16	1.17	−1.06	0.00	0.61

Note. MnSq: mean square error; DIF: differential item functioning.
[a]Based on confirmatory factor analysis. [b]Using Pearson's correlation. [c]DIF contrast >0.5 indicates substantial DIF. [d]DIF contrast across gender = difficulty for females − difficulty for males. [e]DIF contrast across time on gaming = difficulty for participants with median weekly hours or below on gaming (i.e., ≤19 hr) − difficulty for participants with above median weekly hours on gaming (i.e., ≥20 hr).

Table 3. Psychometric properties of the Internet Gaming Disorder Scale – Short Form at scale level

Psychometric testing	Value	Suggested cutoff
Ceiling effects (%)	1.8	<20.0
Floor effects (%)	0.8	<20.0
Internal consistency (Cronbach's α)	.90	>.7
Confirmatory factor analysis		
χ^2 (*df*)	260.15 (27)*	Non-significant
Comparative fit index	0.990	>0.9
Tucker–Lewis index	0.987	>0.9
Root mean square error of approximation	0.061	<0.08
Standardized root mean square residual	0.049	<0.08
Average variance extracted	0.47	>0.5
Composite reliability	0.89	>0.6
Standard error of measurement	2.27	The smaller the better
Item separation reliability from Rasch	1.00	>0.7
Item separation index from Rasch	28.79	>2.0
Person separation reliability from Rasch	0.86	>0.7
Person separation index from Rasch	2.50	>2.0
Test–retest reliability by Pearson's correlation	0.87	>0.4

*$p < .001$.

Table 4. Concurrent validity of the Internet Gaming Disorder Scale – Short Form using regression models with adjustment for age and gender

Criterion	β (*p* value)	ΔR^2	Overall R^2 (adjusted R^2)
Depression	.139 (<.001)	.02	.137 (.136)
Anxiety	.148 (<.001)	.02	.133 (.132)
Stress	.103 (<.001)	.01	.130 (.129)
Weekly hours on Internet gaming	.663 (<.001)	.39	.509 (.508)

Note. Depression, anxiety, and stress were measured using Depression Anxiety Stress Scales. ΔR^2 = overall R^2 − the R^2 derived from age and gender; that is, indicating the explained variance of the criterion.

DISCUSSION

This study demonstrated that the Persian IGDS-SF9 had a one-factor structure stably (or unidimensionality), and confirmed all previous validation findings (Monacis et al., 2016; Pontes & Griffiths, 2015, 2016). The Persian IGDS-SF9 also demonstrated robust psychometric properties including acceptable internal consistency, adequate test–retest reliability, promising criteria-related validity, and minimal floor and ceiling effects. Moreover, this is the first study to investigate the effects of gender and time spent online gaming per week in relation to IGDS-SF9 scoring. It was found that gender did not appear to affect the response patterns, while the time spent online gaming per week might have small effects. Thus, factorial structure and order of

Table 5. Measurement invariance across gender and across weekly hours on Internet gaming through confirmatory factor analysis

Model and comparisons	Fit statistics							
	χ^2 (*df*)	$\Delta\chi^2$ (Δdf)	CFI	ΔCFI	SRMR	ΔSRMR	RMSEA	ΔRMSEA
Gender								
M1: Configural	348.73 (54)*		0.985		0.053		0.068	
M2: Plus all loadings constrained	408.57 (62)*		0.983		0.057		0.069	
M3: Plus all intercepts constrained	444.30 (70)*		0.981		0.060		0.067	
M2−M1		59.84 (8)*		−0.002		0.004		0.001
M3−M2		35.73 (8)*		−0.002		0.003		−0.002
Weekly hours on Internet gaming[a]								
M1: Configural	367.50 (54)*		0.979		0.055		0.070	
M2: Plus all loadings constrained	865.16 (62)*		0.947		0.082		0.105	
M2P: Partial constrained loadings[b]	477.16 (60)*		0.973		0.061		0.077	
M3: Plus all intercepts constrained[b]	611.45 (68)*		0.964		0.070		0.082	
M2−M1		497.66 (8)*		−0.032		0.027		0.035
M2P−M1		109.66 (8)*		−0.006		0.006		0.007
M3−M2P		134.29 (8)*		−0.009		0.009		0.005

Note. M1: Model 1, a configural model; M2: Model 2, a model based on M1 with all factor loadings constrained being equal across groups; M2P: Model 2 with partial invariance, a model based on M2 with some factor loadings relaxed across groups; M3: Model 3, a model based on M2 or M2P with all item intercepts constrained being equal across groups. CFI: comparative fit index; SRMR: standardized root mean square residual; RMSEA: root mean square error of approximation.

[a]Median weekly hours were 19 hr.

[b]Factor loadings of Items 4 and 9 were relaxed across two groups.

*$p < .05$.

item difficulties of the IGDS-SF9 were the same across gender, which made scores of the IGDS-SF9 comparable across gender. However, the comparisons of IGDS-SF9 scores between different time spans spent online gaming should be interpreted with caution.

The unidimensionality of the IGDS-SF9 was supported by both CFA and Rasch analysis, and accords with previous studies (Monacis et al., 2016; Pontes & Griffiths, 2015, 2016). The unidimensionality indicated that items in the IGDS-SF9 belonged to the same concept and each item score can be summed up as an overall score to represent the severity of IGD (i.e., the higher the score, the more severe the IGD). A Rasch analysis was also used to estimate the item difficulties (i.e., severity of the IGD). When individuals respond with "often" or "very often" on items with higher item difficulties, the individual's IGD symptoms were more severe. These results are helpful for clinicians to get an initial assessment about an individual' IGD severity quickly.

The floor and ceiling effects were minimal implying that the IGDS-SF9 is sufficient in discriminating the degree of IGD severity (e.g., high vs. low) among high-school students. These characteristics ensured the utility of the IGDS-SF9 because the scale has the potential to apply across a wide range of individual IGD severity. Given that the participants in this study were recruited in a community (i.e., high-school students), it is tentatively concluded that very few might be real patients diagnosed with IGD. Further studies are recommended to investigate the ceiling effect of the IGDS-SF9 in populations that contain a higher proportion of individuals potentially having IGD.

The CFA results indicated that gender did not affect the interpretations of IGDS-SF9 items, but Rasch analysis showed that Item 4 (i.e., *fail to control or cease gaming activities*) might be interpreted differently. In Rasch analysis, the DIF contrast of Item 4 was −0.55, indicating that Item 4 was more difficult for males. The reason why it is more difficult for the males to control or cease gaming activities may be because males tend to be characterized by lower self-control compared with females (Mitchell & Potenza, 2015; Struber et al., 2008). Nonetheless, it is recommended that Item 4 should be retained rather than deleting it based on the following reasons. First, there is a sense of control reflecting severity of addictive behaviors (APA, 2013), meaning that Item 4 is believed to be representative for individuals with IGD. Second, the results of CFA supported the invariance of the Item 4. Third, the value of DIF was close to our criterion (0.55–0.5). For future use of the IGDS-SF9, users may need to ensure whether DIF occurs on Item 4 and then compare the results of the different genders.

The effect of hours on Internet games was found in both Rasch analysis and CFA on Item 4 (*fail to control or cease gaming activities*) and Item 9 (*jeopardize or lose an important thing because of gaming activity*). In Rasch analysis, the DIF contrast was −0.67 for Item 4 and 0.61 for Item 9. These results imply that individuals spending ≤19 hr online gaming per week more often felt out of control and lost important things less often than individuals spending >19 hr online gaming per week. One possible reason is that individuals spending less time online gaming per week had the insight to maintain reasonable gaming hours and would make the effort to quit, but they may find it difficult to quit. In contrast, individuals spending >19 hr online gaming per week did not have enough time and energy to deal with other issues in their life; therefore, they may not care about the jeopardy or loss of something important in their lives. Given that these two items fitted our expectations and the partial

invariance achieved for the IGDS-SF9, it is suggested that both items are retained rather than deleting them. However, should future users who want to compare both item scores for those with different time spans spent online gaming, they are cautioned with the displayed DIF. Investigating the proper weighted values of the two items is also recommended.

This study is not without limitations. First, the cutoff point of time spent online gaming was performed according to the participants' median. The cutoff may differ from other populations and the results of this study may not be generalized to those with different medians. Second, no history was recorded as to when the participants started to play online games, affecting the inference of association among IGD, depression, anxiety, and stress. Third, there was no diagnosis of IGD for each participant. Therefore, this study was unable to calculate sensitivity and specificity of the IGDS-SF9 in detecting IGD in adolescents. Fourth, this study only investigated the psychometric properties in high-school students, which may restrict the generalization of the results to people in other age groups (e.g., adults). Future studies should explore the psychometric properties of the Persian version of the IGDS-SF9 in adults. Fifth, the AVE of the IGDS-SF9 was slightly lower than the criterion (0.5), implying that the items of the IGDS-SF9 did not sufficiently contribute to the variance of participants' IGD level. Thus, the true score of participants' IGD level may be underestimated or overestimated. Finally, because only Iranian adolescents were recruited in this study, we could not well compare the IGDS-SF9 across Western and Eastern countries. Future studies may want to explore the DIF across countries for the IGDS-SF9.

CONCLUSIONS

This study showed that all items of the Persian IGDS-SF9 were embedded in a one-factor structure. The Persian IGDS-SF9 also had robust psychometric properties. However, the effects of gender (Item 4) and time spent online gaming per week (Items 4 and 9) on the interpretations of the IGDS-SF9 items were found. Thus, further studies are warranted to investigate the proper weighted value of the two items. Consequently, the sum scores can be properly compared in different genders and in individuals spending different weekly hours online gaming.

Funding sources: Nothing declared.

Authors' contribution: AHP created and organized this study and collected the data; T-YW wrote the first draft; C-YL analyzed and interpreted the data; KÅ provided the directions of data analysis; MDG supervised the entire study; AHP, C-YL, KÅ, MDG, and AB critically reviewed the manuscript and provided constructive comments.

Conflict of interest: MDG is the developer of the original version of the Internet Gaming Disorder Scale – Short Form (IGDS-SF9). All authors report no financial or other relationship relevant to the subject of this study.

REFERENCES

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

Asghari, A., Saed, F., & Dibajnia, P. (2008). Psychometric properties of the Depression Anxiety Stress Scales-21 (DASS-21) in a non-clinical Iranian sample. *International Journal of Psychology, 2*(2), 82–102.

Bagheri, Z., Jafari, P., Tashakor, E., Kouhpayeh, A., & Riazi, H. (2014). Assessing whether measurement invariance of the KIDSCREEN-27 across child-parent dyad depends on the child gender: A multiple group confirmatory factor analysis. *Global Journal of Health Science, 6*(5), 142–153. doi:10.5539/gjhs.v6n5p142

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science, 16*(1), 74–94. doi:10.1007/BF02723327

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*(24), 3186–3191. doi:10.1097/00007632-200012150-00014

Chang, K.-C., Wang, J.-D., Tang, H.-P., Cheng, C.-M., & Lin, C.-Y. (2014). Psychometric evaluation, using Rasch analysis, of the WHOQOL-BREF in heroin-dependent people undergoing methadone maintenance treatment: Further item validation. *Health and Quality of Life Outcomes, 12*(1), 148. doi:10.1186/s12955-014-0148-6

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504. doi:10.1080/10705510701301834

Cheng, C.-P., Luh, W.-M., Yang, A.-L., Su, C.-T., & Lin, C.-Y. (2016). Agreement of children and parents scores on Chinese version of Pediatric Quality of Life Inventory Version 4.0: Further psychometric development. *Applied Research in Quality of Life, 11*(3), 891–906. doi:10.1007/s11482-015-9405-z

Crews, F., He, J., & Hodge, C. (2007). Adolescent cortical development: A critical period of vulnerability for addiction. *Pharmacology Biochemistry and Behavior, 86*(2), 189–199. doi:10.1016/j.pbb.2006.12.001

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. doi:10.2307/3151312

Griffiths, M. D., Kuss, D. J., & Pontes, H. (2016). A brief overview of Internet gaming disorder and its treatment. *Australian Clinical Psychologist, 2*(1), 20108.

Jafari, P., Bagheri, Z., Ayatollahi, S. M. T., & Soltani, Z. (2012). Using Rasch rating scale model to reassess the psychometric properties of the Persian version of the PedsQL TM 4.0 Generic Core Scales in school children. *Health and Quality of Life Outcomes, 10*(1), 27. doi:10.1186/1477-7525-10-27

Jafari, P., Bagheri, Z., & Safe, M. (2012). Item and response-category functioning of the Persian version of the

KIDSCREEN-27: Rasch partial credit model. *Health and Quality of Life Outcomes, 10*(1), 127. doi:10.1186/1477-7525-10-127

Jette, D. U., Warren, R. L., & Wirtalla, C. (2005). Functional independence domains in patients receiving rehabilitation in skilled nursing facilities: Evaluation of psychometric properties. *Archives of Physical Medicine and Rehabilitation, 86*(6), 1089–1094. doi:10.1016/j.apmr.2004.11.018

Khan, A., Chien, C.-W., & Brauer, S. G. (2013). Rasch-based scoring offered more precision in differentiating patient groups in measuring upper limb function. *Journal of Clinical Epidemiology, 66*(6), 681–687. doi:10.1016/j.jclinepi.2012.12.014

Kim, N. R., Hwang, S. S.-H., Choi, J.-S., Kim, D.-J., Demetrovics, Z., Király, O., Nagygyörgy, K., Griffiths, M. D., Hyun, S. Y., Youn, H. C., Choi, S. W., & Youn, H. C. (2016). Characteristics and psychiatric symptoms of Internet gaming disorder among adults using self-reported DSM-5 criteria. *Psychiatry Investigation, 13*(1), 58–66. doi:10.4306/pi.2016.13.1.58

Király, O., Griffiths, M. D., & Demetrovics, Z. (2015). Internet gaming disorder and the DSM-5: Conceptualization, debates, and controversies. *Current Addiction Reports, 2*(3), 254–262. doi:10.1007/s40429-015-0066-7

Ko, C.-H., Yen, J.-Y., Chen, C.-C., Chen, S.-H., & Yen, C.-F. (2005). Gender differences and related factors affecting online gaming addiction among Taiwanese adolescents. *The Journal of Nervous and Mental Disease, 193*(4), 273–277. doi:10.1097/01.nmd.0000158373.85150.57

Kook, S. H., & Varni, J. W. (2008). Validation of the Korean version of the pediatric quality of life inventory™ 4.0 (PedsQL™) generic core scales in school children and adolescents using the Rasch model. *Health and Quality of Life Outcomes, 6*(1), 41. doi:10.1186/1477-7525-6-41

Lin, C.-Y., Luh, W.-M., Yang, A.-L., Su, C.-T., Wang, J.-D., & Ma, H.-I. (2012). Psychometric properties and gender invariance of the Chinese version of the self-report Pediatric Quality of Life Inventory version 4.0: Short form is acceptable. *Quality of Life Research, 21*(1), 177–182. doi:10.1007/s11136-011-9928-1

Lin, C.-Y., Yang, S.-C., Lai, W.-W., Su, W.-C., & Wang, J.-D. (2017). Rasch models suggested the satisfactory psychometric properties of the World Health Organization Quality of Life–Brief among lung cancer patients. *Journal of Health Psychology, 22*(4), 397–408. doi:10.1177/1359105315603474

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33*(3), 335–343. doi:10.1016/0005-7967(94)00075-U

Mitchell, M. R., & Potenza, M. N. (2015). Importance of sex differences in impulse control and addictions. *Frontiers in Psychiatry, 6,* 24. doi:10.3389/fpsyt.2015.00024

Monacis, L., Palo, V., Griffiths, M. D., & Sinatra, M. (2016). Validation of the Internet Gaming Disorder Scale – Short-Form (IGDS9-SF) in an Italian-speaking sample. *Journal of Behavioral Addictions, 5*(4), 683–690. doi:10.1556/2006.5.2016.083

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. doi:10.1016/j.chb.2014.12.006

Pontes, H. M., & Griffiths, M. D. (2016). Portuguese validation of the Internet Gaming Disorder Scale–Short-Form. *Cyberpsychology, Behavior, and Social Networking, 19*(4), 288–293. doi:10.1089/cyber.2015.0605

Pontes, H. M., Király, O., Demetrovics, Z., & Griffiths, M. D. (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

Rehbein, F., Kliem, S., Baier, D., Mossle, T., & Petry, N. M. (2015). Prevalence of Internet gaming disorder in German adolescents: Diagnostic contribution of the nine DSM-5 criteria in a state-wide representative sample. *Addiction, 110*(5), 842–851. doi:10.1111/add.12849

Rosseel, Y., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Merkle, E., Hallquist, M., Rhemtulla, M., Katsikatsou, M., Barendse, M., Chow, M., & Jorgensen, T. (2015). Package 'lavaan'. Retrived January 1, 2017, from https://cran.r-project.org/web/packages/lavaan/lavaan.pdf

Shih, C.-L., & Wang, W.-C. (2009). Differential item functioning detection using the multiple indicators, multiple causes method with a pure short anchor. *Applied Psychological Measurement, 33*(3), 184–199. doi:10.1177/0146621608321758

Struber, D., Luck, M., & Roth, G. (2008). Sex, aggression and impulse control: An integrative account. *Neurocase, 14*(1), 93–121. doi:10.1080/13554790801992743

Wang, Y.-S., Wang, H.-Y., & Shee, D. Y. (2007). Measuring e-learning systems success in an organizational context: Scale development and validation. *Computers in Human Behavior, 23*(4), 1792–1808. doi:10.1016/j.chb.2005.10.006