CrossMark

ORIGINAL ARTICLE

# Measurement Invariance of the Nine-Item Internet Gaming Disorder Scale (IGDS9-SF) Across Albania, USA, UK, and Italy

Valeria de Palo[1] · Lucia Monacis[1] · Maria Sinatra[2] ·
Mark D. Griffiths[3] · Halley Pontes[3] · Menada Petro[4] ·
Silvana Miceli[5]

**Abstract** The IGDS9-SF, which assesses Internet Gaming Disorder behaviors, has been validated in a number of countries (Portugal, Italy, Iran, Slovenia), although the psychometric equivalence of the instrument has been assessed only across Australia, the USA, the UK, and India. This research aimed at providing further cross-cultural insights into IGD by assessing the factorial structure of the IGDS9-SF in Albania and investigating its measurement invariance across Albanian, Italian, American, and British gamers. Multi-Group Confirmatory Factor Analyses were performed on a sample of 1411 participants from Albania ($n = 228$), USA ($n = 237$), the UK ($n = 275$), and Italy ($n = 671$). The CFAs confirmed the single-factor structure in the four countries. Measurement invariance supported the configural invariance and partially supported the metric and scalar invariance. Overall, the findings provided evidence for the underlying factor assessing IGD across the countries, although the specific meaning of the construct was non-identical.

✉ Valeria de Palo
  valeria.depalo@unifg.it

✉ Mark D. Griffiths
  mark.griffiths@ntu.ac.uk

  Lucia Monacis
  lucia.monacis@unifg.it

  Maria Sinatra
  maria.sinatra@uniba.it

  Halley Pontes
  halley.pontes@ntu.ac.uk

🙌 Springer

After two decades of research into problematic videogame playing and videogame addiction, Internet Gaming Disorder (IGD) has been recognized by the American Psychiatric Association (APA) as a tentative mental disorder that merits further consideration by clinicians and researchers. Included in 2013 in the appendix (i.e., Section III, "Emerging Measures and Models") of the updated (fifth) version of the *Diagnostic and Statistical Manual for Mental Disorders* (DSM-5), the phenomenon has been defined as a "persistent and recurrent use of the Internet to engage in games, often with other players, leading to clinically significant impairment or distress" (APA 2013, p. 795).

Despite the fact that research on the topic has increased in both quantity and quality, the lack of robust assessment tools has led to the development of various psychometric instruments to assess different aspects of IGD (Griffiths et al. 2014; Kuss 2013). In their review, King et al. 2013) reported the existence of 18 different screening instruments, of which only one, the Problem Video Game Playing Questionnaire (PVP; Tejeiro Salguero and Moran 2002), included the nine IGD diagnostic criteria suggested by the DSM-5. The persistent inconsistency and lack of uniformity of the IGD psychometric tools not only highlight the "chaos and confusion" in IGD conceptualization and measurement but also stress the need for an international consensus on IGD diagnosis (Kuss et al. 2017). Petry and O'Brien (2013) argued that the inclusion of IGD as a separate mental disorder would be possible only after the identification of its features, cross-cultural validation of its specific criteria, determination of its prevalence rates in representative samples across the world, and the evaluation of its biological underpinnings.

The nine-item Internet Gaming Disorder Scale-Short Form (IGDS9-SF; Pontes and Griffiths 2015) was developed to address the need for a unified and robust standardized assessment of IGD which takes account of the nine diagnostic criteria proposed by the DSM-5. The instrument has now been widely used in a number of countries and employed in many research studies across different fields, from clinical and cognitive psychology to sociology and human-computer interaction (e.g., Ko 2014; Lin et al. 2015; Monacis et al. 2017; Pontes et al.2017). Even though the unidimensional factor solution of the scale has been replicated and confirmed in Portuguese (Pontes and Griffiths 2016), Italian (Monacis et al. 2016), Persian (Wu et al. 2017), and Slovenian (Pontes et al. 2016) samples, these findings

Menada Petro
menada_petro@yahoo.com

Silvana Miceli
silvana.miceli56@unipa.it

[1]  Department of Humanities, Centre for Advanced Studies on Cyberpsychology and Ethics (ASPEN), University of Foggia, via Arpi, 175, 71100 Foggia, Italy

[2]  Department of Educational Sciences, Psychology, Communication, Centre for Advanced Studies on Cyberpsychology and Ethics (ASPEN), University of Bari Aldo Moro, Bari, Italy

[3]  International Gaming Research Unit, Division of Psychology, Nottingham Trent University, Nottingham, UK

[4]  Education Faculty; "Aleksander Moisiu", University of Durrës, Durrës, Albania

[5]  Dipartimento di Scienze Psicologiche, Pedagogiche e della Formazione, University degli Studi di Palermo, Palermo, Italy

have not provided sufficient support for the psychometric equivalence of the instrument across countries. To date, only two studies have assessed the measurement invariance (MI) of the IGDS9-SF. More specifically, Stavropoulos et al. (2017) examined the MI of the scale across gamers from Australia, the USA, and the UK, and Pontes et al. 2017assessed the MI across gamers from the USA, India, and the UK.

As response biases in self-reported measures may influence results in psychological research, especially when data are gathered from two or more cultural groups, testing for equivalence of measures makes it possible to verify whether the members of different groups or cultures ascribe the same meanings to the items of a questionnaire (Milfont and Fischer 2010). Four levels of equivalence have been distinguished (e.g., Fontaine 2005; Kühne 2013): *functional equivalence*, which assesses the existence of the construct in all groups considered; structural equivalence or *configural invariance*, which verifies whether the number of factors and the patterns of free factor loadings hold across the groups; metric equivalence or *metric invariance*, which determines whether the loading weights are equivalent across groups and whether the items assess their relevant latent factor using the same metric scale; and full score or *scalar invariance*, which reveals whether intercepts are equivalent across groups (i.e., if the individuals in different groups endorse the same observed level or response category for the same level of the latent trait). MI is generally assessed by performing Multi-Group Confirmatory Factor Analyses (MGCFAs), in which the theoretical model is compared with the observed structure in two or more samples (Milfont and Fischer 2010). According to Jöreskog's (1971, 1993) strategy, nested models are organized in a hierarchical ordering that implies adding parameter constraints one at a time, with the increasingly restrictive models being tested in terms of the fit of the data to the model.

Given "the increasing demand for cross-cultural studies to be carried out using the nine diagnostic criteria for IGD" for IGDS9-SF (Stavropoulos et al. 2017, p. 5), the aim of the present study was to further examine the MI of the instrument across four countries: Albania, Italy, the UK, and the USA. Although the previous studies on the MI of the instrument have investigated two of these four countries (i.e., UK, USA), the data obtained have not yet been compared with those from Italy and Albania. The present research was thus justified for various reasons. First, the instrument was psychometrically assessed in the language fielded in the respective countries. Second, with regard to gaming revenue, in 2017, the USA, the UK, and Italy were ranked within the top 10 countries worldwide (second, fifth, and tenth, respectively), whereas Albania was ranked 94th (Global Games Market Report 2017). Third, with regard to Albania, recent studies have analyzed only the phenomenon of Internet addiction reporting high levels of addiction to the Internet, as well as significant relationships between Internet addiction, self-esteem, shyness, loneliness, locus of control, and anxiety (Agaj 2015, 2016; Agaj and Marku 2015; Hasmujaj 2016; Melonashi 2017). However, to the best of the authors' knowledge, IGD has not yet been investigated. Fourth, cultural differences are expected to influence the way psychopathology is perceived and reported.

When considering cultural variations, given that the individual and the self are conceived as social constructions, the existence of universal dimensions of individual differences which could be assessed equivalently across cultures has been questioned (Walford et al. 2010). For instance, a well-known distinction in cultural orientations involves "individualistic" vs. "collectivistic" differences. These two cultural orientations prescribe social norms, values, and beliefs that direct the individuals' cognition, attitudes, and behavior (Hofstede 1980, 2001). In individualistic cultures, self-fulfillment and self-preservation are emphasized. Individuals are autonomous and independent from their in-groups, strive for their own success by giving

priority to their personal goals, and behave primarily on the basis of their attitudes rather than of the social norms. On the other hand, in collectivistic cultures, individuals are interdependent within their in-groups, give priority to communal goals of their in-groups, shape their behavior in accordance with the in-group norms, and are concerned with relationships (Triandis 2001). In this context, cultural psychologists have warned against the extent to which the differences between individualistic and collectivistic cultures may affect the response patterns and the measurement errors of psychometric tools. Accordingly, traditional rating scales may be less reliable and valid in collectivistic in comparison to individualistic cultures since self-concepts are less clear and behaviors may be more determined by social roles, relationships, and norms in the collectivism orientation (Walford et al. 2010).

In light of such cultural differences, there are reasons to assume that the way in which IGD is experienced and reported by gamers may differ between Albanian, Italian, British, and American cultures. In fact, the social structures and policies of the UK, the USA, and Italy reflect an individualistic culture, whereas Albanian culture reflects a collectivistic orientation (Hofstede et al. 2010). Following on from previous research (Stavropoulos et al. 2017; Pontes et al. 2017), the present study sought to provide further cross-cultural insights into IGD by assessing the factorial structure of the IGDS9-SF in Albania and investigating its MI across four non-probability normative samples of Albanian, Italian, American, and British gamers. Moreover, the cross-cultural equivalence of the IGDS9-SF in the aforementioned cultures provides further insight into its cross-cultural measurement consistency, thus increasing the possibility of the worldwide utilization of the instrument in clinical contexts, thus facilitating the goal of unification in the assessment of IGD (Griffiths et al. 2014; Pontes and Griffiths 2014).

## Method

### Participants and Procedure

The sample comprised 1411 participants (age range 14–70 years, mean age = 25.94 years, SD = 8.91 years; 36.4% females) from Albania (AL; $n = 228$; age range = 18–70 years, mean age = 31.38 years, SD = 10.97; 50.9% females), USA ($n = 237$; age range = 16–69 years, mean age = 29.09 years, SD = 10.72; 21.7% females), the UK ($n = 275$; age range = 16–70 years, mean age = 29.50 years, SD = 9.48; 13.9% females), and Italy (IT; $n = 671$; age range = 14–46 years, mean age = 21.62 years, SD = 3.9; 45.4% females). In relation to data collection methods, the Albanian and Italian participants were recruited from schools, universities, and gaming halls, whereas English-speaking gamers from the USA and the UK were recruited online by advertising the link of the study in popular online gaming forums (e.g., https://us.battle.net/forums/en/wow/; https://www.ea.com/forums). The study was approved by the Ethics Committees of the relevant institutions. The IGDS9-SF was translated from English into Italian and Albanian separately by the Italian and Albanian authors of the present study, following the standard guidelines from Merenda (2006). After translating the IGDS9-SF to Italian and Albanian, all items were back-translated to English by a native speaker to establish their comparability. The resulting Italian and Albanian versions of the IGDS9-SF were subjected to a pilot study with a sample of 25 students to assess content and face validity of the items in order to capture eventual problems emerging from the outlined adaptation process.

## Measures

The IGDS9-SF (Pontes and Griffiths 2015) assesses the construct of IGD according to the DSM-5 criteria (APA 2013) alongside its severity and potential detrimental effects by examining both online and/or offline gaming activities occurring over a 12-month period. The nine questions comprising the IGDS9-SF are answered using a five-point Likert scale: one ("Never"), two ("Rarely"), three ("Sometimes"), four ("Often"), and five ("Very Often"). The scores are obtained by summing the individual's responses, and total scores can range from 9 to 45, with higher scores being indicative of higher degrees of Internet Gaming Disorder. Internal reliability in the present study was high and comparable across the four countries (Table 1).

## Data Management, Analytic Strategy, and Statistical Analyses

After finalizing the recruitment process, several data management steps adopted by previous similar studies (e.g., Pontes and Griffiths 2015, 2016; Pontes et al. 2016) were used to ensure the robustness of the results. First, data cleaning was conducted for each sample by inspecting cases with missing values and the assessment of the univariate and multivariate normality of all items in the IGDS9-SF. The univariate normality was checked by following the standard guidelines of Kim (2013) (i.e., absolute skew value larger than two and an absolute kurtosis larger than seven). Moreover, univariate outliers were identified by inspecting boxplots. The multivariate outliers were inspected using Mahalanobis distances and the critical value for each case based on the chi-squared ($\chi^2$) distribution values. The inspection of the cases yielded no item-level missing values in the four samples. As for univariate normality, no item showed absolute values of skewness greater than two or values of kurtosis greater than seven. A total of 29 univariate and multivariate outliers were found (and consequently removed from further analysis).

Second, descriptive statistics (i.e., minimum, maximum, mean, and standard deviation) for each sample across all nine items of the IGDS9-SF were calculated. After ensuring the assumptions of the statistical analyses were met, a series of confirmatory factor analyses (CFAs) with categorical variables was computed for each country. Following this, MGCFAs were run to establish the MI of the IGDS9-SF. Goodness of fit for the analyses was evaluated on the basis of different fit indices and the following recommended thresholds: chi-squared ($\chi^2$) and its degrees of freedom (test values associated with $p > .05$), the Comparative-of-Fit Index (CFI; values $\geq 0.90$), the root mean square error of approximation (RMSEA; values close to .06) and its 95% confidence interval (CI), and the weighted root mean square residual (WRMR; values $\leq 0.90$).

MI was tested by comparing progressively more constrained models that tested for configural, metric, and scalar invariance (Millsap and Yun-Tein 2004). The Mplus DIFFTEST

Table 1  Descriptive statistics and reliability coefficients for the IGDS9-SF for each country

|  | Minimum–maximum | Mean (SD) | MIC | $\alpha$ |
| --- | --- | --- | --- | --- |
| Albania | 9–35 | 21.10 (7.08) | 0.323 | 0.810 |
| USA | 9–38 | 17.50 (6.01) | 0.383 | 0.841 |
| UK | 9–45 | 17.99 (7.02) | 0.442 | 0.870 |
| Italy | 9–43 | 15.93 (8.96) | 0.766 | 0.963 |

SD standard deviations, MIC mean inter-item correlation, $\alpha$ Cronbach's $\alpha$ reliability coefficient

option was used to calculate and compare the fit of the different models being tested. The nested models were also compared by using cutoff values of $\Delta$CFI < 0.01 and $\Delta$RMSEA < 0.015 for metric and scalar invariance (Chen 2007; Cheung and Rensvold 2002). Modification indices (MIs) were examined to detect the source of non-invariance when full metric and/or scalar invariance is not established (Bentler 1990; Bentler and Bonnet 1980; Hooper et al. 2008; Hu and Bentler 1999). All the analyses were performed on IBM SPSS Statistics 20.0 for Windows and MPlus 7.2 (Muthén and Muthén 1998–2012).

## Results

A CFA with weighted least squares means and variance adjusted (WLSMV) estimation method was computed separately for each country. Fit indices are shown in Table 2. Overall, the results obtained for the Albanian and Italian samples showed poor fit, whereas the results for British sample showed acceptable fit. A careful inspection of the MIs suggested adding covariance paths between the errors terms of items 6 and 7 (MI = 23.96), 6 and 8 (MI = 18.52), 5 and 7 (MI = 17.95), 2 and 6 (MI = 12.47), 4 and 9 (MI = 12.99) for the Albanian sample, between the error terms of items 7 and 9 (MI = 37.06) for the British sample, and between the error terms of items 6 and 7 (MI = 51.73), 1 and 2 (MI = 40.50), 2 and 3 (MI = 43.46) for the Italian sample. These models were used as baseline models for ascertaining the measurement invariance of the IGDS9-SF across all groups. All standardized factor loadings were statistically significant ($p < .001$) and ranged from .399 to .783 for the Albanian sample, .408 to .804 for the American sample, .602 to .886 for the British sample, and .832 to .979 for the Italian sample.

The configural invariance model with factor loadings and intercepts free to vary between groups was assessed following the baseline models. The resulting model proved to have an acceptable fit, $\chi^2 = 265.752$, $df = 99$, $p < .001$; RMSEA = 0.069; CFI = 0.998; WRMR = 1.453. When metric invariance was computed, tests of model fit resulted in a significant worsening of fit, $\chi^2 = 530.187$, $df = 123$, $p < .001$; RMSEA = 0.097; CFI = 0.996; WRMR = 2.60. The inspection of MIs indicated that freeing the factor loadings of item 1 in Albanian and British samples, of items 2 in Italian sample, of item 4 in British and Italian samples, of item 7 in

**Table 2** Fit indices of confirmatory factor analysis (CFA) for each country

| | | $\chi^2$ | df | p value | RMSEA (95% CI) | CFI | WRMR |
|---|---|---|---|---|---|---|---|
| Albania | | 150.736 | 27 | .001 | 0.142 (0.120–0.164) | 0.907 | 1.261 |
| | Covariance between items 6 and 7 | 127.372 | 26 | .001 | 0.131 (0.109–0.154) | 0.924 | 1.153 |
| | Covariance between items 6 and 8 | 110.740 | 25 | .001 | 0.123 (0.100–0.146) | 0.935 | 1.062 |
| | Covariance between items 5 and 7 | 94.093 | 24 | .001 | 0.113 (0.090–0.138) | 0.947 | 0.968 |
| | Covariance between items 2 and 6 | 81.252 | 23 | .001 | 0.105 (0.081–0.131) | 0.956 | 0.899 |
| | Covariance between items 4 and 9 | 68.712 | 22 | .001 | 0.097 (0.071–0.123) | 0.965 | 0.820 |
| USA | | 65.615 | 27 | .001 | 0.078 (0.054–0.102) | 0.977 | 0.721 |
| UK | | 100.013 | 27 | .001 | 0.099 (0.079–0.120) | 0.973 | 0.895 |
| | Covariance between items 7 and 9 | 66.159 | 26 | .001 | 0.075 (0.053–0.098) | .985 | 0.700 |
| Italy | | 222.223 | 27 | .001 | 0.104 (0.091–0.117) | 0.997 | 1.206 |
| | Covariance between items 6 and 7 | 176.461 | 26 | .001 | 0.093 (0.080–0.106) | 0.998 | 1.052 |
| | Covariance between items 1 and 2 | 141.667 | 25 | .001 | 0.083 (0.07–0.097) | 0.998 | 0.919 |
| | Covariance between items 2 and 3 | 105.486 | 24 | .001 | 0.071 (0.058–0.085) | 0.999 | 0.759 |

$\chi^2$ chi square, *df* degree of freedom, *RMSEA* (95% CI) root mean square error of approximation and its 95% confidence interval, *CFI* Comparative-of-Fit Index, *WRMR* weighted root mean square residual

American and British samples, of item 9 in Italian sample, resulted in better model fit, $\chi^2 = 251.339$, $df = 113$, $p < .001$; RMSEA = 0.058; CFI = 0.999; WRMR = 1.65. However, models 2a–b were rejected using the $\Delta\chi^2$ rule. The last partial metric invariance model (model 2c) was retained on the basis of the $\chi^2$ difference test ($p > .01$) and of the $\Delta$RMSEA and $\Delta$CFI rules outlined. As for the scalar invariance, fit indices showed poor fit of the model, $\chi^2 = 1301.389$, $df = 196$, $p < .001$; RMSEA = 0.126; CFI = 0.989; WRMR = 4.39. The inspection of MIs suggested that intercepts of items 1 (for American, British, and Italian samples), 2 (for Italian sample), 3 (for Albanian and American samples), 4 (for Albanian sample), 5 (for British and Italian samples), and 8 (for American, British, and Italian samples) were non-invariant across the samples. Although $\chi^2$ differences were still significant, the values of $\Delta$RMSEA and $\Delta$CFI suggested that the final partial scalar invariance model (model 3e) could be retained (see Table 3). Overall, the factor loadings of items 1, 2, 4, 7, and 9 and the intercepts of items 1, 2, 3, 4, 5, 7, and 8 appeared to be non-invariant across the countries.

## Discussion

On the basis of previous research on the IGDS9-SF measurement invariance (MI) (Pontes et al. 2017; Stavropoulos et al. 2017), the present study psychometrically tested the single-factor model of the IGDS9-SF in Albania, as well as at comparing and examining the measurement equivalence of the instrument across Albania, Italy, the UK, and the USA. Although the unidimensional-factor solution has been established independently in Italy, the UK, and the USA (Monacis et al. 2016; Pontes and Griffiths 2015), the data have not yet been cross-culturally compared. Moreover, there is a lack of Albanian research on IGDS9-SF.

Before carrying out the MI analyses, the factorial structure of the instrument was tested separately for each country. The CFAs' results revealed that the single-factor structure was generally confirmed in the four countries, although additional covariances between some items were added to improve the fit of the model in the Albanian and Italian samples. More specifically, in the Albanian sample, MIs suggested adding various covariances between the error terms of item 2 (*Do you feel more irritability, anxiety, or even sadness when you try to either reduce or stop your gaming activity?*) and item 6 (*Have you continued your gaming activity despite knowing it was causing problems between you and other people?*), item 4 (*Do you systematically fail when trying to control or cease your gaming activity?*) and item 9 (*Have you jeopardized or lost an important relationship, job, or an educational or career opportunity because of your gaming activity?*), item 5 (*Have you lost interests in previous hobbies and other entertainment activities as a result of your engagement with the game?*) and item 7 (*Have you deceived any of your family members, therapists, or others because the amount of your gaming activity?*), item 6 and item 7, item 6, and item 8. In the British sample, MIs suggested adding covariances between the error terms of items 7 and 9. Finally, in the Italian sample, MIs suggested adding covariances between the error terms of items 1 and 2, 2 and 3, and 6 and 7. Therefore, in Albanian, British, and Italian samples, these items shared an amount of variance that was not captured by the construct.

With regard to the MI, although the $\chi^2$ difference test between the nested models was significant, the incremental fit indices values and the cutoff values of $\Delta$CFI and $\Delta$RMSEA provided support for configural invariance and partial support for the metric and scalar invariance, in line with the studies by Pontes et al. (2017) and Stavropoulos et al. (2017). The acceptable fit to the data of the configural invariance model demonstrated that the single-

**Table 3** Tests of invariance of the IGDS9-SF questionnaire between Albanian, the US, the UK, and Italian gamers

| | $\chi^2$ | df | p value | $\chi^2$ (DIFFTEST) | df | p value | RMSEA | ΔRMSEA | CFI | ΔCFI | TLI | WRMR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model 1: configural invariance | 265.752 | 99 | .001 | – | – | – | 0.069 | – | 0.998 | – | 0.998 | 1.453 |
| Model 2: metric invariance | 530.187 | 123 | .001 | 198.202 | 24 | .000 | 0.097 | 0.028 | 0.996 | 0.002 | 0.995 | 2.599 |
| Model 2a: partial metric invariance (item 1 for Albania; item 4 for the UK and Italy; item 7 for USA) | 318.737 | 119 | .001 | 70.13 | 20 | .000 | 0.069 | 0.000 | 0.998 | 0.000 | 0.998 | 1.93 |
| Model 2b: partial metric invariance (item 7 for the UK and item 2 for Italy) | 273.32 | 117 | .001 | 41.871 | 18 | .001 | 0.061 | 0.008 | 0.998 | 0.000 | 0.998 | 1.74 |
| Model 2c: partial metric invariance (item 1 for the UK and item 9 for Italy) | 251.339 | 113 | .000 | 27.074 | 16 | .041 | 0.058 | 0.011 | 0.999 | 0.001 | 0.998 | 1.65 |
| Model 3: scalar invariance | 1301.389 | 196 | .000 | 1011.883 | 81 | .000 | 0.126 | 0.068 | 0.989 | 0.010 | 0.992 | 4.39 |
| Model 3a: partial scalar invariance (item 3 threshold 1 for Albania, item 1 threshold 1 for the USA, the UK, and Italy) | 1104.918 | 192 | .000 | 859.883 | 77 | .000 | 0.116 | 0.058 | 0.991 | 0.008 | 0.993 | 3.949 |
| Model 3b: partial scalar invariance (item 7 threshold 1 for Albania; item 8 threshold 1 for the USA and Italy; item 1 threshold 2 for the UK) | 898.561 | 188 | .000 | 674.426 | 73 | .000 | 0.103 | 0.045 | 0.993 | 0.006 | 0.995 | 3.498 |
| Model 3c: partial scalar invariance (item 4 threshold 1 for Albania; item 1 threshold 2 for the USA; item 8 threshold 2 for the UK; item 2 threshold 1 for Italy) | 727.727 | 184 | .000 | 514.549 | 69 | .000 | 0.091 | 0.033 | 0.995 | 0.004 | 0.996 | 3.09 |
| Model 3d: partial scalar invariance (item 8 threshold 2 for the USA; item 5 threshold 1 for Italy) | 629.664 | 182 | .000 | 412.671 | 65 | .000 | 0.083 | 0.025 | 0.996 | 0.003 | 0.996 | 2.85 |
| Model 3e: partial scalar invariance (item 3 thresholds 1 and 3, item 8 threshold 4 for Italy) | 534.83 | 179 | .000 | 310.877 | 64 | .000 | 0.072 | 0.014 | 0.996 | 0.003 | 0.997 | 2.6 |

$\chi2$ chi square, $df$ degree of freedom, RMSEA (95% CI) root mean square error of approximation and its 95% confidence interval, CFI Comparative-of-Fit Index, TLI Tucker-Lewis Fit Index, WRMR weighted root mean square residual

factor structure of the IGDS9-SF was equivalent across the different countries compared, that is, the IGD construct can be assessed by the common underlying factor across Albania, Italy, the UK, and the USA. Conversely, the support for partial metric invariance suggested that the weights of the relationships between some items and the respective latent factor differed across the four countries. Particularly, the IGDS9-SF items 1, 2, 4, 7, and 9 referred, respectively, to preoccupation/salience, withdrawal symptoms, loss of control, deception, and negative consequences, showed unequal associations to the IGD latent factor across the countries. The IGD construct was differently associated with item 1 in Albania and the UK, with item 2 in Italy, with item 4 in the UK and Italy, with item 7 in the USA and the UK, and with item 9 in Italy. Finally, the support for partial scalar invariance revealed that for the same level of the latent IGD trait, participants across the cultures compared endorsed different response ratings in IGDS9-SF items assessing preoccupation (item 1), withdrawal symptoms (item 2), tolerance (item 3), loss of control (item 4), giving up other activities (item 5), deception (item 7), and escape (item 8).

The inequalities in factor loadings and intercepts may be clarified in terms of the cultural variations related to "individualism" and "collectivism" orientations (Hofstede 1980, 2001; Triandis 2001). In fact, the findings of the present research corroborated the assumptions that the universal dimensions of individual differences cannot be assessed equivalently across cultures and that differences between individualistic and collectivistic cultures may affect the response patterns and the measurement errors of psychometric instruments (Walford et al. 2010). Given that Albania is considered higher in collectivism, the items referred to the interpersonal and relationships difficulties associated with IGD may be answered differently. Individuals living in cultures characterized by a high conformity to cultural norms, values, and attitudes tend to regard family as important and to follow the norms and values dictated by their cultures. On the other hand, individuals with individualistic orientations, giving priority to personal goals would be more IGD vulnerable due to the tendency to focus more on game performance and to be less likely to seek professional psychological help (Raylu and Oei 2004).

Overall, the results of the present study provided further evidence for the underlying factor assessing IGD across the four countries, even though the specific meaning of the construct was non-identical. The cultural variations in the understanding, conceptualization, and assessment of IGD should be overcome, therefore, by using a more emic approach based not on the adaptation/translation of questionnaires, but on their creation taking into account the specific cultural perspectives of participants (Pontes et al. 2017).

The present study had a number of potential limitations that might narrow the conclusions which can be drawn from it, thus calling for more research. Some of these were more methodological in nature, others more conceptual. A potential methodological shortcoming of the study was that gender and age effects were not examined, making it difficult to establish the extent to which these variables might have confounded the results. Moreover, the comparison between convenience (i.e., non-representative) samples limited the generalizability of the results, also given to the reduced number of participants in Albania, the UK, and the USA. Additionally, although the individualist vs. collectivist distinction was theoretically highlighted, the two dimensions were not actually assessed. Consequently, the partial metric and scalar invariance was more theoretical than empirical. Item bias was another potential limitation that was fully scrutinized in the present study. Although the translations of the instrument into Albanian and Italian were each achieved with great rigor following the standard back-translation

procedures, the issues concerning item meaning and cultural appropriateness could have been influenced by the cultural styles themselves.

Despite these limitations, the present study provides further insight into the construct validity of the IGDS9-SF by providing a better understanding of its psychometric properties in a cross-cultural context that may help researchers and practitioners reach a common consensus concerning IGD diagnosis. Nevertheless, further cross-cultural research on IGD remains to be conducted because additional international comparisons would be of value in further examining how specific patterns of responses may differ across countries.

**Compliance with Ethical Standards**

**Conflict of Interest**  The authors declare that they have no conflict of interest.

**Ethical Approval**  All procedures performed in this study involving human participants were in accordance with the ethical standards of the University's Research Ethics Board and with the 1975 Helsinki declaration.

**Informed Consent**  Informed consent was obtained from all participants.

# References

Agaj, D. (2015). Predictive effects of self-esteem, shyness and loneliness on internet addiction, case of Albania. *International Journal of Education and Research, 3*(12), 129–134.

Agaj, D. (2016). The impact of the components of the locus of control in internet addiction, case of Albania. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 18*(1), 40–44.

Agaj, D., & Marku, A. (2015). The phenomenon of Internet addiction in adolescents: case of Albania. *Academic Journal of Interdisciplinary Studies, 4*(3 S1), 122. https://doi.org/10.5901/ajis.2015.v4n3s1p122.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.).* Arlington, VA: Author.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246. https://doi.org/10.1037//0033-2909.107.2.238.

Bentler, P. M., & Bonnet, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–606. https://doi.org/10.1037//0033-2909.88.3.588.

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

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255. https://doi.org/10.1207/s15328007sem0902_5.

Fontaine, J. R. J. (2005). Equivalence. In K.Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Vol. 1) (pp. 803–813). San Diego, CA: Academic Press.

Global Games Market Report (2017). Retrieved November 06, 2017, from:https://newzoo.com/insights/rankings/top-100-countries-by-game-revenues/

Griffiths, M. D., King, D. L., & Demetrovics, Z. (2014). DSM-5 Internet Gaming Disorder needs a unified approach to assessment. *Neuropsychiatry, 4*(1), 1–4. https://doi.org/10.2217/npy.13.82.

Hasmujaj, E. (2016). Internet addiction and loneliness among students of University of Shkodra. *European Scientific Journal, 12*(29), 397–407. https://doi.org/10.19044/esj.2016.v12n29p397.

HofstedeG. (1980). *Culture's consequences*. Beverly Hills, CA: Sage.

Hofstede, G. (2001). *Culture's consequences: comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc..

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: software of the mind*. New York: McGraw-Hill.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53–60.

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. https://doi.org/10.1080/10705519909540118.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*, 409–426. https://doi.org/10.1007/bf02291366.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). Newbury Park, CA: Sage.

Kim, H. Y. (2013). Statistical notes for clinical researchers: assessing normal distribution using skewness and kurtosis. *Restorative Dentistry & Endodontics, 38*(1), 52–54. https://doi.org/10.5395/rde.2013.38.1.52.

King, D. L., Haagsma, M. C., Delfabbro, P. H., Gradisar, M., & Griffiths, M. D. (2013). Toward a consensus definition of pathological video-gaming: a systematic review of psychometric assessment tools. *Clinical Psychology Review, 33*(3), 331–342. https://doi.org/10.1016/j.cpr.2013.01.002.

Ko, C. H. (2014). Internet gaming disorder. *Current Addiction Reports, 1*(3), 177–185. https://doi.org/10.1007/s40429-014-0030-y.

Kühne, R. (2013). Testing measurement invariance in media psychological research. *Journal of Media Psychology, 25*, 153–159. https://doi.org/10.1027/1864-1105/a000096.

Kuss, D. J. (2013). Internet gaming addiction: current perspectives. *Psychology Research and Behavior Management, 6*, 125. https://doi.org/10.2147/prbm.s39476.

Kuss, D. J., Griffiths, M. D., & Pontes, H. M. (2017). Chaos and confusion in DSM-5 diagnosis of Internet Gaming Disorder: issues, concerns, and recommendations for clarity in the field. *Journal of Behavioral Addictions, 6*(2), 103–109. https://doi.org/10.1556/2006.5.2016.062.

Lin, X., Jia, X., Zang, Y. F., & Dong, G. (2015). Frequency-dependent changes in the amplitude of low-frequency fluctuations in internet gaming disorder. *Frontiers in Psychology, 6*(1471). https://doi.org/10.3389/fpsyg.2015.01471.

Melonashi, E. (2017). Social anxiety and self-reported time spent online in a sample of Albanian university students. *CBU International Conference Proceedings, 5*, 717–721. https://doi.org/10.12955/cbup.v5.1013.

Merenda, P. F. (2006). An overview of adapting educational and psychological assessment instruments: past and present. *Psychological Reports, 99*, 307–314. https://doi.org/10.2466/pr0.99.6.307-314.

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: applications in crosscultural research. *International Journal of Psychological Research, 3*(1), 111–121. https://doi.org/10.21500/20112084.857.

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*(3), 479–515. https://doi.org/10.1207/s15327906mbr3903_4.

Monacis, L., de Palo, V. D., 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. https://doi.org/10.1556/2006.5.2016.083.

Monacis, L., de Palo, V., Griffiths, M. D., & Sinatra, M. (2017). Exploring individual differences in online addictions: the role of identity and attachment. *International Journal of Mental Health and Addiction, 15*(4), 853–868. https://doi.org/10.1007/s11469-017-9768-5.

Muthén, L. K., &Muthén, B. O.(1998–2012). Mplus user's guide (7th ed.). Los Angeles, CA: Muthén & Muthén.

Petry, N. M., & O'brien, C. P. (2013). Internet gaming disorder and the DSM-5. *Addiction, 108*(7), 1186–1187. https://doi.org/10.1111/add.12162.

Pontes, H. M., & Griffiths, M. D. (2014). Assessment of Internet gaming disorder in clinical research: past and present perspectives. *Clinical Research and Regulatory Affairs, 31*(2–4), 35–48. https://doi.org/10.3109/10601333.2014.962748.

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. https://doi.org/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 & Social Networking, 19*(4), 288–293. https://doi.org/10.1089/cyber.2015.0605.

Pontes, H. M., Macur, M., & Griffiths, M. D. (2016). Internet gaming disorder among Slovenian primary schoolchildren: findings from a nationally representative sample of adolescents. *Journal of Behavioral Addictions, 5*(2), 304–310. https://doi.org/10.1556/2006.5.2016.042.

Pontes, H. M., Stavropoulos, V., & Griffiths, M. D. (2017). Measurement invariance of the internet gaming disorder scale–short-form (IGDS9-SF) between the United States of America, India and the United Kingdom. *Psychiatry Research, 257*, 472–478. https://doi.org/10.1016/j.psychres.2017.08.013.

Raylu, N., & Oei, T. P. (2004). Role of culture in gambling and problem gambling. *Clinical Psychology Review, 23*(8), 1087–1114. https://doi.org/10.1016/j.cpr.2003.09.005.

Stavropoulos, V., Beard, C., Griffiths, M. D., Buleigh, T., Gomez, R., & Pontes, H. M. (2017). Measurement invariance of the internet gaming disorder scale–short-form (IGDS9-SF) between Australia, the USA, and the UK. *International Journal of Mental Health and Addiction*, 1–16. https://doi.org/10.1007/s11469-017-9786-3.

Tejeiro Salguero, R. A., & Morán, R. M. B. (2002). Measuring problem video game playing in adolescents. *Addiction, 97*(12), 1601–1606. https://doi.org/10.1046/j.1360-0443.2002.00218.x.

Triandis, H. C. (2001). Individualism-collectivism and personality. *Journal of Personality, 69*(6), 907–924. https://doi.org/10.1111/1467-6494.696169.

Walford, G., Tucker, E., & Viswanathan, M. (2010). *The SAGE handbook of measurement*. Sage Publications.

Wu, T. Y., Lin, C. Y., Årestedt, K., Griffiths, M. D., Broström, A., & Pakpour, A. H. (2017). 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? *Journal of Behavioral Addictions, 6*(2), 256–263. https://doi.org/10.1556/2006.6.2017.025.