ESTA ES LA VERSIÓN FINAL ACEPTADA POR LA REVISTA PARA SU PUBLICACIÓN UNA VEZ SUPERADO EL PROCESO DE REVISIÓN

Spanish Validation and Scoring of the Internet Gaming Disorder Scale - Short-Form (IGDS9-SF)

Abstract

Since the inclusion of the Internet Gaming Disorder (IGD) in the DSM-5, the Internet Gaming Disorder Scale-Short Form (IGDS9-SF), a short nine items test, has become one of the most used standardized instruments for its psychometric evaluation. This study presents a validation and psychometric evaluation of the Spanish version of the IGDS9-SF. A sample of 2173 videogame players between 12 and 22 years old, comprising both genders, was employed, achieved with a randomized selection process from educational institutions in the city of Madrid. Participants completed the adapted version of the IGDS9-SF, the GHQ-12 and a negative cognitions scale associated with videogame use, as well as sociodemographic data and frequency of videogame play.

A unifactorial structure with sufficient reliability and internal consistency was found through exploratory and confirmatory analyses. In addition, the instrument was found to have good construct validity; the scoring of the IGDS9-SF were found to show a positive association with gaming frequency, with general health problems, and to a greater extent, with problematic cognitions with regard to videogames. Factorial invariance was found concerning the age of participants. However, even though the factorial structure was consistent across genders, neither metric nor scalar invariance were found; for this reason, we present a scale for the whole sample and a different one for gender. These results suggest that

this Spanish version of the IGDS9-SF is a reliable and valid instrument, useful to evaluate the severity of IGD in Spanish students, and we provide a scoring scale for measurement purposes.

Key words: Internet Gaming Disorder, IGDS9-SF, psychometric properties, invariance.

Spanish Validation and Scoring of the Internet Gaming Disorder Scale - Short-Form (IGDS9-SF)

The use of videogames (VG) in young adults and adolescents is an important cause of social alarm, since it has been more frequently associated with negative consequences such as sleep problems (Lam, 2014), alterations in well-being (Scott and Porter-Armstrong, 2013), increase in mental health problems or lesser degree of self-control (Dinh, Yasuoka, Poudel, Otsuka and Jimba, 2013). Other identified problems include deterioration of the family environment, social isolation, behavioral problems and a generally disorganized life. However, playing VG is a habitual leisure activity, especially amongst young adults and adolescents, and just a low percentage of players develop a problem that requires professional help. Therefore, it seems appropriate to dispose of criteria to discriminate between normal, excessive and problematic videogame gaming (VGG).

Until the inclusion of the Internet Gaming Disorder (IGD) in the DSM-5 (APA, 2013), there was a lack of agreement between researchers regarding the criteria to identify when playing VG is problematic and frequently the identification was based on the criteria for pathological gambling or substance abuse (Kuss and Griffiths, 2012). There also was a lack of agreement on the name to be given to this problem, changing between excessive VGG, addiction to VG or problematic VGG.

The DSM calls the problem IGD and establishes nine criteria for its diagnosis, of which five have to be met for a period of at least 12 months, although these do not determine the level of severity. These nine criteria are: (1) preoccupation about internet gaming; (2) withdrawal symptoms; (3) tolerance; (4) unsuccessful attempts to control participation in internet gaming; (5) loss of interest in previous hobbies and entertainment activities as a result of, and with the exception of, videogames; (6) excessive and continued use of internet games regardless of the acknowledged psychosocial problems; (7) deceived family members, therapists, or other people regarding the amount of time spent playing games on the Internet; (8) use of videogames to escape or relieve negative moods and, (9) jeopardised or lost significant interpersonal relationships, jobs and educational or professional opportunities because of participations in Internet gaming. The contribution of the DSM-5 constitutes an advance, although it is not without controversy (Király et al., 2017; Kuss, Griffiths and Pontes, 2017; Van Rooij and Kardefelt-Winther, 2017).

It is difficult to determine the presence of IGD because of the limitations, or lack of precise criteria in the diagnosis of IGD (or alternative names), resulting in a widespread prevalence number, indicating as reference an estimate between 0.7% and 15.6%, shown in a meta-analysis by Feng, Ramo, Chan and Bourgeois, (2017). This broad range can account for, besides the lack of criteria or their suitability, the description of the problem, the disparity in the instruments available for evaluation and their limitations.

Previous to the publication of the DSM-5, Kuss and Griffiths (2012), and King, Haagsma, Delfabbro, Gradisar and Griffiths (2013), analyzed 18 measurement instruments of IGD (or alternative denominations) and their application, pointing out, between the strengths found, a high internal consistency (between .70 and .96), and a good convergent validity with related measures. Specifically, seven of these instruments found a positive correlation

between severity, or number of symptoms, and time dedicated to VGG. However, the samples used were rather small, biased (very specific of a certain age group or mixed without distinction) and/or of convenience, generally obtained with online recruitment. Kuss and Griffiths (2012) concluded that it was difficult to assess the validity of these tools in discriminating the level of addiction to gaming, recommending the development of manuals with standardized guidelines. At a later time, Griffiths (2017) pointed out that the assessment instruments for IGD were generally inconsistent.

In a more recent revision of the assessment tools for IGD in adolescents and young adults, Bernaldo-de-Quirós, Labrador-Méndez, Sánchez-Iglesias and Labrador (2019), highlighted the significant changes made since the publication of the DSM-5 criteria. The study showed the variety of the instruments is reduced, now focusing on evaluating IGD (compared to the assessment of addiction to the Internet in general), they are often shorter (between 9 and 27 items) and they use an answer format of a 5 point Likert scale type. Between the tools analyzed, the value of the IGDS9-SF stands out (Pontes and Griffiths, 2015).

On the one hand, this could be due to the observed applicability and adjustment to the DSM-5 criteria. The IGDS9-SF features 9 items which account for the nine diagnostic criteria of the DSM-5. Its objective is to assess the severity of IGD and its harmful effects, evaluating the activities of VGG carried out on the Internet and outside of it, during the past 12 months. Furthermore, although the main objective of the instrument is not to diagnose IGD, a cutoff point is established to differentiate between players with or without the disorder. The validation of the original instrument (Pontes and Griffiths, 2015) was carried out with a sample of 1060 players, English speakers, between 16 and 60 years old (*M =27 years old, SD = 9.02*). The evaluation tool showed an internal consistency of .87, a single factor and good

construct validity, with significant correlations with IGD-20 (Pontes, Király, Demetrovics and Griffiths, 2014), and the time dedicated to VGG per week.

On the other hand, the IGDS9-SF is also the most translated and validated instrument thus far: Portuguese (Pontes and Griffiths, 2016); Slovenian (Pontes, Macur and Griffiths, 2016); Italian (Monacis, De Palo, Griffiths and Sinatra, 2017); Persian (Wu et al., 2017); Turkish (Evren et al., 2018); Polish (Schivinski, Brzozowska-Woś, Buchanan, Griffiths and Pontes, 2018); and Spanish (Beranuy et al., 2020). All of these validations studies agree in identifying one single underlying factor, with internal consistency between .82 and .99 and good criterion validity, with significant correlations with time spent VGG. Additionally, a relationship was found between scores on the IGDS9-SF and depression, anxiety and stress (Pontes and Griffths, 2016; Wu el al., 2017), mental health (Pontes, Macur, and Griffiths, 2016), life satisfaction (Pontes et al., 2016), gambling disorders, and quality of life (Beranuy et al., 2020) measures. In the same way, the instrument showed good convergent validity using other tools measuring addiction to VG (Evren et al., 2018; Monacis et al., 2017; Pontes and Griffiths, 2015) or addiction to the Internet in general (Evren et al., 2018; Monacis et al., 2017).

The scoring of the IGDS9-SF ranges from 0 to 45 points, marking the cutoff point to consider a player with IGD, either a score of 35, or answering to five or more items with the maximum value “very often” (Pontes and Griffths, 2016; Pontes et al. 2016). Alternatively, the Italian version sets the cutoff point to identify players with IGD to 21, according to the gold standard of the Game Addiction Scales (Lemmens, Valkenburg and Peter, 2009), showing sensitivity of 86.1% and specificity of 86%.

However, in the application of the IGDS9-SF the problems related to sample selection persist: the use of random sampling is scarce (Pontes et al., 2016; Wu et al., 2017), in the

majority of cases convenience sampling is used, with *links* in *gaming* forums (Evren et al., 2018; Pontes and Griffiths, 2015; Schivinski et al., 2018) or recluting only in schools available to the researchers (Beranuy et al., 2020; Pontes and Griffiths, 2016; Monacis et al., 2017). Only half of these studies use a sample of exclusively adolescents and young adults (Beranuy et al., 2020; Pontes and Griffiths, 2016; Pontes et al., 2016; Wu et al., 2017), and in some cases they only briefly mention the number of participants of different age groups or any differentiation carried out (Monacis et al., 2017; Schivinski et al., 2018), and in one case we can see that age groups are not even discerned (Evren et al., 2018).

Due to the present circumstances, it seems of great interest to validate the IGDS9-SF in Spain, in a population of adolescents and young adults, using a random sample of educational institutions, in order to provide an instrument considered more appropriate for the current time, to assess the presence of problems with VGG. The objectives of this study were: 1) to assess the psychometric properties of the Spanish version of the IGDS9-SF and criterion validity, by analyzing the relationships of their scores with other theoretically related variables, 2) to evaluate reliability with internal consistency and split-half method; construct validity using factor analysis and the measurement invariance across age and gender groups. In addition, we discussed the use of cutoff scores and, arguing that currently there is no adequate gold standard to base those scores on, presented a scoring scale instead.

Method

Participants

The initial sample comprised 2887 participants of both genders, extracted from a representative sample of educational institutions from the city of Madrid. We considered two age groups, 12 to 16 (a total of 2020 students) and 17 to 22 years old (867 students). Of these, 75.0% and 75.8% played videogames, respectively. No relationship was found between age

group and playing videogames, $X^2(1) = 0.173$, $p = .677$. The final sample consisted of 2173 participants (28.8% females). Table 1 shows age and gender distributions for both samples.

Table 1

Age and Gender Distributions for Initial and Final Samples

	Initial sample		Final sample (videogame players)		
Age group (years)	*N*	% players	*N*	N_{female} (%)	M_{age} (SD_{age})
12 to 16	2020	75.0%	1516	491 (32.4%)	13.82 (1.34)
17 to 22	867	75.8%	657	135 (20.5%)	18.82 (1.52)
Total	2887	75.3%	2173	626 (28.8%)	15.33 (2.69)

Instruments and Measures

IGDS9-SF. The Spanish version of the scale (see Appendix A) is an adapted translation of the original IGDS9-SF (Pontes and Griffiths, 2015); it was translated by two independent people and back-translated into Spanish by two native English speakers. An additional researcher overviewed the final version. This single-factor scale, based on the DSM-5 (APA, 2013) criteria for IGD, consists of nine 5-point Likert (1 "Never" to 5 "Very Often") items. The total score is the sum of the items scores; the higher the score, the higher the severity of IGD.

General Health Questionnaire (GHQ-12: Goldberg and Williams, 1988). The 12 items establish a total score for perceived health (α = .569), with an anxiety subscale (7 items, α = .776) and a further one for dysfunction (5 items, α = .771). The correction used was

CGHQ, which aims to avoid underdetection of participants with long term problems. Greater scores indicate poorer health.

Gaming Cognitions Scale (Fernández-Árias et al., 2019). This scale showed an adequate reliability R_{xx} = .931 and internal consistency, α = .907. It comprised sixteen 5-point Likert scale items (1 "Never" to 5 "Always"), ranging from 0 to 80 points (a greater score indicates more problematic cognitions towards videogames). Three factors account for 50.27% of the variance of the scores: Game immersion (α = .829), craving (α = .843), and refusal to stop playing (α = .799).

Frequency of gameplay. In two separate questions, participants reported their average days per week invested in playing VG, and average hours per week (from "less than 1 hour" up to "more than 30 hours" playing VG, increasing in order of five hours (1 to 5 hours, 6 to 10, until reaching 30, with the further option of more than 30 hours).

Procedure

Five independent evaluators with psychology degrees were trained to administer the *Gamertest,* an online assessment tool which includes the instruments used in this study (http://www.famgi14.es/gamertest/index.html). Data on the student population for the 21 districts of the city of Madrid, including their ages, school year and type of schooling (public school, private school and state subsidized school), was retrieved from the website of the city hall statistics service (Ayuntamiento de Madrid, 2017). Schools were divided into groups by district and type of school, and randomly ordered. Then, for each district and type of school, the first school on the list was contacted through a detailed letter, with a follow-up call soon thereafter, and asked to provide access to the set of classes required by the district. If the school refused, the next school on the list was contacted. Once a school agreed to participate in the study, the evaluators delivered informed consent forms for the children's

parents/guardians and a date was set for the evaluator to visit the school to perform the assessment in a classroom chosen using stratified random sampling. After collecting informed consent forms from the parents/guardians, the assessments were administered in groups, using computers in each school's computer room, allowing approximately 30-40 minutes for the students to complete them. Participants' responses were anonymously collected and coded directly in a computerized database.

Ethical issues for this study were audited by the ethics committee of the [name of the universitiy] Faculty of Psychology.

We used several R packages: The CFAs were carried out using lavaan, version 0.6-3 (Rosseel, 2012); estimators α and ω and their CI were computed using MBESS, version 4.4.3 (Kelley, 2018), and the invariance analysis was carried out with semTools (Jorgensen, Pornprasertmanit, Schoemann, and Rosseel, 2018). The rest of the statistical analyses were carried out using SPSS 20, including macros for parallel analysis (O'Connor, 2000) and Mardia's multivariate analysis (DeCarlo, 1997).

Factorial validity. The final sample of participants, from 12 to 16 years old, was divided randomly into two subsamples of videogame players, to assess the validity of the Spanish version of the IGDS9-SF. One of the subsamples (n = 758) was used to carry out an exploratory factor analysis (EFA), after assessing the adequacy of the EFA in our dataset, via Bartlett's sphericity test and KMO estimate. In this subsample, we assessed multivariate normality of the items via Mardia's multivariate kurtosis and skewness coefficients (Mardia, 1970); as the hypothesis of normality was rejected, an unweighted least squares (ULS) extraction method was selected. Three criteria were used to determine the number of components to retain from EFA: the K1 method, the inspection of the scree plot and a

principal components parallel analysis (Horn, 1965). As only one factor was retained, no rotation method was used.

The second subsample (n = 758) was used to confirm the subjacent structure through confirmatory factor analysis (CFA). To address non-normality issues, a maximum likelihood with robust standard errors estimation method (MLR) and Satorra-Bentler statistic were used. Several fit indexes were employed (and compared with the values recommended by Chau, 1997; and Schreiber, Nora, Stage, Barlow, and King, 2006): chi square statistic to degrees of freedom ratio (χ^2/df), root mean square error of approximation (RMSEA) and its 90% confidence interval, standard root mean square residual (SRMR), comparative fit index (CFI) and Tucker-Lewis index (TLI). The magnitude, direction and statistical significance of the standardized parameter estimates were interpreted.

The third subsample (videogame players from 17 to 22 years old, n = 657) was also used in a CFA (using the same specifications as the first one), to extend the assessment of factor validity to that age group. To further examine the generalizability of the scale across younger and older participants in the whole sample, we tested for group measurement invariance using multiple group confirmatory factor analysis (MGCFA). Several nested multiple group models were tested for (a) configural invariance, (b) weak or metric invariance, and (c) strong or scalar invariance, in order to assess the invariance of number of factors, factor loadings and indicator intercepts, respectively. The χ^2 difference ($\Delta\chi^2$) test, CFI difference (ΔCFI), and RMSEA difference (ΔRMSEA) were used as fit indices when testing for metric and scalar invariance. A value of ΔCFI < −.010 supplemented by a ΔRMSEA < .015 would indicate invariance (Chen, 2007).

We also examined the generalizability of the scale across genders (regardless of age), using MGCFA, with the same criteria as for age MGCFA.

We calculated descriptive statistics for the items and the total score of the Spanish version of the IGDS9-SF for the three subsamples. Two estimators (with 95% CI) for internal consistency were used, as well as Cronbach's alpha (α), and omega (ω) to overcome some problems associated with the former (McDonald, 1999). Also, we assessed split-half reliability using the Spearman-Brown formula (arranging the items of the two halves by mean).

ANOVA and correlation tests were used to study the relationship between IGDS9-SF scores and weekly frequency and average hours spent playing VG, respectively.

Lastly, we correlated the IGDS9-SF scores with the GHQ-12 scores and with the Gaming Cognitions Scale scores.

Results

IGDS9-SF Psychometric Properties

Item analysis and reliability. Table 2 displays the mean and standard deviation values of the items as well as the total scale scores. Individually considered, all scores showed positive skewness, but different degrees of kurtosis. When analyzing Subsample 1, used for EFA, the estimates of Mardia's multivariate kurtosis and skewness coefficients were high and significant, 34.69 and 166.36 respectively (both $ps < .001$). The internal consistency (α and ω) and reliability estimators were adequate for the three subsamples (see Table 2).

Table 2

IGDS9-SF. Descriptive Statistics, Internal Consistency and Reliability, by Subsample

	Subsample 1		Subsample 2		Subsample 3	
Item	*M*	*SD*	*M*	*SD*	*M*	*SD*
1	2.45	1.25	2.45	1.29	2.31	1.30
2	1.78	1.09	1.73	1.04	1.54	0.91
3	2.09	1.28	2.03	1.26	1.81	1.10
4	2.03	1.26	1.98	1.16	1.77	1.18
5	1.64	1.13	1.55	1.02	1.60	1.01
6	1.60	1.14	1.64	1.11	1.51	1.07
7	1.64	1.12	1.49	0.87	1.55	1.03
8	2.38	1.51	2.28	1.43	2.23	1.36
9	1.25	0.77	1.21	0.64	1.23	0.69
Total score	16.87	6.89	16.35	5.93	15.55	5.90
α [95% CI]	.825 [.795, .852]		.768 [.734, .791]		.780 [.736, .815]	
ω [95% CI]	.828 [.796, .856]		.778 [.746, .801]		.787 [.745, .821]	
R_{xx}	.822		.770		.790	

Notes: $N = 2173$. Subsample 1 (12 to 16 years old), $n = 758$. Subsample 2 (12 to 16 years old), $n = 758$. Subsample 3 (17 to 22 years old), $n = 657$.

EFA. Both the KMO test, .90, and Bartlett's test of sphericity, $\chi^2(36) = 1751.84$, $p < .001$, showed the data was adequate for EFA. Table 3 shows the communalities of the items after the extraction of a single factor (see Figure 1 for the scree plot and parallel analysis results), as well as its factor loadings, which ranged from .503 to .685. The single factor accounted for 35.78% of the items' total variance.

Table 3

IGDS9-SF. Results from Exploratory Factor Analysis

Item	Communality	Factor loading
1	.405	.637
2	.429	.655
3	.468	.685
4	.341	.585
5	.391	.625
6	.255	.505
7	.384	.620
8	.253	.503
9	.292	.540

Notes: $N = 758$. ULS extraction method. Explained variance 35.78%.

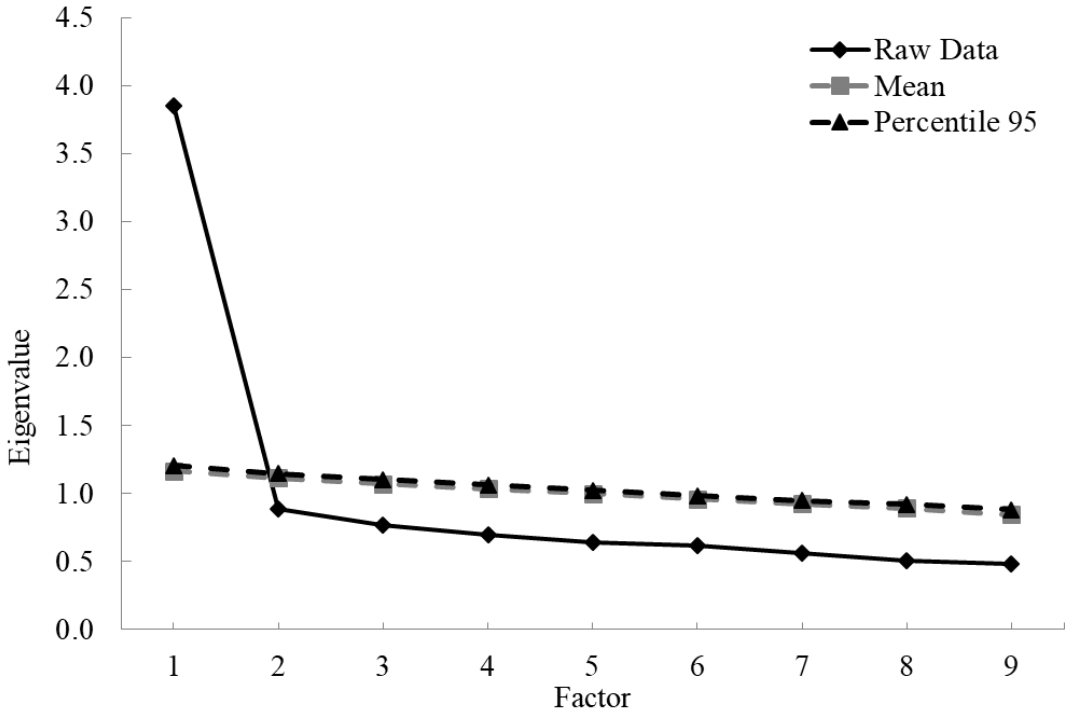


Figure 1. EFA scree plot for the Spanish version of the IGDS9-SF. The plot displays empirical data eigenvalues, mean and 95th percentile eigenvalues of 100 random samples in a parallel analysis. $N = 758$.

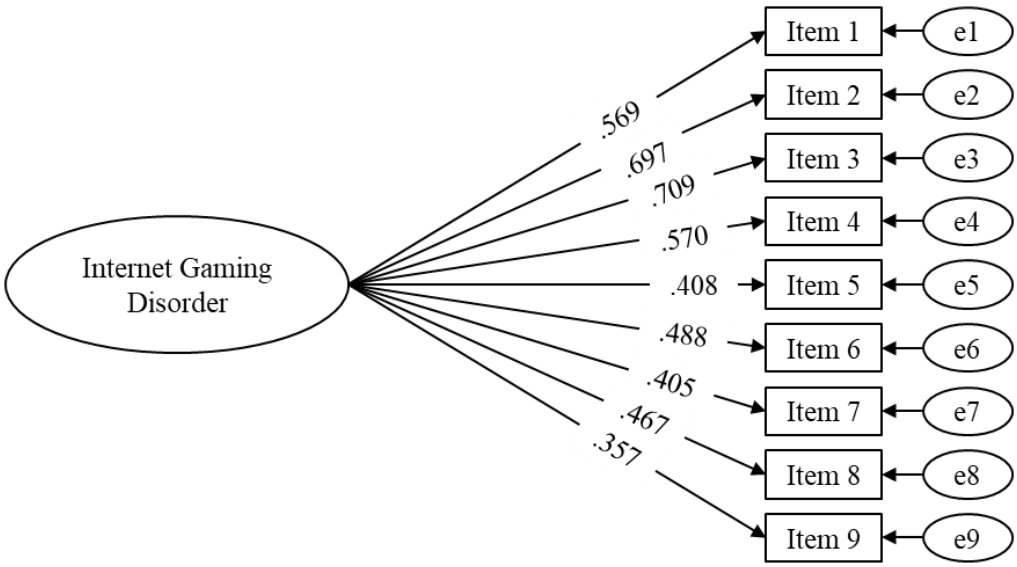
CFA. The fit indices for both CFA models can be found in Table 4. When considering Subsample 2, the CFA showed adequate fit based on all the model fit indices. In Subsample 3, the observed indices did not reach the threshold values on CFI and TLI, but they did on χ^2/df, RMSEA and SRMR, so the model should not be rejected. Moreover, both models showed positive, significant ($p < .001$) item factor loadings, ranging from .357 to .697 for Subsample 2, and from .342 to .732 for Subsample 3 (see Figure 2), reaching the values recommended by Brown (2105).

Table 4

IGDS9-SF. Fit indices for CFA Models

		Subsample 2 (CFA, *n* = 758)	Subsample 3 (CFA, *n* = 657)	RV
Fit index	χ^2/df	2.096	2.744	≤ 3.000
	RMSEA	.038	.052	< .060 to .080
	RMSEA [90% CI]	[.027, .049]	[.040, .063]	< .060 to .080
	SRMR	.038	.044	≤ .080
	CFI	.961	.940	≥ .950
	TLI	.948	.920	≥ .950

Note. RV: Recommended values (Chau, 1997; Schreiber et al., 2006).



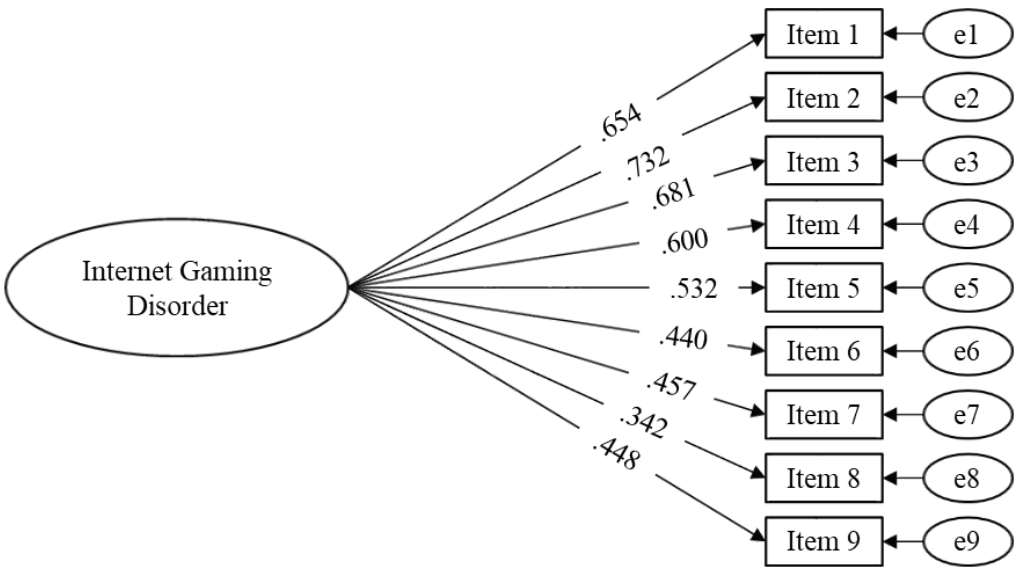


Figure 2. CFA models for Subsample 2 (above, n = 758, from 12 to 16 years old) and Subsample 3 (below, n = 657, from 17 to 22 years old). Standardized coefficients. All factor loadings are statistically significant (ps < .001).

Measurement invariance. The MGCFA yielded a good fit across age group subsamples; the single factor model showed invariance in factor structure (configural invariance), when constraining factor loadings (metric invariance) and when constraining factor loadings and indicator intercepts (scalar invariance). Only the $\Delta\chi^2$ suggested non-invariance in the scalar invariance model, but this index is known for its sensitivity to sample size, and nonetheless ΔCFI and ΔRMSEA showed good fit (Table 5). Therefore, it can be concluded that the model is valid for both age groups.

When considering gender invariance (see Table 5), the MGCFA showed good configurable fit across genders (i.e. same factor structure for both groups). However, the model did not yield a good fit neither for the metric invariance model (the construct does not have the same meaning across genders) nor for the scalar invariance model (the intercepts are not equal across genders), attending to $\Delta\chi^2$ and ΔCFI, although ΔRMSEA was adequate.

Table 5

IGDS9-SF. Measurement Invariance by Age and Gender Groups

Invariance model	χ^2	*Df*	$\Delta\chi^2$	*p* ($\Delta\chi^2$)	CFI	ΔCFI	RMSEA	ΔRMSEA
Age								
Configural	260.09	54	-	-	.958	-	.055	-
Weak/metric	280.59	62	11.58	.171	.956	-.002	.052	.003
Strong/scalar	313.42	70	34.79	< .001	.950	-.006	.052	.000
Invariance model	χ^2	*Df*	$\Delta\chi^2$	*p* ($\Delta\chi^2$)	CFI	ΔCFI	RMSEA	ΔRMSEA
Gender								
Configural	247.02	54	-	-	.959	-	.051	-
Weak/metric	325.29	62	37.28	< .001	.943	-.016	.057	.005

Strong/scalar	399.21	70	65.95	< .001	.926	-.017	.061	.004

Notes: Robust estimators for CFI and RMSEA. All deltas are compared to the previous model.

Frequency of playing VG. There was a significant linear relationship between weekly frequency of playing and IGDS9-SF scores, $F(1, 2165) = 421.82$, $p < .001$, $\eta^2 = .160$. A higher score was associated with more hours of playing VG per week.

Participants played VG during an average of 3.41 days per week (*Mdn* = 3.00, *SD* = 2.01). We found a positive relationship between days of gaming per week and IGDS9-SF scores, $r = .395$, $r^2 = .156$.

GHQ-12 and Cognition scale. The IGDS9-SF scores positively correlated with GHQ-12 scores (with a shared variance, r^2, ranging from .007 to .040) and with the Cognition Scale scores (r^2 from .286 to .433), see Table 6.

Table 6

IGDS9-SF. Convergent Validity with GHQ-12 and with the Cognition Scale

Scale	Factor	*r*
GHQ-12	Total	.182
	Anxiety	.200
	Disfunction	.082
Cognition	Game immersion	.535
	Craving	.658
	Refusal to stop playing	.658

Note: All correlations are significant, $p < .001$.

Scoring and Cutoff Values

The IGDS9-SF scores ranged from 9 to 45 (*M* = 16.29, *SD* = 6.29, *Mdn* = 15.00, *IQR* = 7.00). A significant difference was found between male (*M* = 17.41, *SD* = 6.36, *Mdn* =

16.00, *IQR* = 7.00) and female (*M* = 13.52, *SD* = 5.17, *Mdn* = 12.00, *IQR* = 5.00) participants, $t(1411.15) = 14.82$, $p < .001$, accounting for 7.8% of the scores variance, $r^2 = .078$. Because of this difference, and because the MGCFA showed metric and scalar non-invariance across gender groups, two separate scoring scales were computed for the IDGS9-SF (Table 7).

Table 7

IGDS9-SF. Scoring Scale for Total Sample and by Gender

	Score		
Percentile	Total (*N* = 2173)	Male (*n* = 1547)	Female (*n* = 626)
1	-	9	-
5	9	10	-
10	-	-	-
15	10	11	-
20	11	12	9
25	-	-	-
30	12	13	10
35	-	14	-
40	13	-	-
45	14	15	11
50	15	16	-
55	-	-	12
60	16	17	13
75	19	20	15
80	21	22	17
85	22	24	18
90	25	26	20
91	-	-	21
92	26	27	22
93	27	28	23
94	28	29	24
95	29	30	25
96	30	32	-
97	32	34	26
98	34	35	27
99	37	37	32

Discussion

This study shows evidence for the validity -and reliability- of the Spanish version of the IGDS9-SF, putting forward the use of this short, self-report tool to assess Internet Gaming Disorder (IGD). We found sufficient psychometric evidence of internal consistency, reliability and factor validity for the Spanish version of the IDGS9-SF, when applied to a sample of 2173 videogame players, composed of children, adolescents and young adults (from 12 to 22 years old) of both genders, recruited from a representative sample of educational centers from the city of Madrid.

We came across evidence of a single-factor model that matched the original, English scale structure (Pontes and Griffiths, 2015). This model was found in an EFA, and was confirmed in two distinct CFAs. The measurement invariance showed validity of the scale across age groups, therefore suggesting its use could be appropriate with participants within the age range between 12 and 22 years old. However, in this Spanish version of the scale, and unlike Monacis et al. (2017) or Wu et al. (2017), only configural gender invariance was found; while the single factor structure was the same for male and female participants, the factor loadings and factor intercept were different for each group. Due to these results, we put forward the use of separate scoring scales according to gender.

As part of the validation process, this study investigated the convergent and criterion validity of the Spanish IGDS9-SF with related measurements (GHQ and cognition scale) and other relevant variables such as gaming frequency. Additionally, a greater score in the IGDS9-SF was found to be associated with more gaming days per week and with a greater number of hours of gaming per week (in both cases sharing 16% of variance). This relationship with gaming time agrees with the one found in the validation of the original tool

(Pontes and Griffiths, 2015), as well as with the assessments carried out in other countries in which this variable was examined (Pontes and Griffiths, 2016; Pontes et al., 2016; Schivinski et al., 2018; Wu et al., 2017). Furthermore, a relationship was found between high scores in the IGDS9-SF and poorer health, assessed with the GHQ-12 (sharing between 1% and 4% of variance, according to the factor examined), in agreement with the Portuguese, Slovene and Turkish validations (Pontes and Griffiths, 2016; Pontes et al., 2016; Wu et al., 2017). Ultimately, a stronger relation was found (between 27% and 43% of the variance explained) between questionnaire scores and the problematic cognition scale regarding videogames. Despite the importance of the association between these cognitive distortions and the problematic use of videogames (King and Delfabro, 2016; Moudiab and Spada, 2019), studies assessing the instrument had not included this type of assessment thus far. These results suggest good criterion validity for the adapted version of the IGDS9-SF.

The sample sizes employed in all the analyses (657, 758 and 758 participants for the three subsamples) can be considered sufficient, far exceeding the recommended sample size for factor analysis, even in the presence of low communalities (Lloret-Segura, Ferreres-Traver, Hernández-Baeza and Tomás-Marco, 2014).

Griffith and Pontes (2015) suggested a cutoff value of 36 out of 45 to classify disordered gamers. In our sample, only 28 participants (1.30% of the gamers, and 0.96% of the total sample) scored a minimum of 36 out of 45. Other authors have suggested different cutoff scores based on gold standards (Monacis et al., 2017). As the cutoff scores have –currently- no empirical support, we also followed other authors who support a more strict, clinical approach, with the endorsement of at least five of the criteria of IGD, by answering "very often" (the highest score) in at least five items of the IGDS9-SF (Evren et al., 2018; Pontes et al., 2016; Pontes and Griffiths, 2016; Schivinski et al., 2018). Using this criterion,

65 participants (3.00% of the gamers, 2.25% of the total sample) would be labeled as having an IGD (similar to the 1.90% and 1.00%, respectively, found by Beranuy et al., 2020). Furthermore, instead of relying on a single cutoff score without adequate empirical support, we provide a scoring scale, for the total sample and segregated by gender, for measurement purposes and norm referencing (Table 7). This scale allows comparing the score of a student to a norm group. Describing performance relative to others, even for non-disordered gamers, is a useful tool for measuring the construct.

The present study only examined a sample of Spanish students aged 12 to 22 years old, therefore the results cannot be generalized to the wider population. Also, the potential biases of self-reported data (such as social desirability and memory recall) are well known. Despite these limitations, we present sound evidence for the validity of the adapted version of the IGDS9-SF, a useful instrument to assess the severity of Internet Gaming Disorder in a population of young Spanish students.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (DSM-5®) (5th ed.). American Psychiatric Publishing. https://doi.org/10.1176/appi.books.9780890425596

Beranuy, M., Machimbarrena, J. M., Vega-Osés, M. A., Carbonell, X., Griffiths, M. D., & González-Cabrera, J. (2020). Spanish Validation of the Internet Gaming Disorder Scale–Short Form (IGDS9-SF): Prevalence and Relationship with Online Gambling and Quality of Life. *International Journal of Environmental Research and Public Health, 17*(5), 1562. https://doi.org/10.3390/ijerph17051562

Bernaldo-de-Quirós, M., Labrador-Méndez, M., Sánchez-Iglesias, I., & Labrador, F. (2019). Instrumentos de Medida del Trastorno de Juego en Internet en Adolescentes y Jóvenes Según Criterios DSM-5: Una Revisión Sistemática. *Adicciones*. https://doi.org/10.20882/adicciones.1277

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.

Chau, P. Y. (1997). Reexamining a model for evaluating information center success using a structural equation modeling approach. *Decision Sciences, 28*, 309–334. https://doi.org/10.1111/j.1540-5915.1997.tb01313.x

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

DeCarlo, L. T. (1997). Mardia's multivariate skew (b1p) and multivariate kurtosis (b2p) [SPSS macro]. Retrieved from http://www.columbia.edu/~ld208/

Dinh, T. S., Yasuoka, J., K. C., Poudel, K. C., Otsuka, K., & Jimba, M. (2013). Massively multiplayer online role-playing games (MMORPG): Association between its

addiction, self-control and mental disorders among young people in Vietnam. *International Journal of Social Psychiatry, 59*(6), 570-577. https://doi.org/10.1177/0020764012445861

Evren, C., Dalbudak, E., Topcu, M., Kutlu, N., Evren, B., & Pontes, H. M. (2018). Psychometric validation of the Turkish nine-item InternetGaming Disorder Scale – Short Form (IGDS9-SF). *Psychiatry Research, 265*, 349-354. https://doi.org/10.1016/j.psychres.2018.05.002

Feng, W., Ramo, D. E., Chan, S. R., & Bourgeois, J. A. (2017). Internet gaming disorder: Trends in prevalence 1998-2016. *Addictive Behaviors, 75,* 17-24. https://doi.org/10.1016/j.addbeh.2017.06.010

Fernández-Árias, I., Bernaldo-de-Quirós, M., Sánchez-Iglesias, I., Labrador- Méndez, M.; Estupiñá, F., & González-Álvarez, M. (2019, July 21-24). Identificación de clusters de perfil en el uso y abuso de videojuegos en adolescentes y jóvenes. In F. J. Labrador (Chair), Gamer test: Evaluación e identificación de problemas con los videojuegos [Symposium]. IV Congreso Nacional de Psicología e International Symposium on Psychological Prevention. Vitoria-Gasteiz, Álava, Spain.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179-185. https://doi.org/10.1007/BF02289447

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2018). semTools: Useful tools for structural equation modeling. R package version 0.5-1 [Computer software]. Retrieved from https://CRAN.R-project.org/package=semTools

Kelley, K. (2018). MBESS: The MBESS R Package. R package version 4.4.3. [Computer software]. Retrieved from https://CRAN.R-project.org/package=MBESS

King, D. L., & Delfabbro, P. H. (2016). The Cognitive Psychopathology of Internet Gaming Disorder in Adolescence. Journal of Abnormal Child Psychology, 44(8), 1635–1645. https://doi.org/10.1007/s10802-016-0135-y

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,* 331–342. https://doi.org/10.1016/j.cpr.2013.01.002

Király, O., Sleczka, P., Pontes, H. M., Urbán, R., Griffiths, M. D., & Demetrovics, Z. (2017). Validation of the Internet Gaming Disorder Test (IGDT-10) and evaluation of the nine DSM-5 Internet Gaming Disorder criteria. *Addictive Behaviors, 64*, 253-260. https://doi.org/10.1016/j.addbeh.2015.11.005

Kuss, D. J., & Griffiths, M. D. (2012). Internet gaming addiction: a systematic review of empirical research. *International Journal of Mental Health Addiction 10*(2)*,* 278–296. https://doi.org/10.1007/s11469-011-9318-5

Kuss, D. J., Griffiths, M. D., & Pontes, H. M. (2017a). 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), 103–09. https://doi.org/10.1556/2006.5.2016.062

Lam, L. T. (2014). Internet gaming addiction, problematic use of the internet, and sleep problems: a systematic review. *Current Psychiatry Reports, 16*(4), 444. https://doi.org/10.1007/s11920-014-0444-1

Lemmens, J. S., Valkenburg, P., & Peter, J. (2009). Development and validation of a game addiction scale for adolescents. *Media Psychology, 12*(1), 77–95. https://doi.org/10.1080/15213260802669458

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada [Exploratory item factor analysis: A practical guide revised and updated]. *Anales de Psicología, 30*(3), 1151-1169.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519-530. https://doi.org/10.1093/biomet/57.3.519

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.

Monacis, L., De Palo, V., Griffiths, M. D., & Sinatra, M. (2017). Validation of the Internet Gaming Disorder Scale – Short-Form (IGDS9-SF) in an Italian-speaking sample. *Journal of Behavioral Addictions, 5*(4), 1-8. https://doi.org/10.1556/2006.5.2016.083

Moudiab, S., & Spada, M. M. (2019). The relative contribution of motives and maladaptive cognitions to levels of internet gaming disorder. *AddictiveBehaviors Reports*, 9, 100-160. https://doi.org/10.1016/j.abrep.2019.100160

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*(3), 396–402. https://doi.org/10.3758/BF03200807

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., Király, O., Demetrovics, Z., & Griffitths, M. D. (2014). The conceptualization and measurement of DSM-5 Internet Gaming Disorder: The development of the IGD-20 Test. *PloS ONE, 9*(10)*,* e110137. https://doi.org/10.1371/journal.pone.0110137

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

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of statistical software, 48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

Schivinski, B., Brzozowska-Woś, M., Buchanan, E. M., Griffiths, M. D., & Pontes, H. M. (2018). Psychometric assessment of the Internet Gaming Disorder diagnostic criteria: An Item Response Theory study. *Addictive Behaviors Reports, 8*, 176-184. https://doi.org/10.1016/j.abrep.2018.06.004

Schreiber, J. B. Nora, A., Stage, F. K., Barlow, E. A., & King. J. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research, 99*, 323-338. https://doi.org/10.3200/JOER.99.6.323-338

Scott, J., & Porter-Armstrong, A. P. (2013). Impact of Multiplayer Online Role-Playing Games upon the Psychosocial Well-Being of Adolescents and Young Adults: Reviewing the Evidence. *Psychiatry Journal*, *2013*, 1-8. https://doi.org/10.1155/2013/464685

Van Rooij, A. J., & Kardefelt-Winther, D. (2017). Lost in the chaos: Flawed literature should not generate new disorders: Commentary on: Chaos and confusion in DSM-5 diagnosis of Internet Gaming Disorder: Issues, concerns, and recommendations for

clarity in the field (Kuss et al.). *Journal of Behavioral Addictions, 6*(2), 128-132. https://doi.org/10.1556/2006.6.2017.015

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

Appendix A

Internet Gaming Disorder Scale–Short-Form (IGDS9-SF) (Pontes & Griffiths, 2015)

Instrucciones: Estas cuestiones te preguntan sobre tu actividad de jugar a videojuegos durante el último año (es decir, los últimos 12 meses). Por jugar a videojuegos entendemos cualquier actividad relacionada con jugar a videojuegos, hayas jugado desde un ordenador, videoconsola o cualquier otro tipo de dispositivo (p.ej., móvil, tablet, etc.), *online* u *offline*.

	Nunca	Raras veces	Algunas veces	A menudo	Muy a menudo
1. ¿Te preocupas por jugar a videojuegos? (Algunos ejemplos: ¿Piensas en las veces que has jugado a videojuegos o anticipas la próxima vez que vas a jugar? ¿Crees que jugar a videojuegos se ha convertido en la actividad principal de tu vida diaria?)	○	○	○	○	○
2. ¿Te sientes más irritable, ansioso o incluso triste cuando intentas reducir o parar tu actividad de jugar a videojuegos?	○	○	○	○	○
3. ¿Sientes la necesidad de aumentar el tiempo que dedicas a jugar a videojuegos para conseguir satisfacción o placer?	○	○	○	○	○
4. ¿Cuándo intentas controlar o dejar de jugar a videojuegos fracasas habitualmente?	○	○	○	○	○
5. ¿Has perdido interés en hobbies y actividades de entretenimiento que realizabas antes como resultado de tu enganche con los videojuegos?	○	○	○	○	○
6. ¿Has continuado jugando a videojuegos a pesar de saber que te estaba causando problemas con otras personas?	○	○	○	○	○
7. ¿Has engañado a alguno de tus familiares, psicólogos u otras personas acerca de cuanto juegas a videojuegos?	○	○	○	○	○
8. ¿Juegas para escapar o aliviar de forma temporal una emoción negativa (p.ej., impotencia, culpa, ansiedad)?	○	○	○	○	○
9. ¿Has puesto en peligro o perdido una relación personal importante, un trabajo o una oportunidad educativa o profesional por jugar a videojuegos?	○	○	○	○	○