

BRAZILIAN JOURNAL OF BIOMETRICS

ISSN:2764-5290

ARTICLE

Item response theory in the development of a scale to measure harassment against women in online games¹

 Agatha Pequeno de Mello¹;  Adilson dos Anjos¹;  Suely Ruiz Giolo*¹;  Fernanda da Conceição Zanin²

¹ Department of Statistics, Federal University of Parana, Curitiba-PR, Brazil.

² Federal University of Parana, Associação dos Professores da APUFPR, Curitiba-PR, Brazil.

* Corresponding author. Email: giolo@ufpr.br;

(Received: November 21,2021; Revised: January 20,2022; Accepted: February 22,2022; Published: March 1,2023)

Abstract

This study aimed to develop a scale to measure the harassment experienced by women in multiplayer online games. The study sample consisted of 324 Brazilian female players aged 18 to 44 years recruited from an online survey. Item Response Theory (IRT), particularly the graded response model, was used to evaluate the performance of the items proposed to measure the cited harassment. Moreover, linear regression was used to investigate the association between some women's characteristics and the scale. The results showed that the proposed scale is a nine-item unidimensional psychometric instrument suitable for measuring and better understanding the harassment experienced by women in online gaming spaces. In addition, the association analysis evidenced a high level of harassment among women who reported playing using female nicknames and playing more than three types of games. Finally, most of those who reported a preference for playing with female players had already experienced a high level of harassment.

Keywords: Graded response model; IRT; Measuring instrument; Psychometrics

1. Introduction

Multiplayer online games let people play against friends and other people they meet on a game's network. While online gaming can positively influence the players' cognitive, emotional, and social skills, some forms of harassment can happen in gaming spaces. Although legislation varies between countries, harassment is commonly defined as a crime characterized by abusive conduct of a part or group through words, behaviors, acts, gestures, and writings that can cause damage to the personality, dignity, or physical or psychological integrity of a

person. The two most common forms of harassment in online gaming spaces are cyberbullying and sexual harassment. Cyberbullying extends bullying from physical to virtual environments like cell phones, computers, and tablets. It is a systematic and continuous form of verbal and psychological aggression, in which the victim is attacked based on their appearance or behavior. In practice, they are attitudes in which the victim is repeatedly humiliated, harmed, disqualified, embarrassed, discredited, isolated, and threatened. Sexual harassment, on the other hand, includes any unwelcome or unwanted sexual advances, requests for sexual favors, and other verbal or physical conduct of a sexual nature (Ballard & Welch, 2017).

Women are the most frequent victims of harassment in online gaming spaces. Some usual harassing behaviors against them are attacking their skills or intelligence and insulting, cursing, threatening, or interfering with their game's progress. Further, unwanted sexual advances, sexist jokes (with connotations of inferiority or inequality), rape jokes, and comments about the female player's appearance are commonly reported (Matthew, 2012; Tang and Fox, 2016; Kowert *et al.*, 2017). Depending on the circumstances, the victim does not always perceive this violence or believe it is a joke. However, regardless of the intensity, constancy, and manner in which the harassment occurs, it ends up reaching the victim's self-esteem (Ewoldsen *et al.*, 2012; Ballard & Welch, 2017).

According to Tang and Fox (2016), harassment against women has become a common issue in online gaming communities. In part, this is because many male players have not welcomed the growing number of women in online gaming spaces (O'leary, 2012). Consequently, when they realize that the opponent is a woman, hostility towards them usually increases (Ballard & Welch, 2017).

Gabbiadini *et al.* (2016) reported that playing sexist games such as Grand Theft Auto can decrease men's empathy for women by making them believe that "real men" are dominant and aggressive. Fox and Potocki (2016) also reported an association between sexist games and men's aggression against women. Kuznekoff and Rose (2012) conducted a study using recorded female and male voices to determine the effect of gender voice on communication directed toward that voice. The female voice received, on average, three times as many negative comments as the male voice or no voice.

Women's strategies for coping with harassment while playing online, preserving themselves from inopportune behaviors, include camouflaging their gender, avoiding voice chat, avoiding playing with strangers, leaving online gaming, and adopting an aggressive persona (Cote, 2017; Fox & Tang, 2017).

Since harassment in online gaming spaces has been identified as a relevant problem for female players, this study aimed to (1) develop an instrument based on item response theory to measure the harassment experienced by women in these spaces; (2) investigate characteristics of the female players correlated with high levels of harassment, and (3) compare the levels of harassment estimated from the proposed instrument with situations experienced in online gaming spaces by the female players surveyed in this study.

The paper is organized as follows. Section 2 presents: the set of items proposed to measure the harassment experienced by women in online games, the study participants, and the statistical methodology used to construct and interpret the harassment scale. The results of the analysis performed are presented in Section 3. Section 4 discusses the main findings, followed by a brief conclusion.

2. Materials and Methods

2.1 The measuring instrument

Based on existing literature, including Freitas (2001), Heloani (2004), Dias (2008), Fox and Tang (2014, 2017), and Higa (2016), amongst others, items reflecting the two most common forms of harassment that occur in online gaming spaces (cyberbullying and sexual harassment) were proposed in the current study to assess the latent trait "harassment against women in multiplayer online games." Response options for each item were scored on a five-point Likert scale ranging from 1 to 5 [1- nunca (never), 2- raramente (rarely), 3- algumas vezes (sometimes), 4- frequentemente (often), and 5- muito frequentemente (very often)]. After removing the items with similar content, ten items remained in the measuring instrument. They are presented in Table 1 in Portuguese and translated into English. In our survey, the items were applied in the Portuguese version.

Table 1. The instrument items in the Portuguese version and translated into English

Item description
1 – Ao jogar online, alguém já desqualificou sua habilidade de jogar por causa do seu gênero? 1 – When playing online, has anyone ever disqualified your ability to play because of your gender?
2 – Você já se sentiu prejudicado(a) durante um jogo <i>online</i> por causa do seu gênero? 2 – Have you ever felt harmed during an online game because of your gender?
3 – Você já foi expulso(a) ou pediram para você sair de um jogo <i>online</i> por causa do seu gênero? 3 – Have you ever been kicked out or asked to leave an online game because of your gender?
4 – Você já se sentiu humilhado(a) durante um jogo <i>online</i> por causa do seu gênero? 4 – Have you ever felt humiliated during an online game because of your gender?
5 – Você já foi ameaçado(a) durante um jogo <i>online</i> por causa do seu gênero? 5 – Have you ever been threatened during an online game because of your gender?
6 – Durante um jogo <i>online</i> , já te perguntaram sobre sua vida amorosa ou sua vida pessoal? 6 – Have you ever been asked about your love or personal life during an online game?
7 – Já te chamaram por nomes ofensivos durante um jogo <i>online</i> por causa do seu gênero? 7 – Has someone called you offensive names while playing online because of your gender?
8 – Durante um jogo <i>online</i> , você já recebeu comentários sexuais por causa do seu gênero? 8 – Have you ever received sexual-based comments during an online game because of your gender?
9 – Já te ofereceram itens ou favores durante um jogo <i>online</i> em troca de favores sexuais? 9 – Have you ever been offered items or favors during an online game in exchange for sexual favors?
10 – Algum jogador já te ameaçou ou perseguiu fora de um jogo <i>online</i> ? 10 – Has any player ever chased or threatened you outside of an online game?

Note: The response options for each item are 1- nunca (never), 2- raramente (rarely), 3- algumas vezes (sometimes), 4- frequentemente (often), and 5- muito frequentemente (very often).

2.2 Participants and procedures

The study recruited 324 female players aged 18 to 44 years (mean = 23.6, SD = 4.5). Participants were recruited through social networking sites, including Twitter, Facebook, WhatsApp, and Instagram. The questionnaire was edited into a Google Form®, and its weblink remained available from January 16 to March 26, 2019. The inclusion criteria were: be at least 18 years old, play multiplayer online games, speak Portuguese, reside in Brazil, and provide electronic informed consent. Participation was voluntary and all information remained confidential. Although our study focused on women, players of any gender could participate in the survey to prevent people from feeling excluded.

In addition to the items shown in Table 1, participants were asked about their age, gender, and preferences when playing online games, such as gender of the nickname who prefer to play (male, female, or gender-neutral), types of games who like to play (e.g., action, adventures, running, fighting, sports, shooters, among others), gender of the character who prefer to play (female, male, both, or neutral), and gender of players who prefer to play (male, female, or both). An open-ended question also allowed the participants to report any situation of violence, impertinence, insistence, inconvenience, or something similar experienced by

them in online gaming spaces. The Research Ethics Committee of the Federal University of Parana approved the study (Protocol Number 3111728).

2.3 Data analysis based on item response theory

Item Response Theory (IRT) methods assume that individuals vary along a dominant latent continuum. Thus, as harassment in online games is a latent trait that is difficult to measure, IRT can be a helpful tool to build a metric to measure it. The main assumptions for using unidimensional IRT models are unidimensionality and local independence.

In the current study, Scree plots of the eigenvalues of factors and principal components analyses are presented to determine the number of factors or principal components to retain. Gorsuch (1983) recommends a ratio of the first to second eigenvalues ≥ 3 to indicate essential unidimensionality. Other psychometric indices presented are: Cronbach's alpha and Omega reliability coefficients, explained common variance (ECV), factor determinacy (FD), and construct replicability (H) (Cronbach, 1951; Rodriguez *et al.*, 2016).

Local independence was evaluated using the LD G^2 statistic presented by Chen and Thissen (1997). The discrimination parameters of the Graded Response Model, which represent slopes, were also examined to assess possible violations of the local independence assumption. Items displaying excess dependence may have very high slopes (e.g., > 4.0) relative to other items on the measure (Nguyen *et al.*, 2014).

Item trace plots showing the items' characteristic curves (ICC) are also presented to examine the relationship between the latent ability and the performance on a test item. In addition, the item fit indices $S-X^2$ are calculated for each item as a tool to detect item misfits (Kang & Chen, 2008).

2.4 Item analyses based on the Graded Response Model

The unidimensional Graded Response Model (GRM) proposed by Samejima (1969) was used in the present study to calibrate the items' parameters and select items with good performance. Under this model, the probability of an individual j ($j = 1, \dots, n$) with latent trait θ_j endorsing the category equal to or higher than k from item i is given as

$$P_{i,k}^+(\theta_j) = \frac{1}{1 + \exp[-a_i(\theta_j - b_{i,k})]}$$

for $k = 0, \dots, c_i - 1$, where c_i is the number of response categories of item i . Given that $P_{i,0}^+(\theta_j) = 1$, the parameter for the lowest category $b_{i,0}$ is not estimated. The probability of endorsing the category k is computed by subtracting the adjacent probabilities such that $P_{i,0}(\theta_j) = 1 - P_{i,1}^+(\theta_j)$ and $P_{i,k}(\theta_j) = P_{i,k}^+(\theta_j) - P_{i,(k+1)}^+(\theta_j)$, for $k > 0$.

The discrimination parameter a represents a slope, which refers to how well the item response options discriminate (or differentiate) between female players with high and low levels of harassment. If discrimination is high, the item provides ample information about harassment differences across individuals. Otherwise, the item does not provide much information and may indicate that it needs to be adjusted or removed.

The difficulty parameter b indicates where the item falls on the continuum of the latent trait. The parameters $b_{i,k}$ (i.e., threshold values) are interpreted as how high an individual's harassment level needs to be to have a 0.5 probability of endorsing that given response category or a higher category. In our study, the selected items were those satisfying $0.8 \leq a_i < 4.0$ and $-3 < b_{i,k} < 3$, with $b_{i,1} \leq b_{i,2} \leq \dots \leq b_{i,(c_i-1)}$ (Barbetta *et al.*, 2014).

After estimating the GRM parameters using the maximum likelihood approach, the expectation a posteriori method was used to estimate the latent trait θ , assumed normally distributed with mean zero and unit variance (Baker & Kim, 2004).

2.5 Construction of the Harassment Scale and its interpretation

The Harassment Scale was constructed based on positioning the anchor or quasi-anchor items that characterize each scale point. Since the latent trait θ was assumed normally distributed with mean zero and unit variance, the items with the greatest discrimination were fixed on each scale unit, basically in the range of -3 to $+3$, with $x_{(p-1)} < x_p$ two consecutive levels of the scale. The closer to $+3$, the higher the harassment level, and the closer to -3 , the lower the harassment level.

For category k of an item to be positioned at a certain level x_p of the scale, three conditions must be satisfied simultaneously: 1) the category k has a probability ≥ 0.6 of being endorsed by female players with $\theta \geq x_p$; 2) the category $k-1$ has a probability < 0.5 of being endorsed by female players with $\theta \geq x_{(p-1)}$; and 3) the difference between the probabilities associated with categories k and $k-1$ must be ≥ 0.3 . Items satisfying all three conditions are called anchor items.

As it is difficult to fulfill all of these conditions, items that met two conditions, called quasi-anchors, were considered (Barbetta *et al.*, 2014). The scale levels were defined and interpreted after positioning the anchor and quasi-anchor items on the scale.

2.6 Characteristics correlated with the latent trait

Multiple linear regression was used to examine the correlation of some characteristics with the latent trait θ . The female players' characteristics investigated were: age, the gender of the nickname who prefer to play in online gaming spaces (male, female, or gender-neutral), the number of types of games who prefer to play, the gender of players who prefer to play in online gaming spaces (male, female, or both), and the gender of the character who prefer to play in online gaming spaces (male, female, both, or neutral).

All analyzes were performed in the R software (R Core Team, 2020) using the packages psych (Revelle, 2017), mirt (Chalmers, 2012), lavaan (Rosseel, 2012), and Bifactor Indices Calculator (Dueber, 2021).

3. Results

3.1 Item categories and respective percentages

Table 2 shows the percentages of female players for each item category. For some categories of items, percentages lower than 10% can be observed. A high concentration of answers for the first category of item 10 (72%) can also be observed, indicating that this item does not provide too much information and will probably not perform well in the GRM.

Table 2. Percentages of female players for each item response category

Categories	Items									
	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
1 – Never	12%	20%	44%	26%	46%	12%	24%	18%	55%	72%
2 – Rarely	13%	14%	14%	12%	16%	9%	12%	14%	13%	15%
3 – Sometimes	31%	32%	22%	27%	21%	21%	23%	24%	17%	10%
4 – Often	21%	19%	12%	18%	10%	27%	18%	20%	9%	2%
5 – Very often	23%	16%	8%	17%	6%	31%	24%	25%	7%	2%

3.2 Dimensionality Analysis

The Scree plots shown in Figure 1 suggest that one factor or principal component is sufficient to retain. In addition, the first to second eigenvalues ratios are ≥ 3 , indicating essential unidimensionality according to Gorsuch's (1983) recommendation

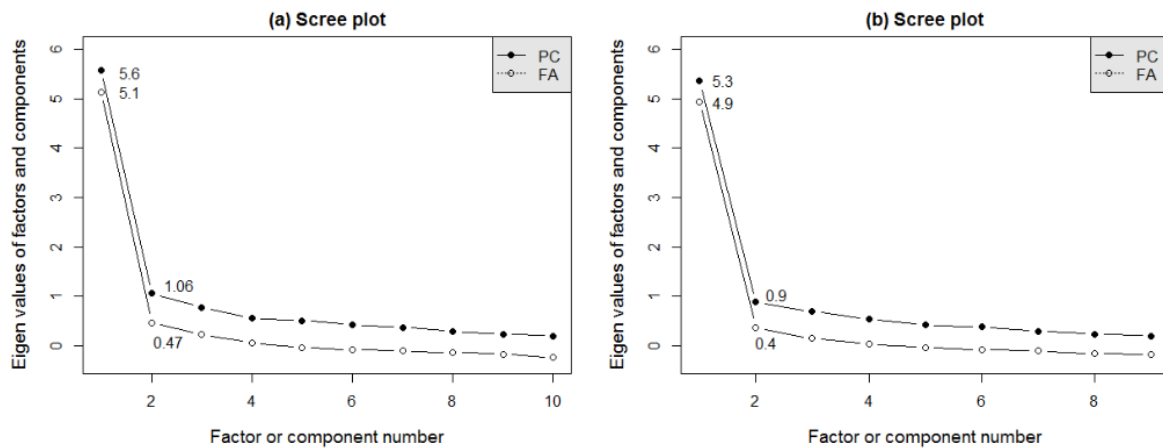


Figure 1. Scree plots of the eigenvalues of factors (FA) and principal components (PC) analyses considering (a) items I1 to I10 and (b) items I1 to I9.

Considering all ten items (I1 to I10), Cronbach's alpha coefficient was 0.89 and 0.91 considering nine items (I1 to I9). Moreover, the Omega coefficient, high factor determinacy, and construct replicability, without item I10, were: Omega = 0.93, FD = 0.97, and H = 0.95, respectively, suggesting a well-defined unidimensional latent variable. Also, the loading values without item I10 varied from 0.60 to 0.91, and the ECV of 0.65 evidenced that the common variance is essentially unidimensional. Therefore, the unidimensional GRM was considered for item parameters estimation and item analysis.

3.3 IRT parameters estimation based on the GRM

The IRT parameter estimates for item discrimination and item difficulty displayed in Table 3 show the nine items that performed well ($0.8 \leq a_i < 4.0$ and $-3 < b_{i,k} < 3$). Some item response categories were grouped to estimate the parameters better. As expected, item 10 did not perform well, thus being removed from the model.

The difficulty parameters (thresholds values) are on the same scale as the Z-scale, where a normal distribution is centered at zero with a unit standard deviation. For example, item I4 has the following threshold values: $b_1 = -0.32$, $b_2 = 0.39$, and $b_3 = 1.05$. These values are the cutoff points that intercept the lines representing the probability of endorsing a category given a certain level on the latent trait represented by the x-axis in the trace plot shown for item I4 in Figure 2(a).

Table 3. Estimates of discrimination and difficulty parameters

Item	a (SE)	Threshold response categories and parameters			
		b_1 (SE)	b_2 (SE)	b_3 (SE)	b_4 (SE)
11	3.31 (0.33)	Rarely -1.34 (0.11)	Sometimes -0.74 (0.09)	Often 0.14 (0.07)	Very often 0.80 (0.09)
12	2.99 (0.30)	Sometimes -0.47 (0.08)	Often 0.44 (0.08)	Very often 1.15 (0.11)	--
13	2.46 (0.27)	Sometimes 0.22 (0.08)	Often 0.98 (0.11)	Very often 1.70 (0.15)	--
14	3.85 (0.42)	Sometimes -0.32 (0.07)	Often 0.39 (0.08)	Very often 1.05 (0.10)	--
15	1.51 (0.19)	Sometimes 0.48 (0.11)	Often or very often 1.51 (0.18)	--	--
16	1.70 (0.18)	Sometimes -1.13 (0.13)	Often -0.26 (0.10)	Very often 0.70 (0.11)	--
17	3.55 (0.38)	Sometimes -0.40 (0.08)	Often 0.22 (0.07)	Very often 0.78 (0.09)	--
18	2.64 (0.26)	Rarely -1.09 (0.11)	Sometimes -0.54 (0.09)	Often 0.15 (0.08)	Very often 0.78 (0.10)
19	1.29 (0.19)	Sometimes, often or very often 0.77 (0.14)	--	--	--

Note: a is the discrimination parameter and b_k the difficulty parameters. SE: standard error.

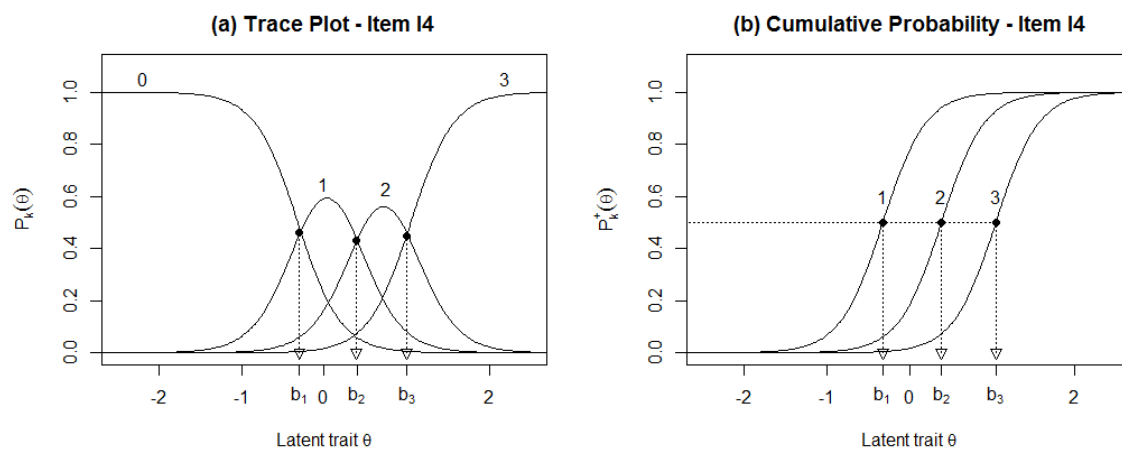


Figure 2. (a) Trace plot for item I4 where each line reflects the probability of endorsing the category k ($0 =$ never or rarely, $1 =$ sometimes, $2 =$ often, and $3 =$ very often), (b) $P_k^+(\theta)$ plot where each line reflects the probability of endorsing the category equal to or higher than k ($k = 1, 2, 3$); $b_1 = -0.32$, $b_2 = 0.39$, and $b_3 = 1.05$ are the thresholds values.

From Figure 2(a), female players with a harassment level below the first threshold (i.e., $b_1 = -0.32$) have a high probability of endorsing the category $k = 0$ (never or rarely) of item I4. In addition, female players with a harassment level between -0.32 and 0.39 are most likely to endorse the category $k = 1$ (sometimes), and those with a harassment level between 0.39 and 1.05 , the category $k = 2$ (often). Finally, those with a harassment level above 1.05 are most likely to endorse the category $k = 3$ (very often). Figure 2(b) shows that the cumulative probability curves $P_k^+(\theta)$ for item I4, $k = 1, 2, 3$, do not cross and continually increase at higher levels of the latent trait.

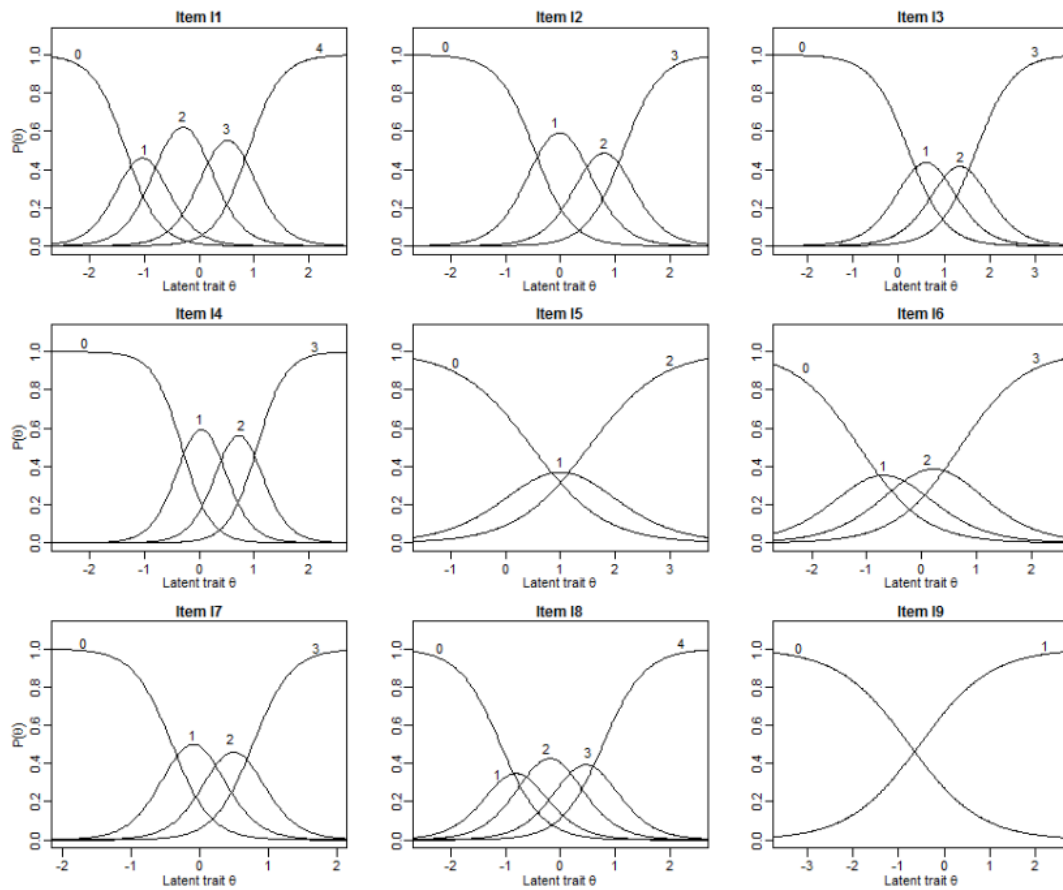


Figure 3. Trace plots for items I1 to I9. Each line reflects the probability of endorsing an item response category given the latent trait level.

The trace plots for all nine items can be seen in Figure 3. They show the item-level information regarding the performance linked to each response category. Items with larger discrimination levels tend to provide information about the female players' harassment in a narrow range (e.g., item I4, which had the largest discrimination value of 3.85).

The test information curve shown in Figure 4 indicates that the instrument is best suited for measuring the harassment level in the range from -2 to 2 (where the curve is highest). Then, the instrument provides more information (or assesses more accurately) female players with harassment in this range.

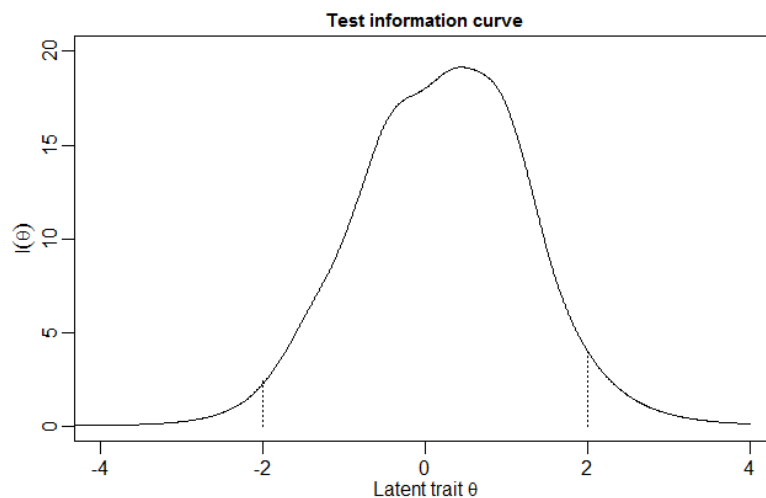


Figure 4. The test information curve $I(\theta)$ of the harassment instrument.

3.4 Item fit and local dependence

After controlling for the false discovery rate, a correction of multiple comparisons better suited in IRT models than the Bonferroni correction (Depaoli *et al.*, 2018), the S-X² item fit statistic evidenced that the items fit the questionnaire well (p -values ≥ 0.07 , Table 4). There were also no issues with LD, as the local independence assumption among items on the Harassment Scale was supported by the LD G² test (p -values > 0.05), as well as by discrimination parameters smaller than 4.0 (see Table 4).

Table 4. Item fit statistics with p -values adjusted for false discovery rate

Item	S-X ²	Df	p -value	Item	S-X ²	Df	p -value
I1	51.05	26	0.07	I6	48.23	41	0.34
I2	27.04	30	0.62	I7	43.76	30	0.17
I3	22.26	25	0.62	I8	38.27	37	0.51
I4	42.38	26	0.11	I9	21.08	20	0.51
I5	40.97	34	0.34				

3.5 Harassment Scale levels and interpretation

The anchor items were: I2, I4, I5, I6, and I8, and quasi-anchor: I1, I3, I7, and I9. After positioning the items that characterize each scale point, five levels were defined for the Harassment Scale. These levels and their interpretations are shown in Table 5.

Table 5. The five levels of the Harassment Scale and their respective interpretations

Harassment level	Anchoring of items	Interpretation of the Harassment Scale at each level
Very low $\theta < -1$	Never	Female players at this level are highly likely to endorse that they never experienced harassment while playing online.
Low $-1 \leq \theta < 0$	I1b ₁ I6b ₁ I8b ₁	At this level, female players are highly likely to endorse that they rarely experienced harassment while playing online. They sometimes are asked about their love or personal life.
Moderate $0 \leq \theta < 1$	I1b ₂ I6b ₂ I2b ₁ I7b ₁ I4b ₁ I8b ₂	Female players at this level are highly likely to endorse that they sometimes experience harassment while playing online. Sometimes, they feel harmed and humiliated, and sometimes someone calls them offensive names. Sometimes, they also receive sexual-based comments and often are asked about their love or personal life.
High $1 \leq \theta < 2$	I1b ₃ I6b ₃ I2b ₂ I7b ₂ I3b ₁ I8b ₃ I4b ₂ I9b ₁ I5b ₁	At this level, female players are highly likely to endorse that they often experience harassment while playing online. They often are disqualified by their ability to play and often feel harmed and humiliated. Sometimes they are threatened and asked to leave the online game. Moreover, they are very often asked about their love or personal life; someone often calls them offensive names, and they often receive sexual-based comments. Besides, they often or sometimes receive item offers in exchange for sexual favors.
Very high $\theta \geq 2$	I1b ₄ I3b ₃ I2b ₃ I3b ₂ I4b ₃ I5b ₂ I7b ₃ I8b ₄	Female players at this level are highly likely to endorse that they very often experience harassment while playing online. They very often are offended by disqualifying their ability to play and very often feel harmed and humiliated. In addition, they often or very often are threatened and asked to leave the game. Someone very often calls them offensive names, and they very often receive sexual-based comments.

Note: b₁, b₂, and b₃ indicate the item's category shown in Table 3.

Based on the five levels defined for the Harassment Scale and the θ values estimated from it, 54 (16.7%)

of the female players in the present study were classified as having a very low level of harassment and 114 (35.2%) as having a low level of harassment. In addition, the number of female players classified as having moderate, high, and very high levels of harassment was 108 (33.3%), 39 (12%), and 9 (2.8%), respectively.

For three of the five levels of the Harassment Scale, Table 6 illustrates the instrument's suitability for measuring the harassment level experienced by women in multiplayer online games. This table makes it possible to note the agreement between the levels assigned from the scale and the situations reported by the female players in the open-ended question.

Table 6. Agreement between three levels assigned to women from the Harassment Scale and the most common situations experienced by them in online gaming spaces

Estimated Harassment level	Most common situations reported by women classified at each level
Very low $\hat{\theta} < -1$ $n = 54$	They reported any experience of harassment since they usually only play with friends. Some disable the voice chat or audio to avoid unwanted situations, even though they have never experienced them.
Moderate $0 \leq \hat{\theta} < 1$ $n = 108$	They sometimes experienced insults, verbal aggressions, offensive names, comments disqualifying their skills to play, hate messages, sexist comments, and sometimes received offers in exchange for sexual favors. Many were asked about their love or personal life.
High $1 \leq \hat{\theta} < 2$ $n = 39$	They often experienced sexist comments, gender-based insults, verbal sexual assaults, and impertinent questions accompanied by comments about their appearance and tone of voice. They often are asked about their relationship status. Some were expelled from online games because male players said that games are not for women. Some also reported avoiding voice chat and using female nicknames because of persistent gender-based insults.

Note: n = number of female players in the study classified at each displayed level.

Concerning gaming choices, 61.4% of the female players surveyed in the current study reported playing four or more different types of online games. The most cited were: action, adventure, running, sports, real-time strategy, fighting, RPG, simulation, and shooters. Furthermore, 57.1% reported playing with female nicknames and 24.4% with female players only (mainly because of previous experiences of harassment).

When gaming choices, motives, play behavior, and performance are considered across genders, the literature has pointed out several gender differences and similarities. Men, for instance, are more competitive than women are, and they are also more likely to play action and simulation games, while women are more likely to play logic and skill training games (Veltri *et al.*, 2014; Cary *et al.*, 2020).

Finally, the multiple linear regression results displayed in Table 7 show that the female players' age group and the gender of the character they prefer to play in online games did not significantly correlate with harassment. When taken as a continuous variable, age also showed no significant association with harassment. On the other hand, the preference to use female nicknames, play with female players only, and play more than three different types of games positively correlated with harassment. Possible reasons to explain this correlation are discussed in the next section.

Table 7. Analysis of the female characteristics associated with the Harassment Scale

Characteristics	<i>n</i> (%)	Estimates	
		B (95%CI)	<i>p</i> -value
Age groups			
18 to 23 years	182 (56.2)	Reference	
24 to 44 years	142 (43.8)	-0.17 [-0.38, 0.03]	0.10
Number of types of games who like to play			
1 to 3	125 (38.6)	Reference	
4 or above	199 (61.4)	0.24 [0.03, 0.44]	0.04*
Gender of players who prefer to play			
Male	19 (5.8)	Reference	
Female	79 (24.4)	0.82 [0.35, 1.29]	< 0.001**
Both	226 (69.8)	0.32 [-0.11, 0.76]	0.15
The gender of character who prefer to play			
Male	20 (6.2)	Reference	
Female	181 (55.9)	-0.28 [-0.74, 0.18]	0.24
Both	107 (33.0)	-0.36 [-0.83, 0.10]	0.12
Neutral	16 (4.9)	-0.40 [-1.03, 0.21]	0.20
The nickname used in online games			
Male	24 (7.4)	Reference	
Female	185 (57.1)	0.48 [0.05, 0.90]	0.03*
Gender-neutral	115 (35.5)	0.36 [-0.06, 0.78]	0.10

Note: Intercept $B_0 = -0.57$ (SE = 0.32), B: regression coefficient and CI: confidence interval. Statistically significant at the significance level of *0.05 and **0.001.

4. Discussion

Aggressive and hostile behaviors against women in online gaming spaces have been systematically documented. Matthew (2012) found that women are four times more likely than men to experience sex-based taunting or harassment. Brehm (2013) reported that more than 63% of their surveyed females experienced sexism or misogynistic behavior from other players. The study by Kuznekoff and Rose (2012) also confirmed sex-based discrimination in online gaming spaces. Thus, considering that harassment against women is recurrent in online gaming spaces, the present study proposed a scale to assess the level of harassment experienced by them in these spaces.

The motivations for harassment targeted at female players in online games have been the object of investigation and discussion by some researchers. Among the potential reasons for harassment in online games, Tang and Fox (2016) argue that some male players may see female players as a threat to their social identity. Thus, they try to demonstrate their superiority and defend what they consider "their" dominance through harassment and insults directed at female players. The socialization of video gameplay as a "boy's toy" and the stereotype that male players outperform their female counterparts also contribute to female gamers' exclusion or marginalization (Kowert *et al.*, 2017). Furthermore, the gaming industry itself is highly skewed, with female designers, programmers, and producers comprising only a fragment of its workers. By allowing the sexist discourse to continue, the game industry (from game developers to marketers) should be aware that they may be shutting out half of the global market (Fox; Tang, 2014; Veltri *et al.*, 2014).

In our study, women who reported playing more than three types of games (61.4%) experienced higher levels of harassment than those who usually played less than three. One reason could be that exposure to more types of games is likely related to an increased risk of encountering male players with aggressive behavior. Furthermore, as many video games tend to be male-centric, they fail to provide female characters and end up creating an atmosphere of exclusion for female players (Assunção, 2016; Kowert *et al.*, 2017).

Kuznekoff and Rose (2012) and Fox and Tang (2017) reported that women are the most frequent victims of harassment in online gaming spaces. In the present study, the use of female nicknames in online games, reported by 57.1%, was positively associated with harassment, corroborating that women are frequent victims of harassment. In this regard, harassment against women in online gaming spaces may be related to female characters being underrepresented in multiplayer online games (Waddell *et al.*, 2014).

Furthermore, there is a crucial distinction in the clothing of male and female characters. Most female characters in video games are more sexualized than male characters, which can be viewed as a primer indicator of sex roles in video games, possibly cultivating or reinforcing biased gender roles or sexist beliefs and attitudes (Kowert *et al.*, 2017). Some video games also include content that female players may find offensive or demeaning, and others contribute to the proliferation of narratives that video games are primarily "boy's toys" (Dickey, 2006; Kowert *et al.*, 2017). Thus, the gaming industry should be aware that it is necessary to avoid stereotyping women in online games, increasing efforts to diminish hostile attitudes towards female players before they become an accepted part of gamer cultures (Kaye & Pennington 2016; Kowert *et al.*, 2017).

Finally, some studies have shown that it is usual for women who already had experienced harassment in online games to adopt coping strategies, preserving themselves from men's inopportune behaviors (Cote, 2017, Fox & Tang, 2017). In our study, the preference of playing with female players positively correlated with harassment. Among the 24.4% who reported this preference, around 90% had already experienced a high level of harassment. In addition to the preference for playing with female players, they reported having already adopted other strategies such as disabling voice chat and choosing to play with male or gender-neutral nicknames. Some of them also stopped playing particular games permanently because of sexual harassment or verbal aggression.

5. Conclusions

Our findings indicated that the Harassment Scale proposed in this study is a suitable unidimensional nine-item psychometric instrument to assess harassment against women in online gaming spaces. As this is the first study to examine the theoretical structure and psychometric properties of the Harassment Scale, its application in other populations would be valuable, thus providing additional foundations for its use. Furthermore, a more detailed analysis of situations reported in open-ended questions using, for instance, text analysis could be advantageous and is suggested for future studies.

Acknowledgments

We thank all respondents who voluntarily participated in the survey and those who helped disseminate the survey during the data collection process.

Conflicts of Interest

The authors declare no conflict of interest.

6. References

1. Assunção, C. No girls on the internet: the experience of female gamers in the masculine space of violent gaming. *Press Start*. **3** (1), 46-65 (2016).
2. Ballard, M. E.; Welch, K. M. Virtual warfare: cyberbullying and cyber-victimization in MMOG play. *Games and Culture*. **12** (5), 466-491 (2017).
3. Baker, F. B.; Kim, S. H. *Item Response Theory: parameter estimation techniques*. (2.ed. New York: Marcel Dekker, Inc, 2004. 528p).
4. Barbeta, P. A.; Trevisan, L. M. V.; Tavares, H.; Azevedo, T. C. A. M. Aplicação da teoria da resposta ao item uni e multidimensional. *Estudos em Avaliação Educacional*. **25** (57), 280-302 (2014).
5. Brehm, A. Navigating the feminine in massively multiplayer online games: gender in World of Warcraft. *Frontiers in Psychology*. **4**, 903 (2013).

6. Cary, L. A.; Axt, J.; Chasteen, A. L. The interplay of individual differences, norms, and group identification in predicting prejudiced behavior in online video game interactions. *Journal of Applied Social Psychology*. **50** (11), 623-637 (2020).
7. Chalmers, R. P. A. Multidimensional item response theory package for the R environment. *Journal of Statistical Software*. **48** (6), 1-29 (2012).
8. Chen, W-H.; Thissen, D. Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*. **22** (3), 265-289 (1997).
9. Cote, A. C. I can defend myself: women's strategies for coping with harassment while gaming online. *Games and Culture*. **12** (2), 136-155 (2017).
10. Cronbach, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika*. **16**, 297-334 (1951).
11. Depaoli, S.; Tiemensma, J.; Felt, J. M. Assessment of health surveys: fitting a multidimensional graded response model, *Psychology, Health & Medicine*. **23**, 1299-1317 (2018).
12. Dias, I. Violência contra as mulheres no trabalho: o caso do assédio sexual. *Sociologia, Problemas e Práticas*. **57**, 11-23 (2008).
13. Dickey, M. D. Girl gamers: the controversy of girl games and the relevance of female-oriented design for instructional design. *British Journal of Educational Technology*. **37** (5), 785-793 (2006).
14. Dueber, D. M. *R Package: BifactorIndicesCalculator*. Available at <https://cran.r-project.org/web/packages/BifactorIndicesCalculator/BifactorIndicesCalculator.pdf>, (2021).
15. Ewoldsen, D. R.; Eno, C. A.; Okdie, B. M.; Velez, J. A.; Guadagno, R. E.; Decoster, J. Effect of playing violent video games cooperatively or competitively on subsequent cooperative behavior. *Cyberpsychology, Behavior, and Social Networking*. **15** (5), 277-280 (2012).
16. Fox, J.; Potocki, B. Lifetime video game consumption, interpersonal aggression, hostile sexism, and rape myth acceptance: a cultivation perspective. *Journal of Interpersonal Violence*. **31** (10), 1912-1931 (2016).
17. Fox, J.; Tang, W. Y. Sexism in online video games: the role of conformity to masculine norms and social dominance orientation. *Computers in Human Behavior*. **33**, 314-320 (2014).
18. Fox, J.; Tang, W. Y. Women's experiences with general and sexual harassment in online video games: rumination, organizational responsiveness, withdrawal, and coping strategies. *New Media & Society*. **19** (8), 1290-1307 (2017).
19. Freitas, M. E. Assédio moral e assédio sexual: faces do poder perverso nas organizações. *Revista de Administração de Empresas*. **41** (2), 8-19 (2001).
20. Gabbiadini, A.; Riva, P.; Andrighetto, L.; Volpato, C.; Bushman, B. J. Acting like a tough guy: violent-sexist video games, identification with game characters, masculine beliefs, & empathy for female violence victims. *PLoS One*. **11** (4), 1-14 (2016).
21. Gorsuch, R. L. *Factor analysis*. (2.ed. Hillsdale, NJ: Erlbaum, 1983. 425p).
22. Heloani, R. Assédio moral: um ensaio sobre a expropriação da dignidade no trabalho. *RAE-eletrônica*. **3** (1), (2004).
23. Higa, F. C. Assédio sexual no trabalho e discriminação de gênero: duas faces da mesma moeda? *Revista Direito GV*. **12** (2), 484-515 (2016).

24. Kang, T.; Chen, T. T. Performance of the generalized S-X2 item fit index for polytomous IRT Models. *Journal of Educational Measurement*. **45** (4), 391-406 (2008).
25. Kaye, L. K.; Pennington, C. R. Girls can't play: the effects of stereotype threat on females' gaming performance. *Computer in Human Behavior*. **59**, 202-209 (2016).
26. Kowert, R.; Breuer, J.; Quandt, T. "Women are from FarmVille, men are from ViceCity: the cycle of exclusion and sexism in video game content and culture," in *New Perspectives on the Social Aspects of Digital Gaming. Multiplayer*, eds R. Kowert and T. Quandt (London: Routledge), **2**, (2017).
27. Kuznekoff, J. H.; Rose, L. K. Communication in multiplayer gaming: Examining player responses to gender cues. *New Media & Society*. **15** (4), 541-556 (2012).
28. Matthew, E. Sexism in video games [Study]: there is sexism in gaming. *Price Charting*, 2012. Retrieved from <https://blog.pricecharting.com/2012/09/emilyami-sexism-in-video-games-study.html>
29. Nguyen, T. H.; Han, H. R.; Kim, M. T.; Chan, K. S. An introduction to item response theory for patient-reported outcome measurement. *Patient*. **7** (1), 23-35 (2014).
30. O'Leary, A. In virtual play, sex harassment is all too real. *The New York Times*. 2012. Retrieved from <https://www.nytimes.com/2012/08/02/us/sexual-harassment-in-online-gaming-stirs-anger.html>
31. R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>, (2020).
32. Revelle, W. R. *psych: procedures for personality and psychological research*. <https://CRAN.R-project.org/package=psych/>, (2017)
33. Rodriguez, A.; Reise, S. P.; Haviland, M. G. Applying bifactor statistical indices in the evaluation of psychological measures, *Journal of Personality Assessment*. **98** (3), 223-237 (2016).
34. Rosseel, Y. lavaan: An R package for structural equation modeling. *Journal of Statistical Software*. **48** (2), 1-36 (2012).
35. Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*. **34** (4), (1969).
36. Tang, W. Y.; Fox, J. Men's harassment behavior in online video games: Personality traits and game factors. *Aggressive Behavior*. **42** (6), 513-531 (2016).
37. Veltri, N. F.; Krasnova, H.; Baumann, A.; Kalayamthanam, N. Gender Differences in Online Gaming: A Literature Review. In: *Twentieth Americas Conference on Information Systems*, p.1-11, 2014, Savannah. *Anais eletrônicos...*, Savannah, 2014. Available at <https://core.ac.uk/download/pdf/301361965.pdf> Accessed in: 20 Jul. 2021.
38. Waddell, T. F.; Ivory, J. D.; Conde, R.; Long, C.; McDonnell, R. White man's virtual world: A systematic content analysis of gender and race in massively multiplayer online games. *Journal of Virtual Worlds Research*. **7** (2), 1-14 (2014).