Measuring Perceived Severity of Intimate Partner Violence against Women (IPVAW) among the General Population and IPVAW Offenders

Manuel Martín-Fernández, Enrique Gracia, and Marisol Lila

*Autonomous University of Madrid, Spain; ‡University of Valencia, Spain

ARTICLE INFO
Received 12 February 2022
Accepted 18 April 2022

Keywords:
Intimate partner violence against women
Perceived severity
Intimate partner violence offenders
Measurement invariance
Item response theory

ABSTRACT
Public perceptions of the severity of intimate partner violence against women (IPVAW) incidents are an important factor that has been linked to key issues regarding this type of violence, such as acceptability or tolerance, personal sense of responsibility, attitudes toward intervention, and the public’s, professionals’, and victims’ responses to IPVAW. The aim of the present study was to provide further validity evidence for the perceived severity of IPVAW scale (PS-IPVAW), by assessing its measurement invariance between gender and age groups, and between men from the general population and male IPVAW offenders. Item response theory was also used to assess the discrimination of the items and their position on the measured latent trait continuum (i.e., perceived severity of IPVAW). To this end, the psychometric properties of the scale were examined in four different samples from the general population (N = 2,627) and in one clinical sample of male IPVAW offenders (N = 200). Our findings showed that the PS-IPVAW scale has excellent internal consistency (α = .89-.90) and a clear one-factor latent structure (CFI = .91-.96, RMSEA = .055-.086), and that partial strict invariance holds across different gender and age groups. We also found that IPVAW offenders’ perceptions of the severity of IPVAW may follow a different pattern to that of men from the general population. The PS-IPVAW scale is able to yield accurate assessments of the perceived severity of this type of violence among the general population and IPVAW offenders.

Evaluación de la gravedad percibida de la violencia de pareja contra la mujer en la población general y en hombres agresores de pareja

RESUMEN
La percepción pública de la gravedad de los incidentes de violencia contra la mujer en las relaciones de pareja (VMP) es un factor importante relacionado con aspectos clave de este tipo de violencia, como su aceptabilidad o tolerancia, el sentido de responsabilidad personal, las actitudes hacia la intervención y las respuestas a la VMP de la población general, los profesionales y las víctimas. El objetivo de este estudio fue proporcionar nuevas evidencias de validez de la escala PS-IPVAW [Perceived Severity of Intimate Partner Violence against Women], analizando su invarianza factorial entre géneros y distintos grupos de edad y entre hombres de la población general y hombres agresores de pareja. También se utilizó la teoría de respuesta al ítem para evaluar la discriminación de los ítems y su posición en el continuo del rasgo latente evaluado (i.e., la gravedad percibida de la VMP). Se examinaron las propiedades psicométricas de la escala en cuatro muestras de la población general (N = 2,627) y en una muestra clínica de hombres agresores de pareja (N = 200). Los resultados mostraron que la escala PS-IPVAW tuvo una excelente consistencia interna (α = .89-.90), una estructura latente unifactorial (CFI = .91-.96, RMSEA = .055-.086) y que se puede mantener el nivel de invarianza factorial estricta parcial entre géneros y distintos grupos de edad. Se encontró también que los hombres agresores de pareja mostraban un patrón diferencial en la evaluación de la gravedad percibida de la VMP en comparación con hombres de la población general. La escala PS-IPVAW es un instrumento capaz de evaluar de forma precisa la gravedad percibida de la VMP en la población general y en hombres agresores de pareja.

Cite this article as: Martín-Fernández, M., Gracia, E., & Lila, M. (2022). Measuring perceived severity of intimate partner violence against women (IPVAW) among the general population and IPVAW offenders. Psychosocial Intervention, 31(2), 109-119. https://doi.org/10.5093/pi2022a8

Funding: This study has been supported by FEDER/Spanish Ministry of Science, Innovation, and Universities – State Research Agency (PSI2017-84764-P) and by the program of stimuli for young researchers of the Autonomous Community of Madrid (S3/JIII/2021-00135). Correspondence: manuel.martinfernandez@uam.es (M. Martín-Fernández).

ISSN:1132-0559/© 2022 Colegio Oficial de la Psicología de Madrid. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Intimate partner violence is the most common form of violence suffered by women (Devries et al., 2013; Stöckl et al., 2013). With an average global lifetime prevalence of 27% and regional prevalence estimates ranging from 16% to 51%, intimate partner violence against women (IPVAW) is considered by the World Health Organization (WHO) as a major public health problem of epidemic proportions (WHO, 2013, 2021). IPVAW is a human rights violation and a barrier to gender equality and women’s empowerment, and constitutes one of the main challenges that modern societies must address to align with United Nations (UN) Sustainable Development Goal 5.2 to “eliminate all forms of violence against all women and girls in the public and private spheres, including trafficking and sexual and other types of exploitation”, set to be achieved by 2030 (United Nations [UN], 2015).

Beyond individual and relational factors, the phenomenon of IPVAW is also rooted in its socio-cultural context (Gracia et al., 2020; Heise, 1998, 2011; WHO, 2002). Public perceptions and attitudes toward IPVAW reflect the social and cultural context in which IPVAW takes place. The degree to which this type of violence in intimate relationships is tolerated or accepted, perceived as serious and of public concern, and thus deemed worthy of personal or institutional involvement, contributes to creating a social climate in which IPVAW can be facilitated or deterred (Carlson & Worden, 2005; Gracia, 2004; Powell & Webster, 2018; Waltermaurer, 2012; WHO, 2002). In this regard, a large body of research has shown that public perceptions and attitudes toward IPVAW are closely related to their perpetration, reporting and incidence rates, a victim’s help-seeking behavior, and public and professional responses to cases of IPVAW (Archer & Graham-Kevan, 2003; Capaldi et al., 2012; Copp et al., 2019; Faramarzi et al., 2005; Gracia et al., 2015; Rizo & Macy, 2011; Sardinha & Catalão, 2018; Worden & Carlson, 2005).

Research has typically focused on four broad categories of public perceptions and attitudes toward IPVAW (Gracia et al., 2020): legitimation of IPVAW (e.g., victim blaming, justification), acceptability of IPVAW (e.g., tolerance, approval), attitudes toward intervention in cases of IPVAW (e.g., willingness to help, attitudes toward reporting, propensity to intervene), and perceived severity of IPVAW (e.g., seriousness, or minimization). Interestingly, of these four categories, perceived severity of IPVAW has received the least scholarly attention (Gracia et al., 2020). Perceived severity of IPVAW is, however, an important factor that has been linked to key issues regarding this type of violence, such as its acceptability or tolerance, personal sense of responsibility, attitudes toward intervention, and responses to incidents of this type of violence among the general public, professionals, and victims of IPVAW (Gracia et al., 2020).

The public’s, professionals’, and victims’ responses may depend on the perceived severity of IPVAW incidents. If some IPVAW incidents are perceived as normal, trivial, or not serious enough, and only certain types of IPVAW behaviors that involve repeated, severe, or extreme violence are considered sufficiently serious, it is more likely that some violence toward women in intimate relationships may be deemed acceptable or tolerable, and therefore not deserving the involvement and intervention of people who are aware of the incidents or even efforts by the victim to seek help (Gracia, García, & Lila, 2009; Gracia et al., 2018; Leon et al., 2021). For example, according to psychosocial models of bystander intervention (Latané & Darley, 1970; Piliavin et al., 1981), the perceived severity of an incident is a precondition to the decision to intervene and, therefore, if some incidents of IPVAW are perceived as not serious enough, those aware of the incident will be less willing to intervene (Gracia, Herrero, et al., 2009; Gracia et al., 2018). In this regard, a substantial body of research has illustrated the link between the perceived severity of IPVAW incidents and the type of public and professional responses and intervention, as well as victims’ help-seeking behaviors (Baldray & Pagliaro, 2014; Djikanovic et al., 2012; Ergöçmen et al., 2013; Fanslow & Robinson, 2010; Frías, 2013; Gracia et al., 2008; Gracia, García, et al., 2009; Gracia, Herrero et al., 2009; Hyman et al., 2009; Kiss et al., 2012). Finally, another important reason to address the perceived severity of IPVAW is that perception of incidents of this type of violence as normal, trivial, or not serious enough is still quite prevalent. For example, a report on attitudes toward violence against women in the European Union (Gracia & Lila, 2015) found that relevant percentages of women respondents from different countries and age groups did not want to talk about or report violence because “the incident was too trivial”, because “they do not consider aggressions as something serious”, or because “the victims considered the experienced violence as normal or not serious.”

In light of the above, it becomes apparent that psychometrically sound instruments are needed to accurately assess the perceived severity of IPVAW across community and clinical samples, thereby advancing our knowledge on public perceptions, attitudes, and responses to IPVAW, and contributing to the prevention efforts to reduce and gradually eliminate this type of violence in our societies. To this end, in this study we present further validity evidence for the perceived severity of IPVAW (PS-IPVAW) scale, a widely used questionnaire to measure this construct.

The Present Study

The perceived severity of IPVAW (PS-IPVAW) scale is an 8-item measure developed initially to assess this construct both among the general population and law enforcement professionals (Gracia et al., 2008; Gracia, García, et al., 2009). It has been widely used in different settings (including Ecuador, Portugal, Spain, and USA), with different purposes and samples, including the general public, law enforcement professionals, and IPVAW offenders (Catalá-Miñana et al., 2013; El Sayed et al., 2020; Expósito-Álvarez et al., 2021; Gracia et al., 2011, 2014; Lila et al., 2013; Lila et al., 2019; Martín-Fernández, Gracia, & Lila, 2018; Martin-Fernández, Gracia, Marco, et al., 2018; Martín-Fernández et al., 2021; Sani et al., 2018; Vargas et al., 2017; Villagrán et al., 2020, 2022; Vitoria-Estruch et al., 2017).

It is important, however, to continue examining the validity evidence for the PS-IPVAW scale. On the one hand, although this scale has been used as a one-dimensional measure, to the best of our knowledge, no study has tested its latent structure using factor analysis. On the other hand, the measurement invariance of the PS-IPVAW scale has not been examined across different gender and age groups. Establishing measurement invariance is an important step before making comparisons between groups to rule out a potential measurement bias. Ascertainment whether respondents from different gender and age groups conceptualize the construct and interpret the items in a similar way allows researchers to make proper and valid comparisons across such groups (Putnick & Bornstein, 2016). Neither has the measurement invariance of the PS-IPVAW scale been tested across men from the general population and IPVAW offenders, which is important not only to rule out measurement bias between these populations, but to establish whether IPVAW offenders show a different pattern in the way they perceive the severity of this type of violence. Testing measurement invariance between men from the general population and IPVAW offenders is also crucial to develop new intervention strategies and improve existing ones used with this clinical population.

Hence, the aim of the present study was to provide further validity evidence for the PS-IPVAW scale. To this end, the psychometric properties of the scale were examined in four different samples by assessing the scale’s internal consistency and latent structure, and testing its measurement invariance across gender and age groups. The measurement invariance of the scale between men from the general population and a clinical sample of IPVAW offenders was also tested. In addition, the discrimination of the items and their position on the latent trait continuum (i.e., perceived severity of IPVAW) were assessed through item response theory (IRT). Validity
Method

Participants and Procedure

A total of 2,627 participants were recruited for the present study through online sampling using social media and e-mail snowballing, an effective and cost-efficient sampling method (Thornton et al., 2016; Topolovec-Vranic & Natarajan, 2016). Participants provided their informed consent to take part in the study before accessing and completing an online form. Participation was anonymous and voluntary. Two inclusion criteria were used: being adult and responding to all the questions and items in the online form. Participants were randomly split into four sample groups with a similar proportion of respondents of the same gender, age, educational level, and nationality. In addition, a clinical sample consisting of 200 male IPVAW offenders court-mandated to a community-based intervention program was also recruited. The sociodemographic characteristics of the samples are shown in Table 1.

The first and second samples were used to assess the psychometric properties of the PS-IPVAW scale and to cross-validate the latent structure of this measure. The third sample included a larger number of participants in order to examine the measurement invariance across gender and age groups. The fourth sample was used to inspect the measurement invariance between men from the general population and the clinical sample of IPVAW offenders.

Measures

Perceived severity of IPV A W (PS-IPVAW; Gracia et al., 2008). The original PS-IPVAW scale is an 8-item instrument in which each item presents an IPVAW scenario (e.g., “In an argument, a man hits his partner and later apologizes to her”). Participants in this study were asked to rate the level of severity of each scenario on a scale from 0 to 10 (the higher the number, the higher the severity). Previous research has related the PS-IPVAW scale scores to other constructs such as ambivalent sexism, empathy, personal responsibility, attitudes of acceptability of IPVAW, victim-blaming attitudes, and willingness to intervene in IPVAW cases (Gracia et al., 2011; Gracia et al., 2018; Lila et al., 2013; Martín-Fernández, Gracia, & Lila, 2018; Martín-Fernández et al., 2021).

Table 1. Sociodemographic Characteristics of the Samples

<table>
<thead>
<tr>
<th>Sample</th>
<th>Gender</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>299</td>
<td>45.6</td>
<td>311</td>
<td>47.6</td>
<td>540</td>
<td>48.3</td>
<td>200</td>
<td>100.0</td>
<td>200</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>356</td>
<td>54.4</td>
<td>343</td>
<td>52.4</td>
<td>578</td>
<td>51.7</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-24</td>
<td>163</td>
<td>24.9</td>
<td>215</td>
<td>32.9</td>
<td>304</td>
<td>27.2</td>
<td>35</td>
<td>17.5</td>
<td>25</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td>25-34</td>
<td>158</td>
<td>24.1</td>
<td>201</td>
<td>30.7</td>
<td>309</td>
<td>27.6</td>
<td>72</td>
<td>36.0</td>
<td>70</td>
<td>35.0</td>
<td></td>
</tr>
<tr>
<td>35-54</td>
<td>230</td>
<td>35.1</td>
<td>171</td>
<td>26.1</td>
<td>297</td>
<td>26.6</td>
<td>93</td>
<td>46.5</td>
<td>105</td>
<td>52.5</td>
<td></td>
</tr>
<tr>
<td>55+</td>
<td>104</td>
<td>15.9</td>
<td>67</td>
<td>10.2</td>
<td>208</td>
<td>18.6</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>M(SD)</td>
<td>37.16 (14.07)</td>
<td>33.34 (13.47)</td>
<td>36.64 (15.09)</td>
<td>34.37 (11.32)</td>
<td>37.01 (8.52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immigrant</td>
<td>Yes</td>
<td>67</td>
<td>10.2</td>
<td>76</td>
<td>11.6</td>
<td>112</td>
<td>10.0</td>
<td>42</td>
<td>21.0</td>
<td>61</td>
<td>30.5</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>588</td>
<td>89.8</td>
<td>578</td>
<td>88.4</td>
<td>1006</td>
<td>90.0</td>
<td>158</td>
<td>79.0</td>
<td>139</td>
<td>69.5</td>
</tr>
<tr>
<td>Educational level</td>
<td>Primary</td>
<td>44</td>
<td>6.7</td>
<td>34</td>
<td>5.2</td>
<td>76</td>
<td>6.8</td>
<td>19</td>
<td>9.5</td>
<td>16</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td>Secondary</td>
<td>177</td>
<td>27.0</td>
<td>170</td>
<td>26.0</td>
<td>293</td>
<td>26.2</td>
<td>67</td>
<td>33.5</td>
<td>86</td>
<td>43.0</td>
</tr>
<tr>
<td></td>
<td>Upper-secondary</td>
<td>201</td>
<td>30.7</td>
<td>223</td>
<td>34.1</td>
<td>337</td>
<td>30.1</td>
<td>64</td>
<td>32.0</td>
<td>73</td>
<td>36.5</td>
</tr>
<tr>
<td></td>
<td>University</td>
<td>233</td>
<td>35.6</td>
<td>227</td>
<td>34.7</td>
<td>412</td>
<td>36.9</td>
<td>50</td>
<td>25.0</td>
<td>25</td>
<td>12.5</td>
</tr>
</tbody>
</table>

Note. The clinical sample comprised male IPVAW offenders. Upper-secondary education is the stage before university education and includes vocational, technical and employment education.
used in this study, which presents different scenarios tackling three
specific factors corresponding to three types of behavioral intentions:
reporting the incident to the authorities (e.g., “If I found out that a
woman in my neighborhood was frequently beaten by her partner,
but did not want to report it to the authorities, I would call the police”);
personal involvement in the case (e.g., “If I heard a man shouting
violently at his partner in the communal area of my building, I
would intervene to stop the situation”); and ignoring the situation (e.g., “If
a young couple was shouting and insulting each other in the street, I
would ignore them”). Participants responded to the items on a 6-point
Likert-type scale (1 = very unlikely, 6 = extremely likely). Individuals’
scores in the WI-IPVAW scale have been associated with attitudes of
acceptability of this type of violence, victim-blaming attitudes, and
hostile sexism (Gracia et al., 2018). The internal consistency of the
overall scale was also good across all samples (αsample1 = .80, αsample2 = .78, αsample3 = .76, αsample4 = .84).

Hostile sexism (Expósito et al., 1998; Glick & Fiske, 1997).
The short version of the hostile sexism subscale of the original
Ambivalent Sexism Inventory was used in this study (Rollero et al.,
2014). This 6-item version evaluates hostile sexism with a 6-point
Likert-type scale (e.g., “Women seek to gain power by getting control
over men”; 1 = totally disagree, 6 = totally agree). This scale has been
adapted to more than 20 countries, showing adequate psychometric
properties (Glick et al., 2000; Glick et al., 2002). Scores on this scale
have been related to attitudes toward intervention, responsibility
attribution, justification of partner violence, acceptance of myths
about IPV, victim-blaming attitudes, and beliefs about domestic
violence in European countries (Gracia et al., 2020). The internal
consistency of this shortened version in this study was excellent
(αsample1 = .88, αsample2 = .89, αsample3 = .90, αsample4 = .90).

Data Analysis

The following analyses were carried out to examine the
psychometric properties of the PS-IPVAW scale. First, the descriptive
statistics of the items (i.e., mean, standard deviation, skewness,
and kurtosis), as well as the item-test corrected correlation, were
computed for all items using the first sample.

To assess the latent structure of the scale, an exploratory factor analysis (EFA) and a confirmatory factor analysis (CFA) were
conducted using the first and second samples, respectively. Before conducting the EFA, the Kaiser-Meyer-Olkin (KMO) statistic was
obtained, and Bartlett’s sphericity test was performed to assess the
suitability of the dataset to perform this analysis. A parallel analysis
was also carried out to determine the number of factors to extract
(Garrido et al., 2013). An EFA was conducted afterwards using the
robust maximum-likelihood (MLR) estimation method, as this
procedure yields accurate model parameter estimates and standard
errors with either normally or non-normally distributed continuous
The goodness of fit of the factorial solution was evaluated using a
combination of fit indices: comparative fit index (CFI) and Tucker-
Lewis index (TLI) values above .90 and .95 indicate an acceptable
and good fit, respectively; root mean square error of approximation
(RMSEA) values below .05, .08, and .06 indicate poor, mediocre,
and excellent fit, respectively; and standardized root mean square
residual (SRMR) values below .08 indicate good fit (Hu & Bentler,
1999; MacCallum et al., 1996).

To further inspect the properties of the items and their information,
an item response (IRT) model was fitted. Considering the continuous
nature of the data, the continuous response model was applied to the
first sample (Samejima, 1973), estimated using maximum-likelihood
with the expectation-maximization algorithm (Zopluoglu, 2012,
2013). This model allows us to identify which levels of the latent trait
continuum (i.e., perceived severity) are better assessed by each item.

A CFA was then conducted to test the stability of the latent structure
found in the first sample by replicating the same factorial model in the second sample. The model was also fitted using the
MLR estimation method. The goodness of fit was assessed using the
same combination of fit indices (i.e., CFI/TLI ≥ .95, RMSEA ≤ .06, and
SRMR ≤ .08). The internal consistency of the scale was evaluated by
computing Cronbach’s alpha for each sample.

Once the latent structure of the scale had been established, we
examined whether the PS-IPVAW was invariant across gender and
age groups using the third sample. To do so, four age groups were
established, grouping participants aged between 18 and 24, 25 and 34,
35 and 54, and above 55. This age categorization is similar to the one
used to map Spanish internet users (Acebes-Arribas, 2016). A series of
multi-group CFA were carried out to test for configural, metric,
scalar, and strict invariance (Putnick & Bornstein, 2016). Configural
invariance tests whether the same factorial model is applicable to the
different gender and age groups. Metric invariance evaluates whether
the same item loadings can be estimated across groups. Scalar
invariance adds a new constraint to the metric invariance, assessing
whether the values of the item loadings and intercepts can be equal
across groups. Strict invariance adds another constraint to the scalar
invariance, testing whether the residual variances of the items can be
fixed to the same value across gender and age groups. All models
were estimated using MLR.

To assess the fit of the measurement invariance models, we
obtained the changes in CFI (ΔCFI) and RMSEA (ΔRMSEA) indices. We
followed the usual guidelines proposed for measurement invariance
testing (Milfont & Fischer, 2010; Putnick & Bornstein, 2016),
considering that there were no differences between two invariance
models if ΔCFI ≤ |.010| and RMSEAs ≤ |.015| (Chen, 2007; Cheung &
Rensvold, 2002). If at least scalar invariance holds across groups, then participants’ responses can be compared across gender or age
groups through a latent means analysis.

Results

Descriptive Analysis

The first sample was used to conduct the descriptive analysis of
the PS-IPVAW scale. The descriptive statistics of the items showed
that most of the participants rated the severity of the scenarios highly,
with mean values around 9 and standard deviations around 1.30. The
skew and kurtosis statistics also showed this tendency as all items
presented negatively skewed and leptokurtic response distributions,
showing that most of the participants’ ratings were in the upper
values of the response scale. The corrected item-test correlations
were above .40 for all items, indicating a strong relationship between
Exploratory Factor Analysis and Item Response Theory

To test the latent structure of the PS-IPVAW, an EFA was conducted with the first sample. The Kaiser-Meyer-Olkin statistic (KMO = .89) and Bartlett’s sphericity test, $\chi^2(7) = 1,147.6, p < .001$, indicated that the dataset was suitable to conduct this analysis. The parallel analysis showed that only one factor was needed, and hence a one-factor solution was estimated. The goodness of fit of the resulting model was poor, albeit close to an acceptable fit ($\text{CFI} = .89$, $\text{TLI} = .85$, $\text{RMSEA} 90\% \text{CI} = .085 [.071, .101]$, and $\text{SRMR} = .053$). All item loadings presented values above .50, except for item 5, which was .40 (Table 3).

To further inspect the item properties, the continuous response IRT model was fitted afterwards. The discrimination $a$-parameters indicate how informative each item is to assess the latent trait continuum (i.e., perceived severity of IPVAW), whereas the position $b$-parameters show the point on the latent trait continuum where the information reaches its peak. In this case, the $a$-parameters were above 0.80 for most of the items —except for item 5, which was 0.55—and the $b$-parameters were all negative and below -2, suggesting that the items are especially informative for very low levels of perceived severity (Table 3). The item response curves showed that as the perceived severity increases, so does the responses given by the participants to the items (Figure 1).

After the latent structure of the scale had been determined, its internal consistency was examined. Cronbach’s $\alpha$ was .89 in the first sample and .90 in the second sample, indicating a good internal consistency for the PS-IPVAW scale.

Measurement Invariance

The measurement invariance of the PS-IPVAW scale across gender and age groups was assessed using the third sample (Table 4). Configural invariance was first tested across gender as the baseline model. It showed a good fit to the data, supporting that the same relationship between each item and the factor of perceived severity (Figure 2).
factorial model could be applied to men and women. When the item loadings were fixed to the same value across gender, the fit indices of the model improved slightly, albeit exceeding the $\Delta$CFI ≤ |.010| cutoff suggested by Cheung and Rensvold (2002). No differences were found, however, between the configural and metric invariance models when they were compared using the Satorra-Bentler scaled chi-square difference test, $\Delta\chi^2(7) = 0.38, p = .999$, and hence metric invariance was supported. The full scalar invariance model was not supported across gender ($\Delta$CFI = -0.21). A close inspection of the modification indices suggested that freeing the intercept parameters of items 2 and 4 was necessary. A partial scalar invariance model was tested (Table 3), and the fit indices obtained were below $\Delta$CFI ≤ |.010| and $\Delta$RMSEA ≤ |.015|, indicating that partial scalar invariance held. Finally, freeing the intercept of item 4 across groups. The changes in the fit indices of this partial invariance model were below the $\Delta$CFI and $\Delta$RMSEA cutoffs, and hence partial scalar invariance was supported. Fixing the item residual to the same value across age groups did not substantially decrease the fit indices of the model, thus indicating that strict invariance also held.

After establishing strict invariance across gender and age groups, the factor means were compared by a latent means analysis. In this comparison, men presented lower perceived severity levels than women, $Z_{\text{men}} = -0.264, SE = 0.061, p < .001$. No differences were found between participants aged between 18 and 24 and those between 25 and 34, $Z_{25-34} = 0.067, SE = 0.084, p = .426$, and above 55, $Z_{55+} = 0.137, SE = 0.097, p = .160$. Significant differences were found, however, between participants aged between 18 and 24 and those aged between 35 and 54, with the latter group presenting higher perceived severity levels, $Z_{35-54} = 0.189, SE = 0.080, p = .018$. The goodness of fit of both models was good (Table 4).

**Validity Evidence Based on Its Relation to Other Variables**

Participants’ total scores in the PS-IPVAW scale were correlated to the A-IPVAW, VB-IPVAW, WI-IPVAW, and hostile sexism scales in the first two samples (Table 5). Results indicated that participants with higher levels of perceived severity also presented lower levels of acceptability, victim-blaming attitudes, and hostile sexism. We also found that participants who showed higher levels of perceived severity tend to present higher willingness to intervene in cases of IPVAW.

**Measurement Invariance between Men from the General Population and IPVAW Offenders**

The latent structure of the scale was first assessed in a clinical sample of IPVAW offenders. The one-factor model was replicated, and the results showed an acceptable goodness of fit (CFI = .95, TLI = .93, RMSEA [90% CI] = .063 [.037, .088], SRMR = .051). The internal consistency of the PS-IPVAW was also good ($\alpha = .86$). A new multi-group CFA was performed to compare the factor means of men from the general population and IPVAW offenders (Table 6). Configural and metric invariance were supported. Although the metric model presented a better fit than the configural model, these differences were not significant, $\Delta\chi^2(6) = 5.75, p = .548$, thus supporting metric invariance. The scalar invariance model, however, presented a CFI substantially lower than the metric invariance model, and hence we proceeded to check the modification indices and fit three additional models to test for partial metric invariance. The intercept of three items (i.e., items 2, 4, and 6) needed to be freed across the model. However, these differences were not significant, $\Delta\chi^2(21) = 14.92, p = .827$, thus supporting metric invariance. The full scalar model again showed a decrement above the $\Delta$CFI ≤ |.010| cutoff, and a partial scalar invariance model was tested by freeing the intercept of item 4. The changes in the fit indices of this partial invariance model were below the $\Delta$CFI and $\Delta$RMSEA cutoffs, and hence partial scalar invariance was supported. Fixing the item residual to the same value across age groups did not substantially decrease the fit indices of the model, thus indicating that strict invariance also held.

**Validity Evidence Based on Its Relation to Other Variables**

Participants’ total scores in the PS-IPVAW scale were correlated to the A-IPVAW, VB-IPVAW, WI-IPVAW, and hostile sexism scales in the first two samples (Table 5). Results indicated that participants with higher levels of perceived severity also presented lower levels of acceptability, victim-blaming attitudes, and hostile sexism. We also found that participants who showed higher levels of perceived severity tend to present higher willingness to intervene in cases of IPVAW.

**Measurement Invariance between Men from the General Population and IPVAW Offenders**

The latent structure of the scale was first assessed in a clinical sample of IPVAW offenders. The one-factor model was replicated, and the results showed an acceptable goodness of fit (CFI = .95, TLI = .93, RMSEA [90% CI] = .063 [.037, .088], SRMR = .051). The internal consistency of the PS-IPVAW was also good ($\alpha = .86$). A new multi-group CFA was performed to compare the factor means of men from the general population and IPVAW offenders (Table 6). Configural and metric invariance were supported. Although the metric model presented a better fit than the configural model, these differences were not significant, $\Delta\chi^2(6) = 5.75, p = .548$, thus supporting metric invariance. The scalar invariance model, however, presented a CFI substantially lower than the metric invariance model, and hence we proceeded to check the modification indices and fit three additional models to test for partial metric invariance. The intercept of three items (i.e., items 2, 4, and 6) needed to be freed across the IPVAW offenders and male participants from the general population to obtain a CFI below $\Delta$CFI = |.010|. The intercept of these items was lower for the IPVAW offenders ($\tau_2 = 3.17, \tau_4 = 3.16, \tau_6 = 4.34$) than for the male participants ($\tau_2 = 6.17, \tau_4 = 6.39, \tau_6 = 8.75$), indicating that, on average, IPVAW offenders tend to systematically rate these items as less severe than male participants from the general population. The
strict invariance model showed a poor goodness of fit, and hence the item residuals were left unconstrained across groups.

### Table 6. Measurement Invariance across Men from the General Population and IPVAW Offenders (Sample 4 and Clinical Sample)

<table>
<thead>
<tr>
<th>Model</th>
<th>Χ² (df)</th>
<th>CFI</th>
<th>RMSEA (90% CI)</th>
<th>ΔCFI</th>
<th>ΔRMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configural</td>
<td>57.39 (28)</td>
<td>.936</td>
<td>.072 [.054, .091]</td>
<td>0.001</td>
<td>0.007</td>
</tr>
<tr>
<td>Metric</td>
<td>62.55 (34)</td>
<td>.937</td>
<td>.065 [.047, .082]</td>
<td>0.001</td>
<td>0.007</td>
</tr>
<tr>
<td>Scalar</td>
<td>85.39 (41)</td>
<td>.903</td>
<td>.074 [.058, .090]</td>
<td>0.034</td>
<td>0.009</td>
</tr>
<tr>
<td>Partial Scalar 1</td>
<td>78.53 (40)</td>
<td>.915</td>
<td>.069 [.053, .086]</td>
<td>-0.022</td>
<td>0.004</td>
</tr>
<tr>
<td>Partial Scalar 2-4-6</td>
<td>73.64 (39)</td>
<td>.924</td>
<td>.067 [.050, .083]</td>
<td>-0.013</td>
<td>0.002</td>
</tr>
<tr>
<td>Partial Scalar 2-4</td>
<td>71.32 (38)</td>
<td>.928</td>
<td>.066 [.049, .083]</td>
<td>-0.009</td>
<td>0.001</td>
</tr>
<tr>
<td>Partial Scalar 2-4-6</td>
<td>241.45 (45)</td>
<td>.569</td>
<td>.148 [.136, .160]</td>
<td>-3.68</td>
<td>0.082</td>
</tr>
</tbody>
</table>

Note. Χ² = scaled chi-square; df = degrees of freedom; CFI = comparative fit index; RMSEA = root mean squared error of approximation; ΔCFI = change in the CFI; ΔRMSEA = change in the RMSEA.

- Intercepts of item 4 freed across groups.
- Intercepts of items 2 and 4 freed across groups.
- Intercepts of items 2, 4, and 6 freed across groups.

The latent means of both groups were thereafter freely estimated. The resulting model showed an acceptable fit (CFI = .94, TLI = .93, RMSEA [90% CI] = .064 [.045, .080], SRMR = .070), and we found that male participants from the general population presented a higher latent mean in the factor (i.e., perceived severity) than IPVAW offenders, Z = 0.456, SE = 0.109, p = .007. The density distributions of the factor scores in the PS-IPVAW of both groups are displayed in Figure 3. Although the distributions of both groups were centered around moderate levels of perceived severity, the factor scores of the male participants from the general sample presented a leptokurtic form, indicating that most of the participants were around moderate levels of perceived severity. The IPVAW offenders, on the other hand, were more scattered between low and moderate levels of perceived severity, presenting a higher density for negative factor scores on the scale, which in turn suggests that IPVAW offenders are more likely to be found among participants with low levels of perceived severity than men from the general sample.

### Discussion

In this study we aimed to provide further validity evidence for the PS-IPVAW scale, an instrument widely used to assess perceived severity of IPVAW, by assessing its measurement invariance between gender and age groups, and between men from the general population and male IPVAW offenders. In line with previous research, our findings supported that the PS-IPVAW scale presented adequate psychometric properties to assess perceived severity of this type of violence. We found that the scale has a solid factorial structure and a good internal consistency across several samples from the general population. The IRT analyses allowed us to refine this measure by removing the least informative item from the original measure. We also found that this scale was invariant across gender and age groups, which in turn indicated that it could be used to make appropriate comparisons between men and women, and between individuals from different age groups, as they conceptualize and interpret perceived severity in a similar way (Putnick & Bornstein, 2016). Our results show that, on average, men presented lower levels of perceived severity than women, which is consistent with previous research (Webster et al., 2018). As for age, we found a similar U-trend as the one reported by Sánchez-Prada et al. (2020), where younger participants showed lower perceived severity levels than middle aged participants.

Regarding the validity evidence based on the relations of the PS-IPVAW to other relevant constructs, our results are consistent with previous research linking perceived severity to attitudes of acceptability of IPVAW, attitudes of victim blaming, ambivalent sexism, and willingness to intervene in cases of IPVAW (Gracia et al., 2018; Lelaurain et al., 2021; Martín-Fernández, Gracia, & Lila., 2018; Martín-Fernández et al., 2021; Martín-Fernández, Gracia, Marco, et al., 2018). In particular, our findings show that participants with higher levels of perceived severity also tended to present lower levels of acceptability of IPVAW, victim-blaming attitudes, and hostile sexism, and showed higher levels of willingness to intervene in cases of this type of violence.

The measurement invariance analysis comparing the item functioning of a sample of men from the general population with a clinical sample of IPVAW offenders constitutes one of the main strengths of this paper. This analysis showed that three intercepts of the items were substantially lower for IPVAW offenders, which in turn indicated that offenders tend to rate these items as less severe than men from the general population. Our results show that, in a similar way (Putnick & Bornstein, 2016). Our results show that, on average, men presented lower levels of perceived severity than women, which is consistent with previous research (Webster et al., 2018). As for age, we found a similar U-trend as the one reported by Sánchez-Prada et al. (2020), where younger participants showed lower perceived severity levels than middle aged participants.

This study has some limitations. First, the general population samples were collected through online sampling, which presents some tradeoffs: on the one hand, it is a cost-efficient way to gather a large pool of responses; on the other hand, it makes it difficult to verify the sociodemographic characteristics of the participants and to overcome possible self-selection bias, as participants who are more motivated by the aims and the topic of the study may be more willing to take part (Thornton et al., 2016; Topolovec-Vranic & Natarajan, 2016). Second, the cross-sectional study design prevented us from assessing how individuals’ levels of perceived severity might change over time. Third, although the PS-IPVAW is especially informative for individuals with low and moderate levels of perceived severity, the precision of this measure is lower for individuals with high and very high levels on this construct. Finally, the study was conducted in Spain, one of the European Union countries with the lowest rates of IPVAW (Gracia & Merlo, 2016; Martín-Fernández et al., 2019, 2020; Vives-Cases et al., 2011), highlighting the need for more research to examine the psychometric properties of the PS-IPVAW in other socio-cultural contexts.

![Figure 3](image-url)
Perceived severity is a key variable that can lead to a better understanding of intimate partner violence, as it is closely related to its acceptability, personal sense of responsibility, attitudes toward intervention, and public and professional responses to IPVAW (Baldry & Pagliaro, 2014; Gracia et al., 2011; Gracia et al., 2018; Hyman et al., 2009; Lelaurain et al., 2018). The availability of psychometrically sound measures to assess perceived severity is an important step for researchers and practitioners, as it would allow them to make precise and valid evaluations of this construct. The PS-IPVAW can be useful to assess perceived severity at individual level in the general population in large-scale studies where time and space can be an issue, such as demographic surveys (Gracia et al., 2018; Martin-Fernández et al., 2021). It could also be used to monitor attitudes and attitudinal change as a result of IPVAW education or awareness campaigns (Gracia et al., 2020). This instrument can also be useful for high-risk populations, such as IPVAW offenders, either as a screening tool or to study the efficacy of prevention and intervention strategies targeted to these individuals (Bowen et al., 2008; Cunha & Gonçalves, 2015; Santirso et al., 2020). In sum, the PS-IPVAW scale could potentially fill this gap in the literature, as it is able to yield accurate assessments of perceived severity of this type of violence among the general population and among IPVAW offenders.

Conflict of Interest

The authors of this article declare no conflict of interest.

References

Acebes-Arribas, B. (2016). Estudio de medios de comunicación online [Study of online communication media]. IAB Spain.


## Appendix

**PS-IPVAW Scale**

Below are seven scenarios that can occur between a male-female couple. On a scale from 0 to 10, please indicate how serious these scenarios seem to you (the higher the number, the higher the severity of the scenario).

[A continuación se describen siete situaciones que pueden ocurrir entre una pareja de hombre y mujer. Indica en una escala de 0 a 10 hasta qué punto esas situaciones te parecen graves: a mayor número, mayor gravedad].

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Spanish Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A woman has reported her partner for assaulting her, but the man continues to threaten her.</td>
<td>Una mujer ha denunciado a su pareja por haberle agredido, pero el hombre continúa amenazándola</td>
</tr>
<tr>
<td>In an argument, a man hits his partner and later apologizes to her</td>
<td>En una discusión, un hombre pega a su pareja y después le pide perdón</td>
</tr>
<tr>
<td>A woman is frequently beaten by her partner, sometimes causing small injuries and bruises, although she does not want to report these acts.</td>
<td>Una mujer es golpeada frecuentemente por su pareja, causándole a veces pequeñas lesiones y hematomas, aunque no quiere denunciar los hechos</td>
</tr>
<tr>
<td>A couple argues, the man insults the woman and threatens to hit her</td>
<td>Una pareja discute, el hombre insulta a la mujer y amenaza con pegarle</td>
</tr>
<tr>
<td>A woman is constantly belittled and humiliated by her partner</td>
<td>Una mujer es despreciada y humillada continuamente por su pareja</td>
</tr>
<tr>
<td>A woman is repeatedly threatened and insulted by her partner, who sometimes pushes her or hits her.</td>
<td>Una mujer es amenazada e insultada continuamente por su pareja, quien a veces le llega a empujar o golpear</td>
</tr>
<tr>
<td>A couple argues constantly, insulting and threatening each other, often coming to blows.</td>
<td>Una pareja discute continuamente, insultándose y amenazándose mutuamente, llegando a las manos con frecuencia</td>
</tr>
</tbody>
</table>

Note: In brackets the original instructions and items in Spanish.