



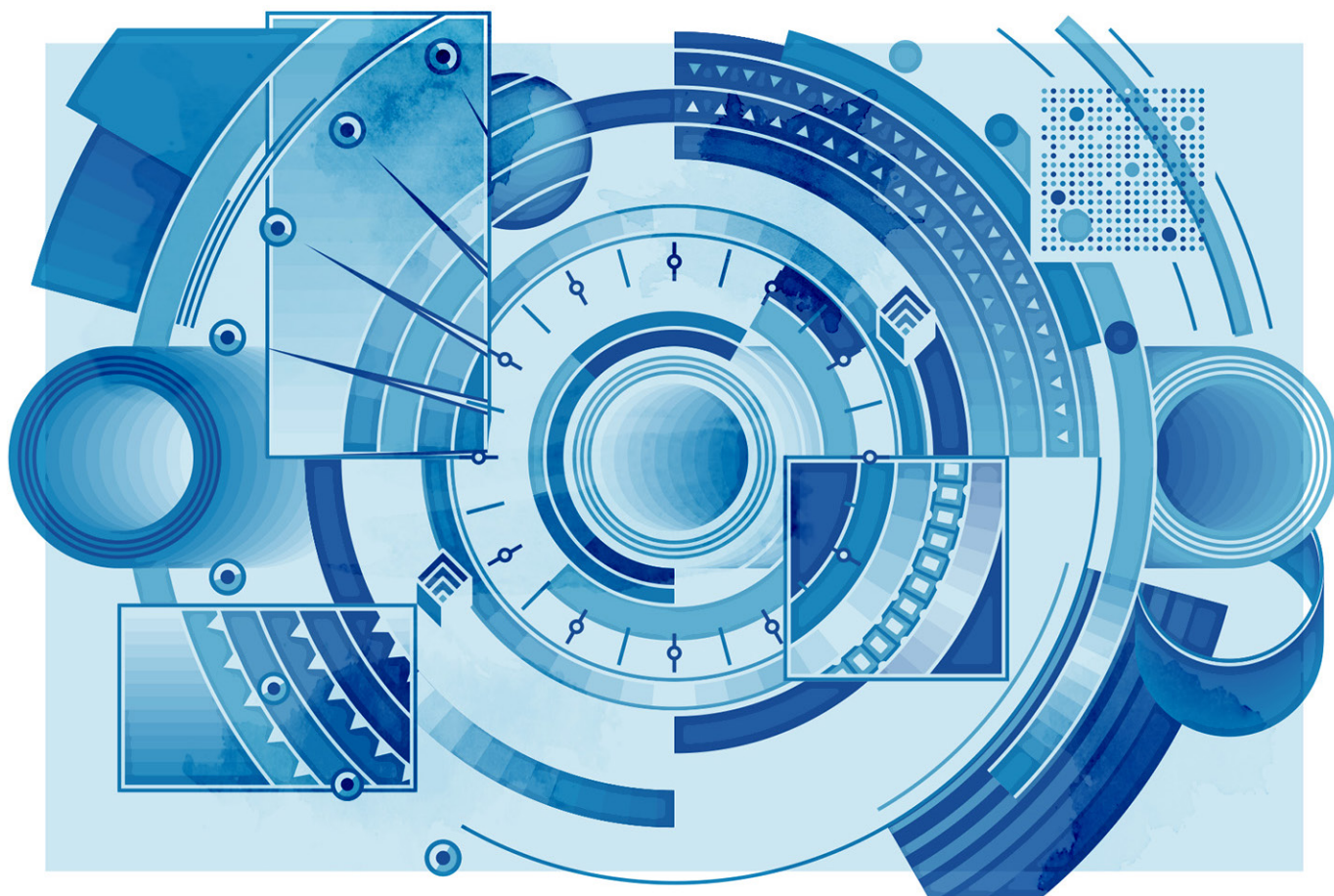
EVALUAR

Revista Evaluar

Laboratorio de Evaluación Psicológica y Educativa
Facultad de Psicología de la Universidad Nacional de Córdoba

2022

VOL 22 - Nº2
ISSN 1667-4545



Argentinian Version of the Fear of COVID-19 Scale (FCV-19S): A Review of Possible Structural Models and its Relationship with Fear of Death

Versión argentina de la Escala de Miedo al COVID-19 (FCV-19S): Una revisión de posibles modelos estructurales y su relación con el miedo a la muerte

Hernán Furman *¹, Mark D. Griffiths², Amir H. Pakpour³, Hugo Simkin^{4, 5}

1 - Carrera de Psicología, Universidad de Flores, CABA, Argentina.

2 - International Gaming Research Unit, Psychology Department, Nottingham Trent University, Nottingham, UK.

3 - Department of Nursing, School of Health and Welfare, Jönköping University, Jönköping, Sweden.

4 - Carrera de Sociología, Universidad de Buenos Aires, CABA, Argentina.

5 - Consejo Nacional de Investigaciones Científicas y Técnicas, CABA, Argentina.

Introduction
Method
Results
Discussion
Conclusion
References

Recibido: 01/05/2022 Revisado: 15/05/2022 Aceptado: 24/05/2022

Abstract

During the pandemic caused by SARS-CoV-2, the Fear of COVID-19 Scale (FCV-19S) has become one of the main psychometric instruments used in research globally. The present study examined the psychometric properties of the Argentinian version of the FCV-19S. More specifically, the study analyzed and compared different possible models in order to determine which internal structure fits better for the Argentinian FCV-19S. The sample comprised 505 adults from Argentina between 18 and 85 years ($M = 45.44$ years; $SD = 17.08$), of both genders (males = 33.5%; females = 66.5%). The results provided empirical support for a bi-factor structure with a general factor ($\alpha = .89$) and two specific factors, that is, emotional response ($\alpha = .87$) and physiological response ($\alpha = .83$). Additionally, the associations between the FCV-19S and the Brief Fear of Death Scale (BFODS) provide evidence of validity related to other variables.

Keywords: *fear of COVID-19, fear of death, FCV-19S, psychometric validation, Argentina*

Resumen

Durante la pandemia provocada por el SARS-CoV-2, la Escala de Miedo al COVID-19 (FCV-19S) se ha convertido en uno de los principales instrumentos psicométricos utilizados en investigación a nivel mundial. El presente estudio examinó las propiedades psicométricas de la versión argentina de la FCV-19S. Más específicamente, el estudio analizó y comparó diferentes modelos posibles con el fin de determinar qué estructura interna encaja mejor para la versión argentina de la FCV-19S. La muestra estuvo compuesta por 505 adultos argentinos de entre 18 y 85 años ($M = 45.44$ años; $DE = 17.08$), de ambos sexos (hombres = 33.5%; mujeres = 66.5%). Los resultados proporcionaron soporte empírico de la existencia de una estructura bifactorial con un factor general ($\alpha = .89$) y dos factores específicos, es decir, respuesta emocional ($\alpha = .87$) y respuesta fisiológica ($\alpha = .83$). Además, las asociaciones entre la FCV-19S y la Escala Abreviada de Miedo a la Muerte (BFODS) proporcionan evidencia de validez relacionada con otras variables.

Palabras clave: *miedo al COVID-19, miedo a la muerte, FCV-19S, validación psicométrica, Argentina*

Author's note: Hernán Furman <https://orcid.org/0000-0001-5597-3174>, Mark D. Griffiths <https://orcid.org/0000-0001-8880-6524>, Amir H. Pakpour <https://orcid.org/0000-0002-8798-5345>, Hugo Simkin <https://orcid.org/0000-0001-7162-146X>. The authors declare no conflict of interest. *Acknowledgment.* The authors wish to thank the participants who voluntarily collaborated in this study. *Funding.* This study was conducted within UBACyT 20020190200244BA, PICT-2019-2019-00408 and UFLO 02OR20, founded by Universidad de Buenos Aires (UBA), Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) and Universidad de Flores (UFLO).

How to cite: Furman, H., Griffiths, M. D., Pakpour, A. H., & Simkin, H. (2022). Argentinian version of the Fear of COVID-19 Scale (FCV-19S): A review of possible structural models and its relationship with fear of death. *Revista Evaluar*, 22(2), 1-13. Retrieved from <https://revistas.unc.edu.ar/index.php/revaluar>

*Correspondence to: Hernán Furman, Universidad de Flores, 5741 Av. Rivadavia, Ciudad Autónoma de Buenos Aires, C1406GLA, Argentina. E-mail: hernan.furman@uflo.edu.ar

Participaron en la edición de este artículo: Abigail Pérez, Gloria Nieve, Juan Cruz Balverdi, Florencia Ruiz, Mónica Serppe, Alicia Molinari, Ricardo Hernández, Jorge Bruera.

Introduction

The pandemic caused by SARS-CoV-2 has generated unprecedented negative consequences and global losses, from the number of deaths to countless psychological, economic and social repercussions (Borio, 2020; Caycho-Rodríguez et al., 2021; Furman et al., 2020). With regard to the psychological repercussions, researchers from all over the world have set themselves the task of studying them and/or developing appropriate approaches for their treatment (Ahorsu et al., 2020; Mamun et al., 2021; Piltch-Loeb et al., 2021). In this area, one of the most studied psychological factors has been the fear of COVID-19 including the fear of contagion, infecting others, or suffering the socioeconomic consequences associated with the pandemic (Caycho-Rodríguez, Tomás et al., 2022; Tzur Bitan et al., 2020). Likewise, as a consequence of the high contagion potential and the high mortality rate of the virus, individuals also experience a greater fear of death (Indacochea-Cáceda et al., 2021; Kumar & Nayar, 2021; Menzies & Menzies, 2020).

The aforementioned fear experiences can cause an increase of mental health problems, either by promoting an increment of risky behaviors (e.g., alcohol and tobacco consumption) or by affecting the prevalence of mental disorders (e.g., post-traumatic stress, and anxiety and depressive disorders, Fitzpatrick et al., 2020; Huarcaya-Victoria et al., 2022; Indacochea-Cáceda et al., 2021). Moreover, due to fear, individuals can think irrationally and react inappropriately, stigmatizing and discriminating against infected individuals (Ahorsu et al., 2020). However, the absence of fear can be equally problematic, fostering the lack of necessary care and hygiene habits during a pandemic (Caycho-Rodríguez, Tomás et al., 2022).

For all these aforementioned reasons, it is

necessary to have valid and reliable instruments to interpret the fear of COVID-19 (Ahorsu et al., 2020; Pakpour & Griffiths, 2020). Such an evaluation makes it possible to analyze its consequences on mental health and explore significant differences among countries (Caycho-Rodríguez et al., 2021; Piqueras et al., 2021). Although a number of psychometric instruments which assess fear and anxiety in relation to COVID-19 are available, the Fear of COVID-19 Scale (FCV-19S; Ahorsu et al., 2020) has been the most widely used scale in the extant literature (Ahorsu et al., 2020; Furman et al., 2020; Tzur Bitan et al., 2020).

The Fear COVID-19 Scale (FCV-19S)

The FCV-19S was originally developed for the Iranian population. According to the developers (Ahorsu et al., 2020), its items emerged from a bibliographic review in which thirty different measures of fear were identified in different populations and circumstances. The authors grouped the most relevant items and discarded those with similar content. The first version consisting of 28 items was analyzed by a panel of experts that included psychologists, virologists, psychiatrists, general physicians and nurses, who suggested discarding another 11 items. The second version comprising 17 items was evaluated by a new panel that included health education specialists, pulmonologists, social psychologists and sociologists, who suggested discarding an additional seven items. Finally, the third version comprising 10 items was administered in a pilot study ($N = 46$; $M = 39.63$; 43% women) and in the general Iranian population ($N = 717$; $M = 31.25$; 42% women). The results suggested discarding another three items, resulting in a final version comprising seven items (Ahorsu et al., 2020).

Since its initial publication in March 2020,

the FCV-19S has been translated, adapted and administrated in over 30 different languages and contexts, including Africa (Elemo et al., 2020; Fawzy El-Bardan, & Lathabhavan, 2021; Giordani et al., 2021), Asia (Al-Shannaq et al., 2021; Alyami et al., 2021; Chang et al., 2022; Chi et al., 2022; Doshi et al., 2021; Dadfar et al., 2021; Masuyama et al., 2020; Nazari et al., 2021; Pang et al., 2022; Reznik et al., 2021; Sakib et al., 2020; Tzur Bitan et al., 2020; Wakashima et al., 2020), Europe (Haktanir et al., 2022; Iversen et al., 2021; Magano et al., 2021; Martínez-Lorca et al., 2020; Nikopoulou et al., 2022; Piqueras et al., 2021; Reznik et al., 2021; Sánchez-Teruel & Robles-Bello, 2021; Satici et al., 2021; Soraci et al., 2020; Stănculescu, 2022; Tsipropoulou et al., 2021), Latin America (Broche-Pérez et al., 2020; Cassiani-Miranda et al., 2022; Cavalheiro & Sticca, 2020; Caycho-Rodríguez, Tomás et al., 2022; Caycho-Rodríguez et al., 2020; de Medeiros et al., 2021; Furman et al., 2020; García-Reyna et al., 2022; Giordani et al., 2020; Huarcaya-Victoria et al., 2022; Mercado-Lara et al., 2022; Moreta-Herrera et al., 2021; Piqueras et al., 2021), North America (Attieh et al., 2022; Perz et al., 2022), and Oceania (Rahman et al., 2021; Winter et al., 2020).

Most of the aforementioned validation studies reported a one-factor structure for FCV-19S (Ahorsu et al., 2020; Al-Shannaq et al., 2021; Alyami et al., 2021; Broche-Pérez et al., 2020; Cassini-Miranda et al., 2022; Cavalheiro & Sticca, 2020; Chang et al., 2022; de Medeiros et al., 2021; Doshi et al., 2021; Elemo et al., 2020; Fawzy El-Bardan & Lathabhavan, 2021; Furman et al., 2020; García-Reyna et al., 2022; Giordani et al., 2020; Giordani et al., 2021; Haktanir et al., 2022; Martínez-Lorca et al., 2020; Mercado-Lara et al., 2022; Nazari et al., 2021; Nikopoulou et al., 2022; Pang et al., 2022; Perz et al., 2022; Piqueras et al., 2021; Sakib et al., 2020; Sánchez-Teruel &

Robles-Bello, 2021; Satici et al., 2021; Soraci et al., 2020; Stănculescu, 2022; Tsipropoulou et al., 2021; Wakashima et al., 2020; Winter et al., 2020). However, there are a few validation studies that reported either a two-factor structure (Caycho-Rodríguez et al., 2021; Caycho-Rodríguez et al., 2020; Magano et al., 2021; Reznik et al., 2021; Tzur Bitan et al., 2020), a higher-order structure (Iversen et al., 2021) or a bi-factor model (Caycho-Rodríguez, Tomás et al., 2022; Chi et al., 2022; Dadfar et al., 2021; Huarcaya-Victoria et al., 2022; Masuyama et al., 2020; Moreta-Herrera et al., 2021). Those who regarded the scale as having more than one factor agreed on the way to group the items, but there were differences regarding the name they assigned to each factor. While some studies have referred to factors in a two-factor structure as *physiological responses* and *emotional responses* to fear of COVID-19 (Caycho-Rodríguez et al., 2021; Masuyama et al., 2020; Reznik et al., 2021), others have named them *physical response* and *fear thoughts* (Chi et al., 2022), *somatic expressions of fear* and *emotional fear reactions* (Huarcaya-Victoria et al., 2022; Tzur Bitan et al., 2020), and *somatic fear* and *cognitive fear* (Iversen et al., 2021). More recently, an alternative bi-factor model that included one general factor and a specific factor called *emotional responses* was proposed (Caycho-Rodríguez, Tomás et al., 2022).

In the Argentinian context, two studies have recently been published reporting the psychometric properties of the FCV-19S. In the first study, the version translated in Spain was used and a solution of two correlated factors was reported (Caycho-Rodríguez et al., 2020). In the other one, a new translation which incorporated Argentinian idiomatic uses and the single-factor model originally proposed was carried out (Furman et al., 2020). Both studies constitute favorable antecedents for the investigation of the fear of COVID-19

in Argentina. However, the psychometric testing was relatively basic. Therefore, the present study analyzed the psychometric properties of the scale in the general Argentinian population in depth. Moreover, it compared the different factorial models proposed in the literature and evaluated both its reliability and the validity criteria according to its internal structure and in relation to other relevant variables.

Method

Participants

The sampling was non-probabilistic and cross-sectional. A total of 505 adults from Argentina between 18 and 85 years ($M = 45.44$ years; $SD = 17.08$; 33.5% males and 66.5% females) voluntarily participated in the study. To determine the sample size, at least 10 to 20 participants were needed per scale item, as suggested in the literature (Kline, 2016; Kyriazos, 2018). With 20 participants per scale item, the minimum number of participants needed would be 140 participants in total. Therefore, the present study's sample size was sufficiently adequate.

Measures

The data were collected through a self-administered survey which included the following measures:

Fear of COVID-19 Scale (FCV-19S). The seven-item FCV-19S (Ahorsu et al., 2020) was used to assess the fear of COVID-19. The items (e.g., *I am very afraid of COVID-19*) are rated on a five-point scale from 1 (*strongly disagreement*) to 5 (*strongly agreement*). The higher the score, the greater the fear of COVID-19. The FCV-19S has

consistently reported acceptable psychometric properties (e.g., Ahorsu et al., 2020; Reznik et al., 2021). For the present study, the validated version adapted to the Argentinian context was used (Furman et al., 2020). The psychometric properties of the scale in the present study are reported in the 'Results' section.

Brief Fear of Death Scale (BFODS). The eight-item BFODS (Collett & Lester, 1969) was used to assess the fear of death. The scale has two dimensions: *fear of own death* (Items 1-4; e.g., *What will it be like to be dead*) and *fear of the death of others* (Items 5-8; e.g., *The loss of a loved one*). Items are rated on a five-point scale from 1 (*Not at all*) to 5 (*A lot*). The first set of items respond to the question of *What degree of worry or anxiety do you experience in relation to the following aspects related to your own death?*, while the second set of items respond to the question of *What degree of worry or anxiety do you experience in relation to the following aspects related to the death of other people?*. For the present study, a version adapted to the Argentinian context was used (Quintero & Simkin, 2017). The scale demonstrated adequate reliability for both the *fear of own death* factor ($\alpha = .73$), *fear of the death of others* ($\alpha = .82$), and the global scale ($\alpha = .82$).

Sociodemographic variables. The survey also included a number of questions regarding socio-demographic information (e.g., age, sex, place of residence).

Procedure and Ethics

In accordance with the specialized literature on this methodology (Jiménez-Ortiz et al., 2020), the survey was hosted on a virtual platform and participants were recruited via the Facebook so-

cial networking platform from November 2020 to February 2021. Participation was voluntary and none of the participants received any financial compensation for taking part in the study. The participants were invited to participate in the study anonymously and they provided informed consent. They were told that the results would be used exclusively for academic-scientific purposes in accordance with National Law 25.326 on the protection of personal data, therefore complying with the codes of ethical conduct established by the National Council for Scientific and Technical Research (CONICET; Res. DN° 2857/06). Ethical approval for the study was provided by the Ethics Committee from the first author's University.

Data Analysis

For the purposes of the present study, both descriptive and inferential statistics were carried out by means of statistical software (SPSS 24 and EQS 8.80). First, the construct validity was analyzed by performing a confirmatory factor analysis with the arbitrary generalized least squares (AGLS) as an estimation method, following the recommendations of previous FCV-19S validation studies (Ahorsu et al., 2020; Chi et al., 2022; Iversen et al., 2021; Masuyama et al., 2020). The χ^2 (Satorra-Bentler), the incremental fit index (IFI), the non-normed fit index (NNFI) and the compar-

ative fit index (CFI) were used while considering values greater than .90 as a cut-off point (Brown, 2015); and the standardized root mean squared residual (SRMR) was used while considering values lower than .1 as indicators of a reasonable adjustment (Harrington, 2009; Schreiber et al., 2006). Subsequently, the parsimony of the five models proposed in the literature was analyzed considering the Akaike information criterion (AIC) and the consistent Akaike information criterion (CAIC), with high values indicating a worse fit (Akaike, 1974; Anderson et al., 1998). The reliability of the scale was subsequently evaluated according to the internal consistency analysis from Cronbach's alpha since it is most widely used in the specialized literature in this line of research (Ahorsu et al., 2020; Chi et al., 2022; Iversen et al., 2021; Masuyama et al., 2020). Values greater than .70 were considered adequate (Brown, 2002). Finally, a correlation analysis was carried out to test the scale in relation to other variables.

Results

Confirmatory Factor Analysis of the Fear Covid-19 Scale

As proposed by specialized studies on this methodology (Peña-Contreras et al., 2020), a confirmatory factor analysis was carried out first, which evaluated the five models presented in the

Table 1
Confirmatory factor analysis of the Fear of COVID-19 Scale.

	$\chi^2_{(df)}$	NNFI	CFI	IFI	SRMR
Unidimensional model	87.544 ₍₁₃₎	.961	.976	.976	.121
Two correlated factors model	66.250 ₍₁₁₎	.966	.982	.982	.095
Higher-order model	66.250 ₍₉₎	.957	.981	.982	.095
Bi-factor model with one specific factor	20.575 ₍₈₎	.989	.996	.996	.035
Bi-factor model with two specific factors	20.575 ₍₄₎	.972	.995	.995	.035

Note. NNFI = non-normed fit index; CFI = comparative fit index; IFI = incremental fit index; SRMR = standardized root mean squared residual.

Table 2
Parsimony indices of the different models.

Model	AIC	CAIC Model	CAIC Independent	Δ CAIC
Unidimensional model	61.544	-6.376	2956.722	2963.098
Two correlated factors model	44.250	-13.220	2956.722	2969.942
Higher-order model	48.250	1.229	2956.722	2955.493
Bi-factor model with one specific factor	4.575	-37.221	2956.722	2993.943
Bi-factor model with two specific factors	12.575	-8.323	2956.722	2965.045

Note. AIC = Akaike information criterion; CAIC = consistent Akaike information criterion.

literature according to the goodness of fit indices (Table 1). Table 1 demonstrates that there were acceptable properties and supports a bi-factor model, especially regarding the standardized root mean squared residual (SRMR) and both for a bi-factor model with one specific factor and for a bi-factor model with two specific factors.

On the other hand, the parsimony of the five models proposed in the literature was analyzed (Table 2). Here the bi-factor model with two specific factors obtained the best fit.

According to the results, the bi-factor model which includes two specific factors (Figure 1) is the one that best fitted the data.

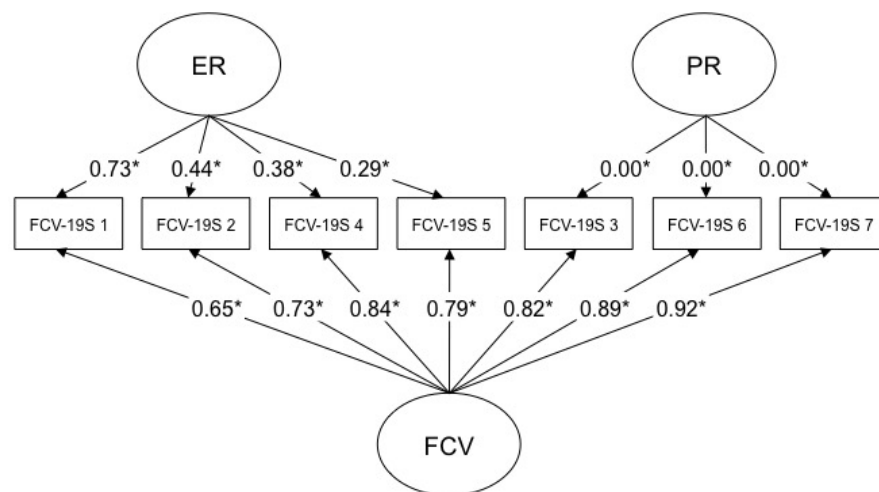
Internal Consistency of the Fear of COVID-19 Scale

Internal consistency was evaluated using Cronbach's alpha (Viladrich et al., 2017), and very good values were obtained for the global scale ($\alpha = .89$), the *physiological response* factor ($\alpha = .83$), and the *emotional response* factor ($\alpha = .87$).

Evidence of Validity in Relation to Other Variables

Regarding validity in relation to other variables, a correlational analysis was performed be-

Figure 1
Structural model of the Fear COVID-19 Scale.



Note. FCV = fear of COVID-19 (general factor); PR = physiological response; ER = emotional response.

Table 3
Correlations between the Fear of COVID-19 Scale and the Brief Fear of Death Scale.

	FOD	FDOT	FD (GF)
FCV	.379**	.271**	.376**
PR	.228**	.114*	.199**
ER	.420**	.328**	.432**

Notes. FCV = fear of COVID-19 (general factor); PR = physiological response; ER = emotional response; FOD = fear of own death; FDOT = fear of the death of others; FD (GF) = fear of death (general factor). **Significant at $p < .01$.

tween the Fear of COVID-19 Scale (FCV-19S) and the Brief Fear of Death Scale (BFODS), taking into account the theoretical relevance of the relationship between both constructs. Positive relationships were obtained between the different dimensions of both scales (Table 3).

Discussion

The present study psychometrically evaluated five structural models for the Fear of COVID-19 Scale (FCV-19S; Ahorsu et al., 2020) which was previously translated, adapted and validated to the Argentinian context (Furman et al., 2020). In this regard, although the original study and most previous studies that have validated the FCV-19S support a one-factor unidimensional solution, few studies have compared different possible models. The minority of studies which have done so tend to opt for other solutions such as those reporting a bi-factor structure (Caycho-Rodríguez, Tomás et al., 2022; Chi et al., 2022; Dadfar et al., 2021; Huarcaya-Victoria et al., 2022; Masuyama et al., 2020; Moreta-Herrera et al., 2021).

In line with this minority of studies, the results of the present study suggest that the bi-factor model that considers a general factor with two specific factors is the one that best fitted the data

collected among Argentinian adults. It is worth mentioning that applications of the bi-factor model have recently proliferated because it allows evaluating the degree to which a measure produces a unique total score and, in a related way, the degree to which the subscales that represent theoretically different constructs produce reliable scores after considering the general factor (Bonifay et al., 2017). Therefore, utilizing a bi-factor model with the FCV-19S offers the possibility of obtaining both a general measure of fear of COVID-19 (composed of seven items), as well as two additional measures, that is, emotional responses (items 1, 2, 4 and 5) and physiological responses (items 3, 6 and 7) that a given population experiences from fear of the virus.

The present study also demonstrated adequate values for the FCV-19S in terms of its internal consistency ($\alpha \geq .89$). This is in line with internal consistencies found by previous research in the same language: Cuba ($\alpha \geq .87$; Broche-Pérez et al., 2020), Mexico ($\alpha \geq .902$; García-Reyna et al., 2022), and Peru ($\alpha \geq .83$; Huarcaya-Victoria et al., 2022). However, it should be noted that one-third of these studies have postulated a bi-factor structure. Furthermore, according to the literature review of Huarcaya-Victoria et al. (2022), similar values have also been observed in the international context. According to these authors, this indicates that the measurements of the different versions of the FCV-19S are inherently stable.

Finally, the correlational analysis with the Brief Fear of Death Scale (BFODS) yielded positive associations. This makes it possible to associate a greater fear of COVID-19 with a greater fear of death, which corroborates the concurrent validity of the FCV-19S in relation to other variables. In this regard, it is worth mentioning that the study of both constructs has been previously addressed and indicates the same relationship between them (Kumar & Nayar, 2020; Menzies &

Menzies, 2020; Pradhan et al., 2022).

Limitations

First, the sample size was a relatively small sample of the general population. It would be of great interest to expand the sample and also work with specific populations, such as the elderly or patients with pre-existing high-risk diseases. Second, participants were surveyed online and, even though this provided a potentially wider reach, the control of variables was less accurate. Third, the study employed a cross-sectional design. Therefore, relationships between the study variables do not provide causal implications. Future research should utilize longitudinal designs in order to assess causal relationships between antecedents, outcomes, and fear of COVID-19 (Caycho-Rodríguez, Tomás et al., 2022). Fourth, as most of the participants in this study were women (66.5%), this may affect the generalizability of our results. Fifth and last, the use of self-report measures may also generate biases in the responses, whether due to social desirability, memory recall effects and/or other method biases. Future studies should use other types of measures (e.g., in-depth interviews).

Conclusion

The present study not only demonstrated that the FCV-19S provides valid and reliable interpretations for the evaluation of fear of COVID-19 in the Argentinian context, but also provides empirical support for a bi-factor model with one general factor and two specific factors for the structure of the scale. As noted, having valid and reliable tools for measuring the psychological consequences of COVID-19 outbreak, including the degree of fear perceived by individuals, constitutes an element

of great value in order to be able to address this delicate issue in all of its complexity.

References

- Ahorsu, D. K., Lin, C-Y., Imani, V., Saffari, M., Griffiths, M. D., & Pakpour, A. H. (2020). The Fear of COVID-19 Scale: Development and initial validation. *International Journal of Mental Health and Addiction*, 20(3), 1537-1545. doi: [10.1007/s11469-020-00270-8](https://doi.org/10.1007/s11469-020-00270-8)
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723. doi: [10.1109/TAC.1974.1100705](https://doi.org/10.1109/TAC.1974.1100705)
- Al-Shannaq, Y., Mohammad, A. A., & Khader, Y. (2021). Psychometric properties of the Arabic Version of the Fear of COVID-19 Scale (FCV-19S) among Jordanian adults. *International Journal of Mental Health and Addiction*, 18, 1-14. doi: [10.1007/s11469-021-00574-3](https://doi.org/10.1007/s11469-021-00574-3)
- Alyami, M., Henning, M., Krägeloh, C. U., & Alyami, H. (2021). Psychometric evaluation of the Arabic Version of the Fear of COVID-19 Scale. *International Journal of Mental Health and Addiction*, 19(6), 2219-2232. doi: [10.1007/s11469-020-00316-x](https://doi.org/10.1007/s11469-020-00316-x)
- Anderson, D. R., Burnham, K. P., & White, G. C. (1998). Comparison of Akaike information criterion and consistent Akaike information criterion for model selection and statistical inference from capture-recapture studies. *Journal of Applied Statistics* 25(2), 263-282. doi: [10.1080/02664769823250](https://doi.org/10.1080/02664769823250)
- Attieh, R., Koffi, K., Touré, M., Parr-Labbé, É., Pakpour, A. H., & Poder, T. G. (2022). Validation of the Canadian French version of the fear of COVID-19 scale in the general population of Quebec. *Brain and Behavior*, 12(5), 32550. doi: [10.1002/brb3.2550](https://doi.org/10.1002/brb3.2550)
- Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science*, 5(1), 184-186. doi: [10.1177/2167702616657069](https://doi.org/10.1177/2167702616657069)

- Borio, C. (2020). The Covid-19 economic crisis: Dangerously unique. *Business Economics*, 55(4), 181-190. doi: [10.1057/s11369-020-00184-2](https://doi.org/10.1057/s11369-020-00184-2)
- Broche-Pérez, Y., Fernández-Fleites, Z., Jiménez-Puig, E., Fernández-Castillo, E., & Rodríguez-Martin, B. C. (2020). Gender and Fear of COVID-19 in a Cuban population sample. *International Journal of Mental Health and Addiction*, 20(1), 83-91. doi: [10.1007/s11469-020-00343-8](https://doi.org/10.1007/s11469-020-00343-8)
- Brown, J. D. (2002). The Cronbach alpha reliability estimate. *Shiken: JALT Testing and Evaluation SIG Newsletter*, 6(1), 17-18. Recuperado de https://hosted.jalt.org/test/bro_13.htm
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research (Second Edition)*. Guilford Press.
- Cassiani-Miranda, C. A., Tirado-Otálvaro, A. F., & Campo-Arias, A. (2022). Adaptation and psychometric evaluation of the Fear of COVID-19 Scale in the general Colombian population. *Death Studies*, 46(3), 595-602. doi: [10.1080/07481187.2021.1874572](https://doi.org/10.1080/07481187.2021.1874572)
- Cavalheiro, F. R. S., & Sticca, M. G. (2020). Adaptation and validation of the Brazilian version of the Fear of COVID-19 Scale. *International Journal of Mental Health and Addiction*, 20(2), 921-929. doi: [10.1007/s11469-020-00415-9](https://doi.org/10.1007/s11469-020-00415-9)
- Caycho-Rodríguez, T., Tomás, J. M., Barboza-Palomino, M., Ventura-León, J., Gallegos, M., Reyes-Bossio, M. & Vilca, L. W. (2022). Assessment of Fear of COVID-19 in older adults: Validation of the Fear of COVID-19 Scale. *International Journal of Mental Health and Addiction*, 20(2), 1231-1245. doi: [10.1007/s11469-020-00438-2](https://doi.org/10.1007/s11469-020-00438-2)
- Caycho-Rodríguez, T., Valencia, P. D., Vilca, L. W., Cervigni, M., Gallegos, M., Martino, P., ... & Burgos-Videla, C. (2021). Cross-cultural measurement invariance of the fear of COVID-19 scale in seven Latin American countries. *Death Studies*, 1-15. doi: [10.1080/07481187.2021.1879318](https://doi.org/10.1080/07481187.2021.1879318)
- Caycho-Rodríguez, T., Vilca, L. W., Cervigni, M., Gallegos, M., Martino, P., Portillo, N., ... & Burgos Videla, C. (2020). Fear of COVID-19 Scale: Validity, reliability and factorial invariance in Argentina's general population. *Death Studies*, 46(3), 543-552. doi: [10.1080/07481187.2020.1836071](https://doi.org/10.1080/07481187.2020.1836071)
- Chang, K-C., Hou, W-L., Pakpour, A. H., Lin, C-Y., & Griffiths, M. D. (2022). Psychometric testing of Three COVID-19-related scales among people with mental illness. *International Journal of Mental Health and Addiction*, 20(1), 324-336. doi: [10.1007/s11469-020-00361-6](https://doi.org/10.1007/s11469-020-00361-6)
- Chi, X., Chen, S., Chen, Y., Chen, D., Yu, Q., Guo, T., ... & Zou, L. (2022). Psychometric evaluation of the Fear of COVID-19 Scale among Chinese population. *International Journal of Mental Health and Addiction*, 20(2), 1273-1288. doi: [10.1007/s11469-020-00441-7](https://doi.org/10.1007/s11469-020-00441-7)
- Collett, L-J., & Lester, D. (1969). The fear of death and the fear of dying. *Journal of Psychology*, 72(2), 179-181. doi: [10.1080/00223980.1969.10543496](https://doi.org/10.1080/00223980.1969.10543496)
- Dadfar, M., Mahoghegh, F., & Eslami, M. (2021). The Fear of COVID-19 Scale (FCV-19S): A study of Iranian university students. *Mankind Quarterly*, 61(3), 707-722. doi: [10.46469/mq.2021.61.3.19](https://doi.org/10.46469/mq.2021.61.3.19)
- De Medeiros, E. D., Reis, L. M., Guimarães, C. L. C., Da Silva, P. G. N., Monteiro, R. P., De Holanda Coelho, G. L., ... & De França, L. L. A. (2021). Psychometric properties of the Brazilian version of the Fear of COVID-19 Scale (FCV-19S). *Current Psychology*. doi: [10.1007/s12144-021-01476-2](https://doi.org/10.1007/s12144-021-01476-2)
- Doshi, D., Karunakar, P., Sukhabogi, J. R., Prasanna, J. S., & Mahajan, S. V. (2021). Assessing coronavirus fear in Indian population using the Fear of COVID-19 Scale. *International Journal of Mental Health and Addiction*, 19(6), 2383-2391. doi: [10.1007/s11469-020-00332-x](https://doi.org/10.1007/s11469-020-00332-x)
- Elemo, A. S., Satici, S. A., & Griffiths, M. D. (2020). The Fear of COVID-19 Scale: Psychometric Properties of the Ethiopian Amharic Version. *International Journal of Mental Health and Addiction*. [Advance online publication]. doi: [10.1007/s11469-020-00448-0](https://doi.org/10.1007/s11469-020-00448-0)
- Fawzy El-Bardan, M., & Lathabhavan, R. (2021). Fear of COVID-19 Scale: Psychometric properties, reliability

- ty and validity in Egyptian population. *Diabetes and Metabolic Syndrome: Clinical Research & Reviews*, 15(4), 102153. doi: [10.1016/j.dsx.2021.05.026](https://doi.org/10.1016/j.dsx.2021.05.026)
- Fitzpatrick, K. M., Harris, C., & Drawve, G. (2020). Fear of COVID-19 and the mental health consequences in America. *Psychological Trauma: Theory, Research, Practice, and Policy*, 12(S1), 17-21. doi: [10.1037/tra0000924](https://doi.org/10.1037/tra0000924)
- Furman, H., Griffiths, M. D., Pakpour, A. H., & Simkin, H. (2020). Evidencias de validez de la Escala de Miedo al COVID-19 (FCV-19S) en el contexto argentino. *Psocial Revista de Investigación en Psicología Social*, 6(2), 99-110. Recuperado de <https://publicaciones.sociales.uba.ar/index.php/psicologiasocial/issue/view/572>
- García-Reyna, B., Castillo-García, G. D., Barbosa-Camacho, F. J., Cervantes-Cardona, G. A., Cervantes-Pérez, E., Torres-Mendoza, B. M., ... & Cervantes-Guevara, G. (2022). Fear of COVID-19 Scale for Hospital Staff in regional hospitals in Mexico: A brief report. *International Journal of Mental Health and Addiction*, 20(2), 895-906. doi: [10.1007/s11469-020-00413-x](https://doi.org/10.1007/s11469-020-00413-x)
- Giordani, R. C., Giolo, S. R., Muhl, C., & Zanoni-Da Silva, M. (2021). Psychometric evaluation of the Portuguese version of the FCV-19 Scale and assessment of fear of COVID-19 in a Southern Brazilian population. *Journal of Human Behavior in the Social Environment*, 31(1-4), 145-153. doi: [10.1080/10911359.2020.1854142](https://doi.org/10.1080/10911359.2020.1854142)
- Giordani, R. C. F., Zanoni Da Silva, M., Muhl, C., & Giolo, S. R. (2020). Fear of COVID-19 scale: Assessing fear of the coronavirus pandemic in Brazil. *Journal of Health Psychology*, 27(4), 901-912. doi: [10.1177/1359105320982035](https://doi.org/10.1177/1359105320982035)
- Haktanir, A., Seki, T., & Dilmaç, B. (2022). Adaptation and evaluation of Turkish version of the Fear of COVID-19 Scale. *Death Studies*, 46(3), 719-727. doi: [10.1080/07481187.2020.1773026](https://doi.org/10.1080/07481187.2020.1773026)
- Harrington, D. (2009). *Confirmatory Factor Analysis*. Oxford University Press.
- Huarcaya-Victoria, J., Villarreal-Zegarra, D., Podestà, A., & Luna-Cuadros, M. A. (2022). Psychometric Properties of a Spanish version of the Fear of COVID-19 Scale in general population of Lima, Peru. *International Journal of Mental Health and Addiction*, 20(1), 249-262. doi: [10.1007/s11469-020-00354-5](https://doi.org/10.1007/s11469-020-00354-5)
- Indacochea-Cáceda, S., Torres-Malca, J. R., Vera-Ponce, V. J., & de La Cruz-Vargas, J. A. (2021). El miedo y la ansiedad ante la muerte en los médicos de varios países latinoamericanos durante la pandemia del Covid-19. *Revista Uruguaya de Medicina Interna*, 6(3), 49-58. doi: [10.26445/06.03.5](https://doi.org/10.26445/06.03.5)
- Iversen, M. M., Norekvål, T. M., Oterhals, K., Fadnes, L. T., Mæland, S., Pakpour, A. H., & Breivik, K. (2021). Psychometric properties of the Norwegian version of the Fear of COVID-19 Scale. *International Journal of Mental Health and Addiction*, 20(3), 1446-1464. <https://doi.org/10.1007/s11469-020-00454-2>
- Jiménez-Ortiz, D., Sánchez-Cardona, I., & Pérez-Pedrogo, C. (2020). Psychometric properties of the OBCS body shame scale in a sample of female residents in Puerto Rico. *Revista Evaluar*, 20(3), 20-33. doi: [10.35670/1667-4545.v20.n3.31706](https://doi.org/10.35670/1667-4545.v20.n3.31706)
- Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). The Guilford Press.
- Kumar, A., & Nayar, K. R. (2021). COVID 19 and its mental health consequences. *Journal of Mental Health*, 30(1), 1-2. doi: [10.1080/09638237.2020.1757052](https://doi.org/10.1080/09638237.2020.1757052)
- Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, 9(8), 2207-2230. doi: [10.4236/psych.2018.98126](https://doi.org/10.4236/psych.2018.98126)
- Magano, J., Vidal, D. G., Pedrosa e Sousa, H. F., Dinis, M. A. P., & Leite, Â. (2021). Validation and psychometric properties of the Portuguese version of the Coronavirus Anxiety Scale (CAS) and fear of COVID-19 Scale (FCV-19S) and associations with travel, tourism and hospitality. *International Journal of Environmental Research and Public Health*, 18(2), 427. doi: [10.3390/ijerph18020427](https://doi.org/10.3390/ijerph18020427)
- Mamun, M. A., Sakib, N., Gozal, D., Bhuiyan, A. I., Hossain,

- S., Bodrud-Doza, M., ... & Pakpour, A. H. (2021). The COVID-19 pandemic and serious psychological consequences in Bangladesh: A population-based nationwide study. *Journal of Affective Disorders*, 279, 462-472. doi: [10.1016/j.jad.2020.10.036](https://doi.org/10.1016/j.jad.2020.10.036)
- Martínez-Lorca, M., Martínez-Lorca, A., Criado-Álvarez, J. J., Cabañas-Armesilla, M. D., & Latorre, J. M. (2020). The Fear of COVID-19 Scale: Validation in Spanish university students. *Psychiatry Research*, 293, 113350. doi: [10.1016/j.psychres.2020.113350](https://doi.org/10.1016/j.psychres.2020.113350)
- Masuyama, A., Shinkawa, H., & Kubo, T. (2020). Validation and psychometric properties of the Japanese version of the Fear of COVID-19 Scale among adolescents. *International Journal of Mental Health and Addiction*, 20(1), 387-397. doi: [10.1007/s11469-020-00368-z](https://doi.org/10.1007/s11469-020-00368-z)
- Menzies, R. E., & Menzies, R. G. (2020). Death anxiety in the time of COVID-19: Theoretical explanations and clinical implications. *The Cognitive Behaviour Therapist*, 13. doi: [10.1017/S1754470X20000215](https://doi.org/10.1017/S1754470X20000215)
- Mercado-Lara, M. F., Campo-Arias, A., & Monterrosa-Castro, Á. (2022). Validity and reliability of the Spanish version of Fear of COVID-19 Scale in Colombian physicians. *International Journal of Mental Health and Addiction*, 20(2), 1122-1129. doi: [10.1007/s11469-020-00430-w](https://doi.org/10.1007/s11469-020-00430-w)
- Moreta-Herrera, R., López-Calle, C., Caycho-Rodríguez, T., Cabezas Guerra, C., Gallegos, M., Cervigni, M., ... & Calandra, M. (2021). Is it possible to find a bifactor structure in the Fear of COVID-19 Scale (FCV-19S)? Psychometric evidence in an Ecuadorian sample. *Death Studies*, 1-11. doi: [10.1080/07481187.2021.1914240](https://doi.org/10.1080/07481187.2021.1914240)
- Nazari, N., Safitri, S., Usak, M., Arabmarkadeh, A., & Griffiths, M. D. (2021). Psychometric validation of the Indonesian version of the Fear of COVID-19 Scale: Personality traits predict the fear of COVID-19. *International Journal of Mental Health and Addiction*, 1-17. [Advance online publication]. doi: [10.1007/s11469-021-00593-0](https://doi.org/10.1007/s11469-021-00593-0)
- Nikopoulou, V. A., Holeva, V., Parlapani, E., Karamouzi, P., Voitsidis, P., Porfyri, G. N., ... & Diakogiannis, I. (2022). Mental Health Screening for COVID-19: A Proposed cutoff score for the Greek Version of the Fear of COVID-19 Scale (FCV-19S). *International Journal of Mental Health and Addiction*, 20(2), 907-920. doi: [10.1007/s11469-020-00414-w](https://doi.org/10.1007/s11469-020-00414-w)
- Pakpour, A. H., & Griffiths, M. D. (2020). The fear of COVID-19 and its role in preventive behaviors. *Journal of Concurrent Disorders*, 2(1), 58-63. doi: [10.54127/wcic8036](https://doi.org/10.54127/wcic8036)
- Pang, N. T. P., Kamu, A., Hambali, N. L. B., Mun, H. C., Kassim, M. A., Mohamed, N. H., ... & Jeffree, M. S. (2022). Malay version of the Fear of COVID-19 Scale: Validity and reliability. *International Journal of Mental Health and Addiction*, 20(1), 263-272. doi: [10.1007/s11469-020-00355-4](https://doi.org/10.1007/s11469-020-00355-4)
- Peña-Contreras, E. K., Lima-Castro, S. E., Arias-Medina, W. P., Bueno-Pacheco, G. A., Aguilar-Sizer, M. E., & Cabrera-Vélez, M. M. (2020). Propiedades psicométricas de la Escala Breve de Resiliencia (BRS) en el contexto ecuatoriano. *Revista Evaluar*, 20(3), 83-98. doi: [10.35670/1667-4545.v20.n3.31715](https://doi.org/10.35670/1667-4545.v20.n3.31715)
- Perz, C. A., Lang, B. A., & Harrington, R. (2022). Validation of the Fear of COVID-19 Scale in a US college sample. *International Journal of Mental Health and Addiction*, 20(1), 273-283. doi: [10.1007/s11469-020-00356-3](https://doi.org/10.1007/s11469-020-00356-3)
- Piltch-Loeb, R., Merdjanoff, A., & Meltzer, G. (2021). Anticipated mental health consequences of COVID-19 in a nationally-representative sample: Context, coverage, and economic consequences. *Preventive Medicine*, 145, 106441. doi: [10.1016/j.ypmed.2021.106441](https://doi.org/10.1016/j.ypmed.2021.106441)
- Piqueras, J. A., Gomez-Gomez, M., Marzo, J. C., Gomez-Mir, P., Falco, R., Valenzuela, B., ... & Group P-SC-19 study. (2021). Validation of the Spanish version of Fear of COVID-19 Scale: Its association with acute stress and coping. *International Journal of Mental Health and Addiction*. doi: [10.1007/s11469-021-00615-x](https://doi.org/10.1007/s11469-021-00615-x)
- Pradhan, M., Chettri, A., & Maheshwari, S. (2022). Fear of

- death in the shadow of COVID-19: The mediating role of perceived stress in the relationship between neuroticism and death anxiety. *Death Studies*, 46(5), 1106-1110. doi: [10.1080/07481187.2020.1833384](https://doi.org/10.1080/07481187.2020.1833384)
- Quintero, S., & Simkin, H. (2017). Validación de la Escala de Miedo a la Muerte Abreviada al contexto argentino. *Psocial Revista de Investigación en Psicología Social*, 3(1), 53-59. Recuperado de <https://publicaciones.sociales.uba.ar/index.php/psicologiasocial/article/view/2340/1985>
- Rahman, M. A., Salehin, M., Islam, S. M. S., Alif, S. M., Sultana, F., Sharif, A., & Cross, W. M. (2021). Reliability of the tools used to examine psychological distress, fear of COVID-19 and coping amongst migrants and non-migrants in Australia. *International Journal of Mental Health Nursing*, 30(3), 747-758. doi: [10.1111/inm.12845](https://doi.org/10.1111/inm.12845)
- Reznik, A., Gritsenko, V., Konstantinov, V., Khamenka, N., & Isralowitz, R. (2021). COVID-19 Fear in eastern Europe: Validation of the Fear of COVID-19 Scale. *International Journal of Mental Health and Addiction*, 19(5), 1903-1908. doi: [10.1007/s11469-020-00283-3](https://doi.org/10.1007/s11469-020-00283-3)
- Sakib, N., Bhuiyan, A. K. M. I., Hossain, S., Al Mamun, F., Hosen, I., Abdullah, A. H., ... & Mamun, M. A. (2020). Psychometric validation of the Bangla Fear of COVID-19 Scale: Confirmatory factor analysis and Rasch Analysis. *International Journal of Mental Health and Addiction*. doi: [10.1007/s11469-020-00289-x](https://doi.org/10.1007/s11469-020-00289-x)
- Sánchez-Teruel, D., & Robles-Bello, M. A. (2021). Escala de miedo al COVID-19 (FCV-19S): Propiedades psicométricas e invariabilidad de la medida en la versión española. *Actas Españolas de Psiquiatría*, 49(3), 96-105. Recuperado de <https://deustofamilypsych.deusto.es/wp-content/uploads/2021/10/Escala-de-miedo-al-COVID-19-Espana.pdf>
- Satici, B., Gocet-Tekin, E., Deniz, M. E., & Satici, S. A. (2021). Adaptation of the Fear of COVID-19 Scale: Its association with psychological distress and life satisfaction in Turkey. *International Journal of Mental Health and Addiction*, 19(6), 1980-1988. doi: [10.1007/s11469-020-00294-0](https://doi.org/10.1007/s11469-020-00294-0)
- 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(6), 323-338. doi: [10.3200/JOER.99.6.323-338](https://doi.org/10.3200/JOER.99.6.323-338)
- Soraci, P., Ferrari, A., Abbiati, F. A., Del Fante, E., De Pace, R., Urso, A., & Griffiths, M. D. (2020). Validation and psychometric evaluation of the Italian version of the Fear of COVID-19 Scale. *International Journal of Mental Health and Addiction*. doi: [10.1007/s11469-020-00277-1](https://doi.org/10.1007/s11469-020-00277-1)
- Stănculescu, E. (2022). Fear of COVID-19 in Romania: Validation of the Romanian version of the Fear of COVID-19 Scale using graded response model analysis. *International Journal of Mental Health and Addiction*, 20(2), 1094-1109. doi: [10.1007/s11469-020-00428-4](https://doi.org/10.1007/s11469-020-00428-4)
- Tsipropoulou, V., Nikopoulou, V. A., Holeva, V., Nasika, Z., Diakogiannis, I., Sakka, S., ... & Parlapani, E. (2021). Psychometric properties of the Greek Version of FCV-19S. *International Journal of Mental Health and Addiction*, 19(6), 2279-2288. doi: [10.1007/s11469-020-00319-8](https://doi.org/10.1007/s11469-020-00319-8)
- Tzur Bitan, D., Grossman-Giron, A., Bloch, Y., Mayer, Y., Shiffman, N., & Mendlovic, S. (2020). Fear of COVID-19 Scale: Psychometric characteristics, reliability and validity in the Israeli population. *Psychiatry Research*, 289, 113100. doi: [10.1016/j.psychres.2020.113100](https://doi.org/10.1016/j.psychres.2020.113100)
- Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). Un viaje alrededor de alfa y omega para estimar la fiabilidad de consistencia interna. *Anales de Psicología*, 33(3), 755-782. doi: [10.6018/analesps.33.3.268401](https://doi.org/10.6018/analesps.33.3.268401)
- Wakashima, K., Asai, K., Kobayashi, D., Koiwa, K., Kamoshida, S., & Sakuraba, M. (2020). The Japanese version of the Fear of COVID-19 Scale: Reliability, validity, and relation to coping behavior. *PloS One*, 15(11), e0241958. doi: [10.1371/journal.pone.0241958](https://doi.org/10.1371/journal.pone.0241958)

Winter, T., Riordan, B. C., Pakpour, A. H., Griffiths, M. D., Mason, A., Poulgrain, J. W., & Scarf, D. (2020). Evaluation of the English version of the Fear of COVID-19 Scale and its relationship with behavior change and political Beliefs. *International Journal of Mental Health and Addiction*. [Advance online publication]. doi: [10.1007/s11469-020-00342-9](https://doi.org/10.1007/s11469-020-00342-9)

Características métricas de la Escala de Actitudes hacia las Personas con Discapacidad (EAPD)

Metric Characteristics of the Scale of Attitudes towards Persons with Disabilities (SAPD)

Miguel Angel Sainz-Palafox ¹, José Ángel Vera-Noriega * ¹, Jesús Tánori-Quintana ²

1 - Centro de Investigación en Alimentación y Desarrollo A. C., Hermosillo, México.

2 - Instituto Tecnológico de Sonora, Cd. Obregón, México.

Recibido: 13/12/2021 Revisado: 14/04/2022 Aceptado: 14/05/2022

Introducción
Método
Resultados
Discusión y Conclusiones
Referencias

Resumen

Debido a la relevancia que supone entender el papel que juega la actitud hacia las personas con discapacidad en el proceso de inclusión educativa, se han elaborado numerosos instrumentos de medición. En el contexto hispano-parlante, se destaca la Escala de Actitudes hacia las Personas con Discapacidad (EAPD). Sin embargo, esta escala no ha sido validada en México. Por lo tanto, el objetivo del presente trabajo es evaluar la evidencia de validez de la EAPD, así como proponer una versión modificada del instrumento en una muestra de 1934 estudiantes universitarios del estado de Sonora, México. Para ello se llevaron a cabo un análisis Rasch, un análisis factorial exploratorio y confirmatorio, y uno de invarianza entre sexos. Los resultados muestran que la escala modificada cuenta con propiedades psicométricas adecuadas y se divide en una estructura bifactorial con actitudes igualitarias y discriminatorias sobre la discapacidad.

Palabras clave: *actitud, discapacidad, validación, educación superior, inclusión, estudiantes*

Abstract

Due to the relevance in the understanding of the role of attitudes towards people with disability in the inclusive education process, numerous psychometric tools have been developed. In the Spanish-speaking context, the Scale of Attitudes towards People with Disabilities (SAPD) stands out. However, its validity has not been demonstrated yet on the Mexican population. For this reason, this work aims at evaluating the validity evidence of the SAPD and proposing a modified version in a sample of 1934 higher education students from Sonora, Mexico. An exploratory and confirmatory factor analysis, as well as a Rasch and invariance analysis were carried out. Results show that the modified version has appropriate psychometric properties and a bifactorial structure with egalitarian and discriminatory attitudes about disability.

Keywords: *attitudes, disability, validation, higher education, inclusion, students*

*Correspondencia a: José Ángel Vera Noriega. Carretera a La Victoria km 0.6, C.P. 83304, Hermosillo, México. Teléfono celular: 662 173 0085. E-mail: avera@ciad.mx

Cómo citar este artículo: Sainz-Palafox, M. A., Vera-Noriega, J. Á., & Tánori-Quintana, J. (2022). Características métricas de la Escala de Actitudes hacia las Personas con Discapacidad (EAPD). *Revista Evaluar*, 22(2), 14-29. Recuperado de <https://revistas.unc.edu.ar/index.php/revaluar>

Participaron en la edición de este artículo: Melani Hajduczyk, María Eugenia Maiorana, Eugenia Barrionuevo, Pablo Carpintero, Stefano Macri, Florencia Ruiz, Mónica Serppe, Ricardo Hernández, Alicia Molinari.

Introducción

De acuerdo con la [Ley General para la Inclusión de las Personas con Discapacidad \(2011\)](#), por persona con discapacidad (PD) entendemos:

Toda persona que por razón congénita o adquirida presenta una o más deficiencias de carácter físico, mental, intelectual o sensorial, ya sea permanente o temporal y que al interactuar con las barreras que le impone el entorno social, pueda impedir su inclusión plena y efectiva, en igualdad de condiciones con los demás (Artículo 2, fracción XXI).

Por otra parte, la inclusión educativa (IE) es un proceso en constante evolución que tiene que ver con la participación de toda la comunidad, las políticas que rigen a las instituciones y la promoción de valores incompatibles con cualquier forma de exclusión y discriminación ([Booth & Ainscow, 2015](#)). Por lo tanto, la IE de las PD implica el esfuerzo de la comunidad en su conjunto, instituciones y normas culturales para eliminar o minimizar las barreras que las PD puedan enfrentar, y que se encuentran no solo en condiciones materiales, arquitectónicas o legales, sino que existen en el plano cultural. Es en este último aspecto, donde toma relevancia el concepto de actitud, identificada por diversos autores como una barrera invisible difícil de superar ([Arellano-Torres, Gaeta-González, Peralta-López, & Cavazos-Arroyo, 2019](#); [Rodríguez-Martín & Álvarez-Arregui, 2015](#)).

Según [Arias-González, Arias-Martínez, Verdugo-Alonso, Rubia-Avi y Jenaro-Río \(2016\)](#), la actitud representa “construcciones teóricas, que constan de componentes cognitivos, afectivos y cognitivo-conductuales y que se infieren a partir de la observación de la conducta o de las manifestaciones verbales de la persona” (p. 9). La actitud se entiende como una predisposición favorable o desfavorable, relativamente duradera, y que puede manifestarse tanto a través de la interacción

directa con el objeto/sujeto como de la idea que se tiene del mismo ([Salinas-Alarcón, 2014](#)).

La actitud hacia las PD puede ser entendida entonces como una característica más o menos estable del individuo, que predispone favorable o desfavorablemente un conjunto de ideas, creencias, sentimientos, preocupaciones, emociones y comportamientos relacionados con la discapacidad y que se manifiestan al interactuar directa o indirectamente con las PD. Puesto que la actitud es un estado latente que el individuo sostiene hacia el estímulo de referencia, esta puede considerarse como mediador entre dicho estímulo y su respuesta conductual en entornos específicos ([Arias-González et al., 2016](#)).

Debido a la importancia de entender el papel de la actitud en el proceso de la IE de PD, diversos autores señalan que desde hace décadas se han elaborado numerosos instrumentos en torno a este constructo ([Arias-González et al., 2016](#); [Fuentes, Pérez-Padilla, de la Fuente, & Aranda, 2021](#); [Salinas-Alarcón, 2014](#)). Sin embargo, se han asociado algunas dificultades a la medición de las actitudes, de las cuales se destacan: Superar el sesgo de la deseabilidad social ([Arellano-Torres et al., 2019](#)), la necesidad de validación diferenciada dependiendo del entorno cultural ([Fuentes et al., 2021](#); [Rodríguez-Martín & Álvarez-Arregui, 2015](#)), la tendencia a valorar positivamente las propias capacidades y las dificultades inherentes para evaluar y predecir con exactitud el propio comportamiento ([Myers & Twenge, 2019](#)). Adicionalmente, el impulso que se le ha dado al tema de la IE de las PD en los últimos años ha ocasionado que los prejuicios y valoraciones negativas se manifiesten de una manera más sutil, ligeramente hostil o ambivalente, bajo la forma de paternalismo y condescendencia ([Nario-Redmond, Kemerlink, & Silverman, 2019](#)).

Uno de los instrumentos más utilizados en español es la Escala de Actitudes hacia las Perso-

nas con Discapacidad (EAPD; Verdugo, Arias, & Jenaro, 1994, como se citó en Asociación Nacional de Universidades e Instituciones de Educación Superior [ANUIES], 2002; Fuentes et al., 2021; Rodríguez-Martín & Álvarez-Arregui, 2015). En el nivel de educación superior, este se ha utilizado para comparar las actitudes de estudiantes de distintas carreras (Araya-Cortés, Gonzales-Arias, & Cerpa-Reyese, 2014; Arellano-Torres et al., 2019; Macías-Gómez, Aguilera-García, Rodríguez-Sánchez, & Gil-Hernández, 2019), de estudiantes de educación (Macías-Gómez, 2016; Macías et al., 2019; Polo-Sánchez, 2017; Polo-Sánchez & Aparicio-Puerta, 2018; Polo-Sánchez & López-Justicia, 2006), estudiantes de ciencias de la salud (Mella-Díaz & González-Quiroga, 2007) y de docentes de nivel superior de psicología y enfermería (Barradas-Alarcón, Robledo-Salinas, Guzmán-Ibáñez, & Gutierrez-Serrano, 2014).

A pesar del uso extendido de la EAPD, la escala no ha sido adaptada lingüística o culturalmente en el contexto latinoamericano, especialmente en el ámbito de la educación superior (Fuentes et al., 2021; Rodríguez-Martín & Álvarez-Arregui, 2015). Para que las interpretaciones derivadas de una prueba psicológica sean válidas es indispensable que exista evidencia de que sus puntuaciones son adecuadas para la población que se está evaluando (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2018). Además, los métodos de validación más actualizados contemplan no solo la teoría clásica del test y el análisis factorial exploratorio (AFE; Hair, Anderson, Tatham, & Black, 1999), sino también la teoría de respuesta al ítem (Kim & Kyllonen, 2006; Zamora-Araya, Smith-Castro, Montero-Rojas, & Moreira-Mora, 2018), el análisis factorial confirmatorio (AFC; Hair et al., 1999; Manzano-Patiño, 2018; Pilatti, Godoy, & Brussino, 2012), el análisis de invariancia (Kline,

2016) y nuevos métodos para medir la fiabilidad, como el coeficiente omega (Ventura-León & Caycho-Rodríguez, 2017).

En un estudio reciente, Fuentes et al. (2021) elaboraron un instrumento para medir la actitud hacia la discapacidad física y sensorial. La construcción de ítems contempla componentes cognitivos, afectivos y conductuales, el interés de mantener la situación de desventaja de las personas con discapacidad y la presencia de actitudes tradicionales y paternalistas. Se utilizó una muestra de un total de 718 participantes: 603 estudiantes y 155 docentes pertenecientes a la Universidad Autónoma de Bolivia Gabriel Moreno. Se inició con una prueba piloto compuesta por 122 ítems, que pasaron a ser 34 en el análisis factorial exploratorio y finalmente 24 en la solución final. Se dividieron los ítems en dos dimensiones que explican el 58.79% de la varianza. La primera se llamó actitudes igualitarias y consta de 11 ítems, con un coeficiente alfa de Cronbach de .91 y la segunda dimensión se llamó actitudes discriminatorias y consta de 16 ítems con un alfa de Cronbach de .92. Para el análisis factorial confirmatorio se utilizó el método de máxima verosimilitud, reportando el *non-normed fit index* (NNFI) = .99, un *comparative fit index* (CFI) = .99 y un *root mean square error of approximation* (RMSEA) = .03, los cuales representan buenos índices de ajuste.

El presente estudio tiene como objetivos evaluar las evidencias de validez de constructo de la EAPD, así también como proponer una versión modificada del instrumento (ver Tabla 1) en una muestra de estudiantes universitarios del estado de Sonora, México, tomando como referencia el trabajo realizado por Fuentes et al. (2021). Se partió de la hipótesis de que existen dos variables latentes relacionadas: ideas discriminatorias e igualitarias, y a su vez, cada una de estas está relacionada con escalas específicas de la EAPD. Para ello, en primer lugar, se buscó establecer un

Tabla 1

Relación de escalas e ítems del instrumento original y la adaptación propuesta.

Escala EAPD original	Total de ítems: 37	Escala EAPD modificada	Total de ítems: 25
Escalas Discriminatorias			
Valoración de Capacidades y Limitaciones	1, 2, 4, 7, 8, 16, 20, 21, 29, 36	Valoración de Limitaciones	1, 2, 4, 8.
Reconocimiento y Negación de Derechos	6, 9, 12, 17, 14, 15, 22, 23, 27, 35, 37	Implicación Personal	10, 25, 26, 31.
Implicación Personal	3, 5, 10, 11, 25, 26, 31	Calificación Genérica	17, 18, 24, 28
Calificación Genérica	13, 18, 24, 28, 34	Escala Igualitaria	
Asunción de Roles	19, 30, 32, 33	Asunción de Roles y Capacidades	19, 21, 30, 33, 36
Bidimensional			
		Reconocimiento y Negación de Derechos	Negación 6, 15, 23, 35, 37 Reconocimiento 14 22 y 27

modelo general de actitudes igualitarias/discriminatorias a través de las escalas de la EAPD y posteriormente se determinó la contribución y pertinencia de cada escala por separado.

Método

Diseño

Se trata de un estudio no experimental, descriptivo y transversal.

Participantes

La muestra está formada por un total de 1934 universitarios, de los cuales 645 se identifican con el sexo femenino (33.4%) y 1289 con el sexo masculino (66.6%), agrupados en las siguientes áreas: Departamento de Industrial (16.4%): Ing. Industrial; Departamento de Metal-

mecánica (27.7%): Ing. Mecánica, Ing. Mecatrónica, Ing. Aeronáutica; Departamento Eléctrica/Electrónica (15.2%): Ing. Eléctrica, Ing. Electrónica, Ing. Biomédica; Departamento Económico-Administrativo (17.3%): Lic. Administración, Ing. Gestión Empresarial; y Departamento de Sistemas e Informática (22%): Ing. Sistemas Computacionales, Ing. Informática. El 33.4% de los participantes se encuentran entre el primer y el segundo semestre, 32.4% entre el tercero y cuarto, el 19.1% entre el quinto y sexto, y el 15.2% en el séptimo o superior. La edad promedio de los estudiantes es de 19.86 años ($DE = 2.84$), con edades de entre 18 y 24 años. Tal como lo indican [Bandalos y Finney \(2019\)](#), realizar el AFE y el AFC en muestras independientes aporta mayor evidencia de validez, por lo que esta fue separada aleatoriamente en dos partes (que corresponden al 50% de la muestra total), y fueron llamadas submuestra 1 y submuestra 2.

Instrumento

El instrumento utilizado fue la Escala de Actitudes hacia las Personas con Discapacidad (EAPD). Para medir el constructo de actitud hacia la discapacidad, Verdugo, Arias y Jenaro (citados en ANUIES, 2002) desarrollaron una estructura multifactorial distribuida teóricamente en cinco dimensiones y 37 ítems: *valoración de capacidades y limitaciones* (VCL con 10 ítems), *reconocimiento/negación de derechos* (RND con 11 ítems), *implicación personal* (IP con siete ítems), *calificación genérica* (CG con cinco ítems) y *asunción de roles* (AR con cinco ítems).

Los ítems 1, 2, 3, 4, 5, 6, 7, 8, 10, 15, 17, 18, 23, 24, 25, 26, 28, 29, 31, 34, 35 y 37 son afirmaciones discriminatorias sobre la discapacidad, por ejemplo, el ítem 1 dice: *Las personas con discapacidad con frecuencia son menos inteligentes que las demás personas*. Para estos ítems, un mayor acuerdo representa actitudes menos favorables hacia la discapacidad. Por el contrario, los ítems 9, 11, 12, 13, 14, 16, 19, 20, 21, 22, 27, 30, 32, 33 y 36 son afirmaciones igualitarias sobre la discapacidad; por ejemplo, el ítem 9 dice: *Las personas con discapacidad deberían tener las mismas oportunidades de empleo que cualquier otra persona*. Para estos ítems un mayor acuerdo representa actitudes más favorables hacia la discapacidad. Por este motivo, antes de realizar los análisis, se invirtieron los valores de los ítems negativos para que una mayor puntuación indique siempre una valoración positiva de la discapacidad.

El instrumento está construido con una escala tipo Likert con seis opciones de respuesta, donde: 1 = *Estoy muy de acuerdo* (MA); 2 = *Estoy bastante de acuerdo* (BA); 3 = *Estoy parcialmente de acuerdo* (PA); 4 = *Estoy parcialmente en desacuerdo* (PD); 5 = *Estoy bastante en desacuerdo* (BD); 6 = *Estoy en total desacuerdo* (TD) para

las afirmaciones igualitarias, mientras que las puntuaciones se invierten para las afirmaciones discriminatorias.

Para el proceso de construcción y validación de la EAPD, se llevó a cabo una revisión exhaustiva de la literatura, que sirvió de insumo para construir un banco de ítems. La validez de contenido de los ítems se evaluó mediante el método de juicio de expertos, con el apoyo de un grupo compuesto por 250 profesionales de distintas áreas. Se determinó la fiabilidad de la escala a través de tres métodos. El primero fue el coeficiente alfa de Cronbach con un valor de .92, el cual mide el nivel de consistencia interna de los ítems. El segundo fue el método de división en mitades y, por último, el tercero fue el método Lambda de Guttman (Polo-Sánchez & López-Justicia, 2006).

Procedimiento

El estudio se llevó a cabo en el Instituto Tecnológico de Hermosillo, ubicado en el centro del estado de Sonora, México. La participación de los alumnos que conformaron la muestra fue posible gracias a la colaboración de los docentes, quienes compartieron la encuesta directamente a sus alumnos a través de correos electrónicos. Previo a ello, se hizo un trabajo de vinculación con los jefes de departamento de los distintos programas académicos, pues resultaba indispensable contar con su aprobación. El instrumento fue transformado a formato digital a través de la plataforma *Google forms*. Antes de iniciar con sus respuestas, los alumnos tenían que leer los objetivos del estudio y aprobar un consentimiento informado. La información fue tratada con estricta confidencialidad y apego a las directrices emitidas por la *Sociedad Mexicana de Psicología* (2007), con el respeto debido a la dignidad y el bienestar de los participantes.

Análisis de datos

Para llevar a cabo el AFE, así como también obtener el índice omega de McDonald, se utilizó el paquete estadístico SPSS v26. Por otra parte, para llevar a cabo el AFC y el análisis de invarianza, se utilizó la extensión AMOS v26. El análisis Rasch se llevó a cabo con el software WINS-TEPS v. 3.74. Finalmente, se utilizó el software G Power 3.1.9.7 para evaluar el tamaño del efecto en la prueba de contraste de hipótesis.

Resultados

El AFE tiene como objetivo establecer una estructura interna que se ajuste mejor a las correlaciones del conjunto de datos con los que se cuenta, así como eliminar variables que sean irrelevantes (Hair et al., 1999). Para llevar a cabo este análisis se utilizó la submuestra 1. El modelo elegido fue de máxima verosimilitud y la pertinencia del análisis se determinó mediante la prueba Kaiser-Meyer-Olkin (KMO), considerando el rango de .50 a .70 como aceptable y de .80 a 1 como excelente, en conjunto con la prueba de esfericidad de Bartlett, con rotación oblimin directa y se eliminaron aquellos ítems cuya carga fuera menor a .30. Se observó que las puntuaciones de asimetría y curtosis de cada ítem presentaran valores superiores a -7 e inferiores a 7, como condición para considerarlos aceptables (Bandalos & Finney, 2019).

El análisis de frecuencias permitió observar que los ítems 3, 11, 12 y 16 se encontraban fuera de rango en asimetría y/o curtosis, por lo que fueron retirados del análisis, lo que dio como resultado 33 ítems. Posteriormente se realizó el AFE para cada escala del instrumento. Dos escalas mostraron una estructura bifactorial: RND con los ítems 37, 23, 15, 35 y 6 en el Factor 1 y los

ítems 14, 12, 22 y 27 en el Factor 2 y VCL con los ítems 1, 4, 8, 2 y 29 en el Factor 1 y los ítems 21, 26, 32 y 36 en el Factor 2. Se observó que los ítems se dividieron de forma tal que las aseveraciones negativas sobre las PD quedaron en el primer factor (con mayor varianza explicada) y las que aludían a opiniones positivas en el segundo. Como consecuencia, la escala VCL fue dividida en Valoración de Capacidades (VC) y Valoración de Limitaciones (VL), mientras que RND fue dividida en Negación de Derechos (ND) y Reconocimiento de Derechos (RD). Para el resto de las escalas se obtuvo una solución unifactorial (IP: 5, 10, 25, 26 y 31; CG: 17, 18, 24 y 28; AR: 19, 30 y 33) (ver Tabla 2).

Se observó que las escalas IP y CG quedaron constituidas únicamente por ítems de carácter negativo, mientras que la escala AR por ítems positivos. Con esto se observa que el instrumento queda prácticamente dividido en escalas discriminatorias (ND, VL, IP y CG) y escalas igualitarias (RD, VC y AR).

Posteriormente se realizó el análisis Rasch. El método utilizado fue el *Rating Scale Model* (RSM) propuesto por Andrich (1978) (Kim & Kyllonen, 2006), pues es apropiado para datos politómicos de escalas tipo Likert. Los indicadores de ajuste del ítem pueden ser usados para identificar los ítems que definen una dimensión singular, sujeta a la restricción de que el modelo es apropiado para ese conjunto de ítems, considerando como aceptables valores de *Infit* y *Outfit* entre .5 y 1.5. Se observa un ajuste adecuado en todos los ítems que conforman el instrumento en relación con la escala a la que pertenecen.

Acto seguido, se procedió a ejecutar el AFC, para lo cual se utilizó la submuestra 2. El AFC proporciona la serie de estimaciones de ecuaciones simultáneas que resulta más adecuada y eficiente mediante regresiones múltiples y permite evaluar la contribución de cada ítem de la escala,

Tabla 2

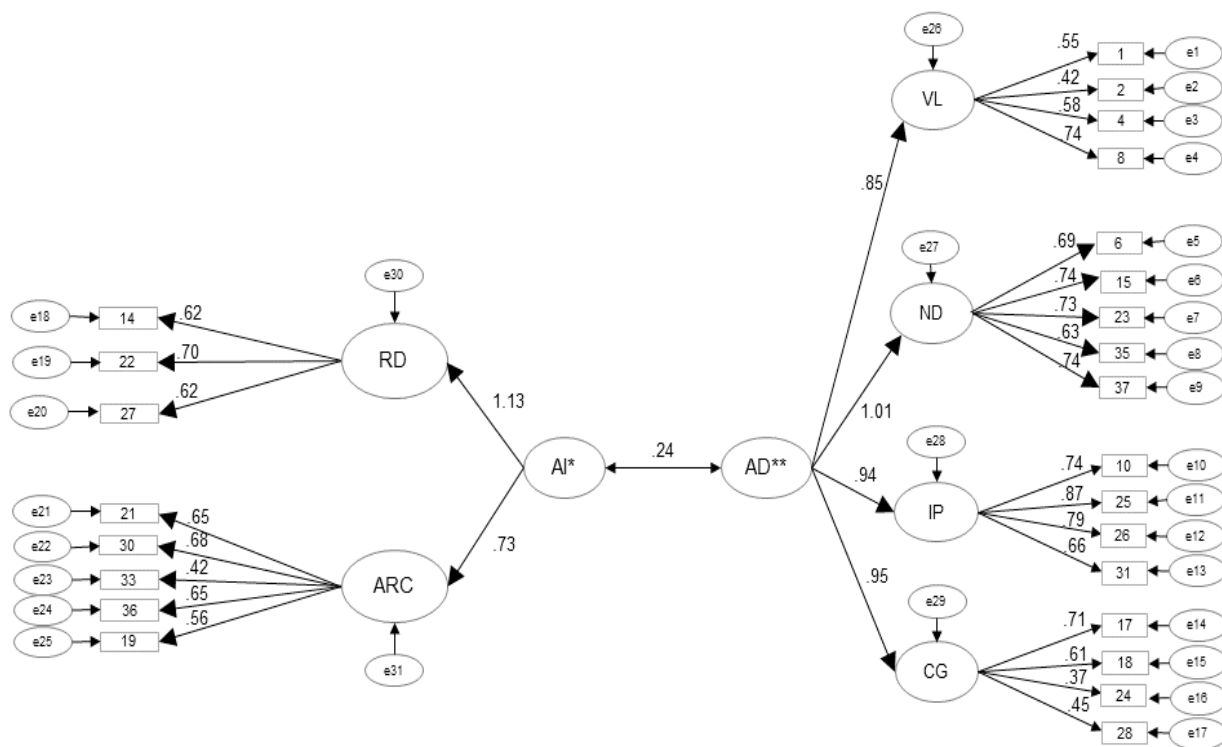
Valores iniciales obtenidos por AFE y varianza total explicada para cada factor.

Escala	KMO	Dimensión	No. de reactivos	Varianza total explicada	Cargas factoriales	
					Mín	Máx
Actitudes hacia las personas con discapacidad (EAPD)	.77	VCL	9	34.94	.40	.66
	.83	R/ND	9	48.17	.42	.73
	.81	IP	5	60.82	.68	.87
	.69	CG	4	36.01	.47	.75
	.70	AR	3	36.41	.43	.65

así como incorporar la fiabilidad de la escala en la estimación de las variables (Hair et al., 1999).

Se estableció un modelo a través del método de máxima verosimilitud con *bootstrap* 500 para determinar la pertinencia de las siete escalas bajo los criterios de los índices de bondad de ajuste; CMIN/DF menor que 3 (Hu & Bentler, 1995; en Pilatti et al., 2012); CFI cercano a 1, RMSEA y SRMR < .05 (Manzano-Patiño, 2018).

Se buscó establecer un modelo que explique dos variables latentes correlacionadas pero independientes: actitudes igualitarias y discriminatorias. Por ese motivo, las escalas RND y VCL fueron separadas, de modo que contribuyeran simultáneamente a ambas variables latentes, pero se mantuvieran correlacionadas estructuralmente. Se procedió a elaborar el modelo en AMOS 26 observando que la varianza entre las escalas VC

**Figura 1.** Coeficientes estandarizados del modelo propuesto (A2).

Nota. *AI = Actitudes igualitarias, *AD = Actitudes discriminatorias**

Tabla 3

Índices de bondad de ajuste obtenidos en el AFC para los modelos probados.

Modelo		Indicadores de ajuste							
Modelo general	χ^2	GL	CMIN/GL	CFI	RMSEA	LO	HI	AIC	SRMR
A1	1220	345	4.1	.90	.05	.54	.61	1386.30	.05
A2	1000	264	3.7	.91	.05	.51	.58	1172.07	.04
Escala									
IP	4	2	2.1	.99	.03	.00	.08	28.21	.01
RND	49	17	2.9	.98	.04	.03	.06	103.33	.02
CG	17	2	8.7	.96	.91	.05	.13	41.59	.03
VL	9	1	9.0	.98	.09	.06	.11	35.01	.01
ARC	60	4	15.18	.94	.12	.09	.15	92.73	.04

y AR es igual a 1, por lo que se decidió unir estas dos escalas en una sola, a la cual se le llamó Asunción de Roles y Capacidades (ARC). Posteriormente, se elaboró un modelo general con las seis escalas finales (A1).

Como resultado se observó que el modelo no se ajustaba a los índices esperados (ver Tabla 3). Se procedió a evaluar los índices de modificación y se obtuvo la mejor solución (A2, ver Tabla 3) tras eliminar el ítem 32, así como establecer las siguientes covarianzas de error: $e1 - e2$ (.24), $e5 - e8$ (-.08), $e8 - e9$ (.23) y $e23 - e24$ (.05). La solución final contempla 25 ítems. Cada una de las dimensiones está formada por ítems cuyos pesos de regresión nos permiten indicar su importancia, dados sus valores y la variación entre ellos.

Como consecuencia del AFC se reestructuraron algunas escalas. En primer lugar, la escala VCL se redujo simplemente a VL, con los ítems 1, 2, 4 y 8. En segundo lugar, las variables de VC y AR se combinaron en una sola escala llamada ARC, conformada por los ítems 19, 21, 30, 33 y 36 (ver Figura 1).

Una vez definido el modelo general, se procedió a analizar cada escala por separado. Se observó que las escalas IP y RND cuentan con un ajuste excelente, mientras que las escalas VL, CG y ARC no se ajustan a los índices esperados (ver Tabla 3).

Para el análisis de fiabilidad de la escala se obtuvieron los valores del coeficiente ω de McDonald (ver Tabla 4), en la cual se obtuvo un valor total de .88, con valores por escala que oscilan entre .63 y .83. La escala CG no supera el mínimo esperado de .65 (Katz, 2011), sin embargo, está muy cerca del límite inferior permitido.

Con el objetivo de verificar la pertinencia de las escalas modificadas y la posible pérdida de la varianza total explicada tras la eliminación de los ítems, se realizó nuevamente el análisis factorial bajo los mismos parámetros establecidos al inicio. Se observa que no hay cambios en términos del KMO ni en la varianza total explicada (ver Tabla 5). Cabe señalar que el cálculo original se hizo en la submuestra 1, mientras que el cálculo de la escala modificada se hizo en la submuestra 2.

Tabla 4

Índices de fiabilidad de las escalas.

ω de McDonald					
RND	VL	IP	CG	ARC	Total
.79	.70	.84	.63	.74	.88

Finalmente, se procedió a realizar el análisis de invarianza entre los participantes que se identificaban con el sexo femenino y el masculino, pues diversos autores (Fuentes et al., 2021; Polo-Sánchez, Fernández-Jiménez, & Fernández-Cabezas, 2018; Rodríguez-Martín & Álvarez-Arregui, 2015; Salinas-Alarcón, 2014) señalan que se han encontrado diferencias entre las puntuaciones de estos grupos en el nivel superior de enseñanza. Este procedimiento contempla la estimación de tres modelos anidados que son progresivamente comparados entre sí. En primer lugar, un modelo sobre el que no se imponen restricciones y se denomina invarianza configuracional (M1). En segundo lugar, se establece un modelo con restricciones en la que los coeficientes de regresión son igualados y se denomina invarianza métrica (M2). Finalmente, se elaboró un modelo donde los interceptos se igualan a las cargas factoriales, el cual se denomina invarianza escalar (M3).

Los tres modelos son comparados en relación al modelo base donde se encuentran anidados, buscando que las diferencias en los valores de Chi-cuadrada ($\Delta\chi^2$) arrojen un valor p no significativo y un ΔCFI menor a .001 (Kline, 2016). Sin embargo, es relevante tener en cuenta que autores como Dominguez-Lara y Medrano (2016) y Ochoa-Meza, Sierra, Pérez-Rodrigo, Aranceta-Bartrina y Esparza-Del Villar (2015) conside-

ran que el CFI es el indicador que toma mayor relevancia para observar el contraste y determinar la invarianza. Para los efectos del presente estudio, se realizó una estimación entre la invarianza configuracional y métrica. En primer lugar, el valor Chi cuadrada ($\Delta\chi^2 = 1163.60, p = .42$) y los índices de ajuste (CMIN/GL = 2.4, CFI = .92, RMSEA = .3) nos llevan a aceptar la hipótesis de invarianza configuracional. En relación con la invarianza métrica, la diferencia del CFI ($\Delta CFI = .002$) entre ambos modelos es menor al límite establecido de .01, pero significativo a $p < .01$. En este sentido, se puede considerar que el modelo presentado alcanza la invarianza métrica entre sexos.

Para finalizar, se realizaron contrastes de hipótesis con el objetivo de observar si sus resultados son congruentes con los que presentan otros estudios, lo que aportaría mayor evidencia de validez (Fuentes et al., 2021). Además del sexo, otras variables de atributo que toman gran relevancia para generar diferencias significativas entre grupos son la frecuencia de contacto con PD y la edad. Se ha reportado que aquellos alumnos de nivel superior que han tenido contacto con la discapacidad presentan actitudes más positivas (González-Cortés & Roses-Campos, 2016), mientras que la edad parece generar diferencias en algunos casos (Luque-de la Rosa & Gutié-

Tabla 5

Comparación de la escala original con la modificada, en términos de KMO, número de reactivos y varianza total explicada.

	KMO		Varianza total explicada			
	*E	**C	E	C	E	C
VCL vs VL	.77	.73	9	4	34.94	37.83
RND vs RND	.83	.81	9	8	48.17	46.17
IP vs IP	.81	.79	5	4	60.82	59.13
CG vs CG	.69	.68	4	4	36.01	36.01
AR vs ARC	.70	.71	3	4	36.41	36.61

Nota. *E = Después del AFE, **C = Después del AFC.

rez-Cáceres, 2014) y en otros no (Fuentes et al., 2021).

En el caso del sexo, se realizó la prueba t de Student para muestras independientes, considerando como válidos los tamaños de efecto mayores a .20 (d) y potencia estadística mayor a .80 ($1-b$) para evitar caer en error tipo II (falsos negativos; Cárdenas-Castro & Arancibia-Martini, 2014; Ventura-León, 2018). Los resultados no varían significativamente dependiendo de si se asumen o no varianzas iguales. Tomando en consideración todos los criterios, se encontraron diferencias significativas en las variables VL ($t = 4.33, p < .001, d = .20$ y $1-b = .99$), IP ($t = 4.28, p < .001, d = .21$ y $1-b = .99$) y RD ($t = 6.15, p < .001, d = .30$ y $1-b = .99$), con puntuaciones medias más altas para el sexo femenino tanto en las escalas igualitarias como en las discriminatorias, lo que indica actitudes más favorables.

Para las variables frecuencia de contacto y edad se realizó la prueba de análisis de varianza de una vía (ANOVA) con seis y tres grupos, respectivamente. En este caso, se consideran aceptables valores mayores a .10 para el tamaño del efecto (f). Con relación a los grupos de edad, aunque se encontraron valores $p < .05$ para IP, CG y ND; solo este último tiene el tamaño del efecto suficiente para considerarlos significativos ($F = 3.96, p = .001, f = .10$ y $1-b = .95$), con diferencias entre los grupos de 20 y de 21 años, con medias más altas para los últimos. Con relación a la frecuencia de contacto, se encontraron diferencias significativas para IP ($F = 2.98, p = .03, f = .12, 1-b = .95$), CG ($F = 3.03, p = .03, f = .12, 1-b = .95$), ARC ($F = 4.22, p = .006, f = .14, 1-b = .95$), RD ($F = 4.11, p = .007, f = .14, 1-b = .95$) y ND ($F = 4.40, p = .04, f = .14, 1-b = .95$). En todos los casos, los que tienen frecuencia de contacto casi nula muestran medias más bajas tanto en las variables igualitarias como en las discriminatorias.

Discusión y conclusiones

La fecha de publicación de la EAPD coincide con la Declaración de Salamanca sobre Necesidades Educativas Especiales (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO], 1994), que supuso un momento decisivo para que el enfoque de educación inclusiva se lograra posicionar a nivel internacional, trayendo consigo cambios en las políticas y sistemas educativos (Duk, Cisternas, & Ramos, 2019). Debido a ello, es de esperarse que desde la publicación de la EAPD los parámetros con los que se miden las expresiones de rechazo o aceptación de la discapacidad hayan cambiado sustantivamente. Diversos autores (Arellano-Torres et al., 2019; Fuentes et al., 2021; Rodríguez-Martín & Álvarez-Arregui, 2015) han señalado como limitantes de la EAPD el hecho de que se basa en prejuicios tradicionales, prejuicios que actualmente son más susceptibles a sesgos de deseabilidad social. Sin embargo, estudios empíricos dejan ver que esta división entre prejuicios tradicionales y modernos no siempre es clara desde el punto de vista métrico (Arellano-Torres et al., 2019).

En el presente estudio el AFE permite visualizar la estructura bifactorial de las escalas caracterizadas por afirmaciones igualitarias, por un lado, y discriminatorias, por el otro. La proporción de afirmaciones discriminatorias comparadas con las igualitarias pasó del 59% en la escala original al 68% en la modificada. Consideramos que la pérdida de ítems con valoraciones positivas se explica por el mayor reconocimiento de capacidades y derechos que ya no se ponen en duda para este grupo de personas. Por ejemplo, para el ítem 3: *Permitiría que su hijo aceptase la invitación a un cumpleaños que le hiciera un niño con discapacidad*, 1695 participantes estuvieron muy de acuerdo, y solo 19 participantes

mostraron algún grado de desacuerdo; mientras que para el ítem 12: *Las personas con discapacidad también pueden divertirse con las demás personas*, 1557 participantes estuvieron muy de acuerdo y solo 19 mostraron algún grado de desacuerdo. Por otra parte, para el ítem 25: *No quiero trabajar con personas con discapacidad*, solo 954 participantes estuvieron muy en desacuerdo, y se observa una mayor distribución entre el resto de las opciones de respuesta, al igual que con el ítem 15: *Las personas con discapacidad deberían ser encerradas en instituciones especiales*, donde 909 personas estuvieron muy en desacuerdo.

Lo anterior nos permite apoyar la hipótesis que sostiene que la EAPD cuenta con afirmaciones sobre la discapacidad que ya no son aplicables a la sociedad moderna, ni al contexto universitario. Sin embargo, el análisis presentado también nos permite visualizar que específicamente las escalas de RND e IP aún son aproximaciones adecuadas para entender aspectos específicos del espectro de actitudes hacia la discapacidad, especialmente esta última, pues cuenta con un 59% de varianza explicada y parámetros de ajuste excelentes en el AFC si se juzga por separado. Por otra parte, la existencia de factores correlacionados y no independientes se explica por el método de rotación utilizado (oblimin directo; Hair et al., 1999). El AFC muestra una covarianza de .24 entre las dos variables latentes (actitudes discriminatorias e igualitarias), lo que indica que es débil, pero suficiente para considerar que los constructos están relacionados (Valdés-Cuervo, García-Vázquez, Torres-Acuña, Urías-Murrieta, & Grijalva-Quirón, 2019).

Se puede observar que las puntuaciones de las escalas se encuentran entre .35 y .87, lo que representa una desventaja en relación con el instrumento desarrollado por Fuentes et al. (2021), que contiene ítems con cargas mayores a .50. En cuanto a la pertinencia de las covarianzas de error,

Brown (2015) señala que cuando se utilizan métodos como el autorreporte para la validación de instrumentos mediante el AFC, es posible que se necesiten errores correlacionados para dar cuenta de la covarianza en los indicadores no explicada por una posible causa común exógena. Por otra parte, en el análisis de confiabilidad encontramos puntuaciones que son comparables con la escala total ($\omega = .88$ vs $\alpha = .92$) en relación con la propuesta de Fuentes et al. (2021). La interpretación de algunas escalas específicas debe llevarse con cautela si se hace de manera aislada, por mostrar un índice bajo, mientras que otras presentan puntuaciones más consistentes.

Adicionalmente, el presente estudio permite apoyar la hipótesis de que los participantes del sexo masculino y aquellos que han tenido contacto casi nulo con personas con discapacidad (PD) muestran medias más bajas en las puntuaciones de actitudes igualitarias y discriminatorias, mientras que la edad por sí misma parece no establecer diferencias considerables, lo cual es congruente con lo establecido por otros autores (Fuentes et al., 2021; González-Cortés & Roses-Campos, 2016; Rodríguez-Martín & Álvarez-Arregui, 2015; Salinas-Alarcón, 2014). El análisis de invarianza aporta mayor certeza a que la medición y comparación establecida entre sexos se encuentra libre de sesgos.

Como limitante del estudio se señala el hecho de que no se hizo un análisis de validez concurrente. Si bien Fuentes et al. (2021) encontraron una correlación significativa de la EAPD con el instrumento que elaboraron, futuras líneas de investigación podrían señalar la posible relación entre la estructura bifactorial aquí descrita y los factores de actitudes igualitarias y prejuiciosas. Otra limitante importante es que no se indagó sobre la invarianza entre distintos tipos de discapacidad. También se indica que los participantes pertenecen a una universidad tecnológica, con

una proporción mayor de participantes del sexo masculino y de titulaciones relacionadas con las ingenierías. Debido a que se ha reportado que estos grupos presentan valoraciones menos positivas hacia la discapacidad (Rodríguez-Martín & Álvarez-Arregui, 2015), se considera prudente replicar los análisis en el contexto de estudiantes que cursen carreras relacionadas con las ciencias sociales.

Se concluye que la revisión de las propiedades métricas de la EAPD era una asignación pendiente, pues aún es reconocida como uno de los instrumentos de mayor uso en el contexto hispanoparlante para evaluar las actitudes hacia las PD. Los resultados aquí presentados permitirán a los investigadores y profesionales que trabajan en el campo de la inclusión de PD en el contexto de las instituciones de educación superior tomar decisiones críticas sobre el uso de instrumentos de medición. A su vez, se considera que el análisis aquí presentado arroja evidencia empírica de que mientras la valoración de capacidades y el reconocimiento de derechos para las PD han ganado una aceptación tal que ya no se ponen en duda, aún se encuentra resistencia en la disposición para asumir roles de apoyo e involucrarse de manera cercana con PD.

Referencias

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2018). *Estándares para pruebas educativas y psicológicas*. Washington, DC: American Educational Research Association.
- Araya-Cortés, A., González-Arias, M., & Cerpa-Reyes, C. (2014). Actitud de universitarios hacia las personas con discapacidad. *Educación y Educadores*, 17(2), 289-305. doi: 10.5294/edu.2014.17.2.5
- Arellano-Torres, A., Gaeta-Gonzales, M. L., Peralta-Lopez, F., & Cavazos-Arroyo, J. (2019). Actitudes hacia la discapacidad en una universidad mexicana. *Revista Brasileira de Educação*, 24, 1-20. doi: 10.1590/s1413-24782019240023
- Arias-González, V., Arias-Martínez, B., Verdugo-Alonso, M. Á., Rubia-Avi, M., & Jenaro-Río, C. (2016). Evaluación de actitudes de los profesionales hacia las personas con discapacidad. *Siglo Cero*, 47(2), 7-41. doi: 10.14201/scero2016472741
- Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES). (2002). *Manual para la Integración de Personas con Discapacidad en las Instituciones de Educación Superior*. Recuperado de www.conapred.org.mx
- Bandalos, D., & Finney, S. J. (2019). Factor analysis: Exploratory and confirmatory. En G. R. Hancock, L. M. Stapleton, & R. O. Mueller (Eds.), *The Reviewer's Guide to Quantitative Methods in the Social Sciences* (2^{da} ed., pp. 98-122). doi: 10.4324/9781315755649
- Barradas-Alarcón, M. E., Robledo-Salinas, M. L., Guzmán-Ibáñez, M. L., & Gutierrez-Serrano, L. (2014). Actitudes hacia las personas con discapacidad en académicos de la Universidad Veracruzana. *Revista Iberoamericana de Producción Académica y Gestión Educativa*, 2. Recuperado de <https://www.pag.org.mx>
- Booth, T., & Ainscow, M. (2015). *Guía para la Educación Inclusiva. Desarrollando el aprendizaje y la participación en los centros escolares*. Recuperado de <http://consultorahumanitas.cl>
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2^{da} ed.). Nueva York: The Guilford Press.
- Cárdenas-Castro, J. M., & Arancibia-Martini, H. (2014). Potencia estadística y cálculo del tamaño del efecto en G * Power: Complementos a las pruebas de significación estadística y su aplicación en psicología. *Salud y Sociedad*, 5(2), 210-224. Recuperado de <http://www.saludysociedad.cl>
- Dominguez-Lara, S. A., & Medrano, L. A. (2016). Inva-

- rianza factorial del Cognitive Emotional Regulation Questionnaire (CERQ) en universitarios limeños y cordobeses. *Universitas Psychologica*, 15(1), 89-98. doi: [10.11144/Javeriana.upsy15-1.ifce](https://doi.org/10.11144/Javeriana.upsy15-1.ifce)
- Duk, C., Cisternas, T., & Ramos, L. (2019). Formación docente desde un enfoque inclusivo. A 25 Años de la Declaración de Salamanca, nuevos y viejos desafíos. *Revista Latinoamericana de Educación Inclusiva*, 13(2), 91-109. doi: [10.4067/s0718-73782019000200091](https://doi.org/10.4067/s0718-73782019000200091)
- Fuentes, V., Pérez-Padilla, J., de la Fuente, Y., & Aranda, M. (2021). Creation and validation of the Questionnaire on Attitudes towards Disability in Higher Education (QAD-HE) in Latin America. *Higher Education Research and Development*. doi: [10.1080/07294360.2021.1927997](https://doi.org/10.1080/07294360.2021.1927997)
- González-Cortés, E., & Roses-Campos, S. (2016). ¿Barre-ras invisibles? Actitudes de los estudiantes universitarios ante sus compañeros con discapacidad. *Revista Complutense de Educación*, 27(1), 219-235. doi: [10.5209/rev_RCED.2016.v27.n1.45702](https://doi.org/10.5209/rev_RCED.2016.v27.n1.45702)
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1999). *Análisis Multivariante* (5^{ta} ed.). Madrid, España: Prentice Hall.
- Katz, M. H. (2011). *Multivariable Analysis* (3^{ra} ed.). New York: Cambridge University Press.
- Kim, S., & Kyllonen, P. C. (2006). Rasch rating scale modeling of data from the Standardized Letter of Recommendation. *ETS Research Report Series*, 2, i-22. doi: [10.1002/j.2333-8504.2006.tb02038.x](https://doi.org/10.1002/j.2333-8504.2006.tb02038.x)
- Kline, R. B. (2016). *Principles and Practices of Structural Equation Modelling* (4^{ta} ed.). New York: The Guilford Press.
- Ley General para la Inclusión de las Personas con Discapacidad. (2011). México. Recuperado de <http://dof.gob.mx>
- Luque de la Rosa, A., & Gutiérrez-Cáceres, R. (2014). La integración educativa y social del alumnado con discapacidad en el EEES: Universidad de Bolonia. *Revista Complutense de Educación*, 25(1), 153-175. doi: [10.5209/rev_RCED.2014.v25.n1.41344](https://doi.org/10.5209/rev_RCED.2014.v25.n1.41344)
- Macías-Gómez, E. (2016). Actitudes de estudiantes de magisterio en educación primaria hacia las personas con discapacidad. *Revista Nacional e Internacional de Educación Inclusiva*, 9(1), 54-69. Recuperado de <https://revistaeducacioninclusiva.es>
- Macías-Gómez, M. E., Aguilera-García, J. L., Rodríguez-Sánchez, M., & Gil-Hernández, S. (2019). Un estudio transversal sobre las actitudes de los estudiantes de pregrado y máster en Ciencias de la Educación hacia las personas con discapacidad. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 22(1), 225-240. doi: [10.6018/rei-fop.22.1.353031](https://doi.org/10.6018/rei-fop.22.1.353031)
- Manzano-Patiño, A. P. (2018). Introducción a los modelos de ecuaciones estructurales. *Investigación en Educación Médica*, 7(25), 67-72. Recuperado de <http://www.riem.facmed.unam.mx>
- Mella-Díaz, S., & González-Quiroga, L. (2007). Actitudes de los estudiantes de la salud de la Universidad de Chile hacia las personas con discapacidad. *Revista Chilena de Terapia Ocupacional*, 7, 3-12. doi: [10.5354/0717-6767.2007.42](https://doi.org/10.5354/0717-6767.2007.42)
- Myers, D. G., & Twenge, J. M. (2019). *Psicología social* (13^a ed.). México DF, México: McGraw Hill Interamericana.
- Nario-Redmond, M. R., Kemerling, A. A., & Silverman, A. (2019). Hostile, benevolent, and ambivalent ableism: Contemporary manifestations. *Journal of Social Issues*, 75(3), 726-756. doi: [10.1111/josi.12337](https://doi.org/10.1111/josi.12337)
- Ochoa-Meza, G., Sierra, J. C., Pérez-Rodrigo, C., Aranceta-Bartrina, J., & Esparza-Del Villar, Ó. A. (2015). Validez e invarianza factorial de un modelo socioecológico para explicar el consumo de fruta en niños escolares mexicanos. *Nutrición Hospitalaria*, 31(2), 649-657. Recuperado de <https://www.nutricionhospitalaria.org>
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (1994). *Declaración de Salamanca y Marco de Acción Sobre Necesidades Educativas Especiales*. Recuperado de <https://unesdoc.unesco.org>

- Pilatti, A., Godoy, J. C., & Brussino, S. (2012). Análisis factorial confirmatorio del Cuestionario de Expectativas hacia el Alcohol para Niños Argentinos (CEA-N). *Acta Colombiana de Psicología*, 15(2), 11-20. Recuperado de <http://www.scielo.org.com>
- Polo-Sánchez, M. T. (2017). Innovación para la formación en inclusión: Actitudes de la comunidad universitaria. *International Journal of Developmental and Educational Psychology. Revista INFAD de Psicología*, 4(1), 185-194. doi: 10.17060/ijodaep.2017.n1.v4.1041
- Polo-Sánchez, M. T., & Aparicio-Puerta, M. (2018). Primeros pasos hacia la inclusión: Actitudes hacia la discapacidad de docentes en educación infantil. *Revista de Investigación Educativa*, 36(2), 365-379. doi: 10.6018/rie.36.2.279281
- Polo-Sánchez, M. T., Fernández-Jiménez, C., & Fernández-Cabezas, M. (2018). The attitudes of different partners involved in higher education towards students with disabilities. *International Journal of Disability, Development and Education*, 65(4), 442-458. doi: 10.1080/1034912X.2017.1406066
- Polo-Sánchez, M. T., & López-Justicia, M. D. (2006). Actitudes hacia las personas con discapacidad de estudiantes de la Universidad de Granada. *REOP-Revista Española de Orientación y Psicopedagogía*, 17(2), 195-211. Recuperado de <http://revistas.uned.es>
- Rodríguez-Martín, A., & Álvarez-Arregui, E. (2015). Universidad y discapacidad. Actitudes del profesorado y de estudiantes. *Perfiles Educativos*, 37(147), 86-102. doi: 10.22201/iissue.24486167e.2015.147.47265
- Salinas-Alarcón, M. (2014). *Actitudes de estudiantes sin discapacidad hacia la inclusión de estudiantes con discapacidad en la educación superior* (Tesis doctoral). Recuperado de <http://www.tdx.cat>
- Sociedad Mexicana de Psicología. (2007). *Código ético del psicólogo*. México: Trillas.
- Valdés-Cuervo, A. A., García-Vázquez, F. I., Torres-Acuña, G. M., Urías-Murrieta, M., & Grijalva-Quiñonez, C. S. (2019). *Medición en investigación educativa con apoyo del SPSS y el AMOS*. México DF, México: Clave Editorial.
- Ventura-León, J. (2018). *Otras formas de entender la d de Cohen*. *Revista Evaluar*, 18(3), 73-78. doi: 10.35670/1667-4545.v18.n3.22305
- Ventura-León, J. L., & Caycho-Rodríguez, T. (2017). Scale of attitudes towards politics among adult population in the City of Bucaramanga, Colombia. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, 12(1), 291-308. doi: 10.11600/1692715x.12117101613
- Zamora-Araya, J. A., Smith-Castro, V., Montero-Rojas, E., & Moreira-Mora, T. E. (2018). Advantages of the Rasch Model for analysis and interpretation of attitudes: The case of the Benevolent Sexism Subscale. *Revista Evaluar*, 18(3), 1-13. doi: 10.35670/1667-4545.v18.n3.22201

Apéndice 1. Escala de Actitudes hacia las Personas con Discapacidad adaptada.

Datos de identificación.

Sexo. Hombre. Mujer. Otro.

Edad en años. 18. 19. 20. 21. 22. 23 o más.

Contacto. Casi Nulo. Frecuente. Muy frecuente.

Instrucciones:

En la encuesta que presentamos a continuación se utiliza el termino personas con discapacidad para referirnos de manera general a toda persona que por razón congénita o adquirida presenta una o más deficiencias de carácter físico, mental, intelectual o sensorial, ya sea permanente o temporal y que al interactuar con las barreras que le impone el entorno social, pueda impedir su inclusión plena y efectiva, en igualdad de condiciones con los demás. Su tarea consiste en opinar si está de acuerdo o no con cada una de las frases que se le van a presentar, teniendo en cuenta que: a) No existen repuestas buenas o malas; cada opinión indica simplemente una diferente forma de pensar, b) Procure contestar todas las frases, incluso aquellas que no se ajustan a circunstancias específicas, c) En caso de duda entre varias opciones, señale aquella que se acerque más a su forma de pensar, d) Lea con atención cada frase, pero no se detenga demasiado en señalar su respuesta y e) Conteste con sinceridad. El cuestionario es totalmente confidencial y anónimo. Los significados de las opciones son las siguientes: MA Estoy Muy de Acuerdo, BA Estoy Bastante de Acuerdo, PA Estoy Parcialmente de Acuerdo, MD Estoy Muy en Desacuerdo, BD Estoy Bastante en Desacuerdo y PD Estoy Parcialmente en Desacuerdo.

1. Las personas con discapacidad con frecuencia son menos inteligentes que las demás personas.	MA	BA	PA	PD	BD	MD
2. Un trabajo sencillo y repetitivo es el más apropiado para las personas con discapacidad.	MA	BA	PA	PD	BD	MD
4. En el trabajo, una persona con discapacidad sólo es capaz de seguir instrucciones simples.	MA	BA	PA	PD	BD	MD
6. Las personas con discapacidad deberían vivir con personas afectadas por el mismo problema.	MA	BA	PA	PD	BD	MD
8. De las personas con discapacidad no puede esperarse demasiado.	MA	BA	PA	PD	BD	MD
10. Las personas con discapacidad deberían mantenerse apartadas de la sociedad.	MA	BA	PA	PD	BD	MD
14. Las personas con discapacidad también pueden casarse si lo desean.	MA	BA	PA	PD	BD	MD
15. Las personas con discapacidad deberían ser encerradas en instituciones especiales.	MA	BA	PA	PD	BD	MD
17. A las personas con discapacidad se les debería impedir votar.	MA	BA	PA	PD	BD	MD
18. Las personas con discapacidad a menudo están de mal humor.	MA	BA	PA	PD	BD	MD
19. Las personas con discapacidad confían en sí mismas tanto como las personas normales.	MA	BA	PA	PD	BD	MD
21. En el trabajo, las personas con discapacidad se entienden sin problemas con el resto de los trabajadores.	MA	BA	PA	PD	BD	MD
22. Sería apropiado que las personas con discapacidad trabajaran y vivieran con personas normales.	MA	BA	PA	PD	BD	MD

23. A las personas con discapacidad se les debería prohibir pedir créditos o préstamos.	MA	BA	PA	PD	BD	MD
24. Las personas con discapacidad generalmente son desconfiadas.	MA	BA	PA	PD	BD	MD
25. No quiero trabajar con personas con discapacidad.	MA	BA	PA	PD	BD	MD
26. En situaciones sociales preferiría no encontrarme con personas con discapacidad.	MA	BA	PA	PD	BD	MD
27. Las personas con discapacidad pueden hacer muchas cosas tan bien como cualquier otra persona.	MA	BA	PA	PD	BD	MD
28. La mayoría de las personas con discapacidad están resentidas con las personas físicamente normales.	MA	BA	PA	PD	BD	MD
30. Las personas con discapacidad son capaces de llevar a cabo una vida social normal.	MA	BA	PA	PD	BD	MD
31. Si tuviera un familiar cercano con discapacidad, evitaría comentarlo con otras personas.	MA	BA	PA	PD	BD	MD
33. La mayor parte de las personas con discapacidad sienten que son tan valiosas como cualquiera.	MA	BA	PA	PD	BD	MD
35. Se debería prevenir que las personas con discapacidad tuvieran hijos.	MA	BA	PA	PD	BD	MD
36. Las personas con discapacidad son en general tan conscientes como las personas normales.	MA	BA	PA	PD	BD	MD
37. Deberían existir leyes que prohibieran casarse a las personas con discapacidad.	MA	BA	PA	PD	BD	MD

Valores de respuesta.

1	2	3	4	5	6
MA	BA	PA	PD	BD	MD
Ítems: 1, 2, 4, 6, 8, 10, 15, 17, 18, 23, 24, 25, 28, 28, 31, 35 y 37.					
6	5	4	3	2	1
MA	BA	PA	PD	BD	MD
Ítems: 14, 19, 21, 22, 27, 30, 33 y 36					

Evidence of validity of the Remotely-administered Visual Search Task

Evidencias de validez de la Tarea de Búsqueda Visual administrada remotamente

Matías Jonás García *¹, Eliana Vanesa Zamora¹, Yesica Aydmune¹,
Ana Comesaña¹, Isabel María Introzzi¹

1 - Instituto de Psicología Básica, Aplicada y Tecnología (IPSIBAT), Universidad Nacional de Mar del Plata -CONICET. Mar del Plata, Argentina.

Introduction
Methods
Results
Discussion
References

Recibido: 10/05/2022 **Revisado:** 30/05/2022 **Aceptado:** 01/06/2022

Abstract

Perceptual inhibition is the executive process that contributes to selective attention by attenuating or reducing the interference effect generated as a result of distracting stimuli present in the environment. Due to the COVID-19 pandemic and other factors, it has been shown that there is a need for tools designed and validated in our environment that enable the assessment in non-face-to-face contexts. For this reason, this study aimed to provide empirical evidence of the internal validity and reliability of a Computerized Conjunction Visual Search (CVS) task. For this purpose, 97 adults aged between 30 and 35 ($M = 32.21$; $SD = 1.73$; 75% females) were assessed under a remote-synchronic administration. The results obtained showed adequate reliability and validity, and that the test respects the postulates of the visual search paradigm on which it is based and that its remote administration form is comfortable for users.

Keywords: *perceptual inhibition, internal validity, reliability, remote administration, conjunction visual search*

Resumen

La inhibición perceptual es el proceso ejecutivo que contribuye a la atención selectiva atenuando o disminuyendo el efecto de interferencia que generan los estímulos distractores presentes en el ambiente. Debido a la pandemia COVID-19 y a otros factores, se ha puesto de manifiesto la necesidad de contar con herramientas diseñadas y validadas en nuestro medio que permitan la evaluación en contextos no presenciales. Por este motivo, el objetivo de este trabajo fue obtener evidencias empíricas de validez interna y confiabilidad de una tarea informatizada de Búsqueda Visual Conjunta (BVC). Para ello, se realizó una evaluación a 97 personas adultas de entre 30 y 35 años ($M = 32.21$; $DS = 1.73$; 75% género femenino) bajo una modalidad de administración remota y sincrónica. Los resultados obtenidos mostraron una confiabilidad y validez adecuadas, y que la prueba respeta los postulados del paradigma de búsqueda visual en la que se sustenta y que resulta cómoda para los usuarios en su formato de administración remota.

Palabras clave: *inhibición perceptual, validez interna, confiabilidad, administración remota, búsqueda visual conjunta*

*Correspondencia a: Matías Jonás García. Tel.: +543813157302. E-mail: matiasjonasg@gmail.com

How to cite: García, M. J., Zamora, E. V., Aydmune, Y., Comesaña, A., & Introzzi, I. M. (2022). Evidence of validity of the Remotely-administered Visual Search Task. *Revista Evaluar*, 22(2), 30-46. Retrieved from <https://revistas.unc.edu.ar/index.php/revaluar>

Participaron en la edición de este artículo: Fiorella Garabano, Andrea Suárez, Eugenia Barrionuevo, Florencia Ruiz, Mónica Serppe, Ricardo Hernández, Alicia Molinari.

Introduction

The multidimensional approach to inhibition. Relevance of its assessment

Executive functions (EFs) are a top-down set of cognitive processes involved in the deliberate and voluntary control of behavior, cognition, and emotions (Diamond & Ling, 2020; Miyake & Friedman, 2012). These processes are activated in new and complex situations that demand cognitive effort where automatic, over-learned responses are insufficient (Diamond, 2013; Introzzi et al., 2020). Evidence has shown that EFs have a significant impact on quality of life. This is due to their involvement in learning and academic achievement, psychological and physical health, social life, and work efficiency, among other domains (e.g., Jacob & Parkinson, 2015; Duckworth, Taxer, Eskreis-Winkler, Galla, & Gross, 2019). The main components of EFs include working memory, cognitive flexibility and inhibitory control (Diamond, 2013; Diamond & Ling, 2020; Miyake et al., 2000). Nevertheless, in recent years it has been postulated that inhibitory control is not an unidimensional process but that related inhibitory processes which have their distinctive operational characteristics are identified (Introzzi et al., 2021). This conception of inhibitory control has been called *multidimensional approach* (e.g., Delalande et al., 2020; Introzzi, Canet-Juric, Aydmune, & Stelzer, 2016), which postulates that inhibitory control processes not only have a primary function of mitigating the effect of interference (Zamora, Richards, Canet-Juric, Aydmune, & Introzzi, 2020), but also particular characteristics that differentiate them from each other. In general, there are three inhibitory processes: perceptual inhibition, cognitive inhibition, and behavioral inhibition (e.g., Friedman & Miyake, 2004; Introzzi et al., 2021). Several studies suggest that they exhibit specific developmen-

tal trajectories (e.g., Aydmune, Introzzi, Olaechea, & López-Ramón, 2022; Gandolfi, Viterbori, Traverso, & Usai, 2014; Vadaga, Blair, & Li, 2016), that are differentially involved in numerous complex cognitive functions and activities of daily living (e.g., Borella, Carretti, & Pelegrina, 2010), and that are particularly impaired in various psychopathological disorders (e.g., Christ, Kester, Bodner, & Miles, 2011; Mammarella et al., 2017) and neurodegenerative diseases (Hazlett, Figueroa, & Nielson, 2015). This shows the necessity and relevance of using techniques designed and validated in our region that allows us to assess each inhibitory process independently.

Cognitive Self-regulation Tasks: An assessment approach from a multidimensional perspective in a remote administration modality

The assessment of inhibitory processes from a multidimensional perspective is not an easy challenge. The proposed tasks must be attractive, easy to understand, presented analogously, and they must not be too long, to avoid the fatigue of the person being assessed. The literature often reports tasks with a large number of trials (e.g., Logan, 1994; Oberauer, 2001), which, coupled with the fact that inhibitory tasks (and executive tasks in general) require significant cognitive effort, can generate fatigue and affect participant performance (Aydmune & Introzzi, 2018). In addition, due to the particularities of each inhibitory process, they must show high demand or requirements for the process, with minimal or low demand for another, or other executive functions (e.g., cognitive flexibility or working memory). In this sense, the Cognitive Self-regulation Tasks -TAC- set (Introzzi & Canet-Juric, 2019) is a technological innovation for several reasons. First, since the beginning, it has been

designed according to a multidimensional model that assumes not only the existence of different EFs (working memory, inhibition, and cognitive flexibility) but also of different inhibitory processes. Therefore, for the design and proposal of activities for each of the tasks, widely used and validated experimental paradigms have been selected, which allow for the specific evaluation of each EF (working memory, cognitive flexibility, and inhibitory processes: behavioral, cognitive, and perceptual inhibition). Second, the TAC is a fully computerized tool. All instances of the assessment process (socio-demographic data entry, test administration, and data collection for scores) are performed automatically using a computer with internet access (Drasgow & Mattern, 2006). This provides certain benefits in comparison to pencil-and-paper tests, including greater control and accuracy in item administration, greater efficiency in scoring and response storage, faster and more efficient psychological reporting, less chance of data entry error, and a more motivating environment and appearance (Medrano & Pérez, 2018). Third, due to the nature of the tasks proposed and the responses required, all of the TAC tasks can be used to assess children from 6 years old to adults over 80 years old. And fourth, due to the characteristics previously mentioned, the TAC test could be used in a remote psychological assessment context, which means that it does not require the direct, face-to-face intervention of the tester. Remote psychological assessment is relevant in the studies with mobility-restricted populations or people who reside far away and have difficulty traveling to the clinic. The recent pandemic context has also demonstrated the importance of this type of assessment (Brearly et al., 2017; Marra, Hoelzle, Davis, & Schwartz, 2020). Although there is a need for neuropsychological assessment tools that support this mode of administration (e.g., Brearly et al., 2017), no studies

have yet been conducted in this regard. Thus, this study aims to analyse the psychometric properties of scores on the remotely administered version of the Visual Search task for the assessment of perceptual inhibition in a population of young adults

Conjunction Visual Search: The TAC task for the assessment of perceptual inhibition

Perceptual inhibition (PI) is the inhibitory process involved in the initial stages of information processing, reducing the arousal level generated by distracting environmental stimuli that interfere with the ongoing task. Hence, it helps to control the input of irrelevant information into consciousness or attentional focus (Friedman & Miyake, 2004; Hasher, Zacks, & May, 1999). It is therefore considered an essential component of selective attention, as it permits relevant information to be highlighted, increasing processing efficiency (e.g., Treisman & Gelade, 1980; Treisman & Sato, 1990). So, the more efficient the inhibitory process is, the faster and more effective the selective attention is supposed to be. Moreover, PI is also strongly implicated in several complex cognitive functions such as mathematics performance (Stolte et al., 2019), reading comprehension (Borella et al., 2010; Borella & de Ribaupierre, 2014; Stevens & Bavelier, 2012), fluid intelligence (Aydumne, Introzzi, & Zamora, 2020; Darowski, Helder, Zacks, Hasher, & Hambrick, 2008; Stelzer, 2014), planning, and cognitive flexibility (Davidson, Amso, Anderson, & Diamond, 2006; Diamond, 2013, 2016).

For the foregoing reasons, we consider that it is important to develop tools designed and validated in our environment which enable us to assess this process specifically. TAC's Visual Search (VS) is a task that serves this purpose. Although it has sufficient empirical evidence of validity and

reliability in the face-to-face administration modality (e.g., Comesaña, Richard's, & Vido, 2019; Richard's et al., 2017a), its psychometric properties have not been analyzed in the remote administration modality.

Evidence of validity and reliability of Visual Search scores in a remote administration version

The VS task has been designed based on the Conjunctive Visual Search (CVS) paradigm proposed by Treisman and Gelade (1980) within the integrative feature theory approach (e.g., Treisman & Gelade, 1980; Treisman & Sato, 1990; Introzzi et al., 2017). In this paradigm, participants must identify the presence or absence of a target -blue square- that is presented among a variable set of distractor stimuli -red squares and blue circles. The stimuli consist of double conjunctions, defined by the combination of two visual features: shape and color. Furthermore, all distractors have one of these visual characteristics in common with the target, a condition that guarantees the visual interference effect and thus the involvement of the PI. The proposed activity is simple, the participant must press one of two keys as quickly as possible depending on the presence or absence of the target (See methodology for more details).

The CVS paradigm suggests the existence of two main effects widely replicated in the literature and based on feature integration theory: *the presence or absence effect of the target* and *the number of distractors effect*. The presence or absence effect of the target is characterized by better performance in trials or conditions in which the target is present compared to those in which it is absent. Prediction assumes the existence of an exhaustive, sequential and obligatory search. In other words, it proceeds in the visual scene by checking one element at a time and it is complet-

ed when the target is detected. Thus, when the target is absent, it is assumed that all elements must be examined; whereas when the target is present it is necessary to examine, on average, half of the elements of the visual presentation before identifying the target. In consequence, according to this proposal, performance in tests where the target is present is expected to be significantly lower than in those where it is absent. The effect of the quantity of distractors is characterized by a decrease in search performance for conjunctions (e.g., a blue square target, between red squares and blue circles, distractors) that depends on the number of distractors added. It is assumed that this decrease is more pronounced as the number of distractors increases, and it is explained by the effect of visual interference (Treisman & Sato, 1990; Introzzi et al., 2017).

Consequently, scores obtained with a task designed based on the CVS paradigm should be consistent with these two main effects, which would provide evidence of the construct (internal) validity of the test. In the VS task, this kind of validity evidence has been obtained in a face-to-face administration modality in typically (Richard's et al., 2017a) and atypically (Richard's et al., 2017b) developing children, adolescents, adults and older adults (Introzzi et al., 2020, 2021; Richard's, Introzzi, Zamora, & López-Ramón, 2022). However, the literature emphasizes that researchers should not assume the equivalence of two different forms or versions of the same tool (Medrano & Pérez, 2018). Consequently, it should not be assumed that the reliability and validity evidence of the instrument obtained in a face-to-face administration context will also be found in the virtual or remote administration modality. Therefore, the present study seeks to obtain empirical evidence of VS scores in a remote administration modality in adults so that it can be used in assessment processes in different contexts

of psychology and other disciplines in the areas of health and education. This aim is part of a larger project to analyze the psychometric properties of different TAC tests in a remote administration modality in different age groups. In this research, it was decided to focus on young adults, as this population is familiar with the use of computers and the internet. In addition, at this age, executive functioning reaches a plateau in its development and its maximum level, while impairment of other functions is usually not observed. Although TAC can be administered without great difficulty in children and older adults, it is considered more appropriate to begin to explore remote administration with young adults.

Methods

Participants

We conducted this study with an instrumental design (Montero & León, 2002) using an independent, non-probabilistic, intentional, snowball sample of adults ($N = 97$), residents of Mar del Plata, between 30 and 35 years old ($M = 32.21$; $SD = 1.73$), 51.5% of whom identified themselves as female. In terms of educational level, all participants had at least a high school education. 71% reported finishing their university studies and 21% reported incomplete university studies.

The inclusion of participants in the sample was decided based on information obtained in a short ad hoc survey designed for this research (see procedure). In addition, inclusion criteria were considered as follows: 30-35 years old, no neurological and/or psychiatric diagnosis, 12 years or more of formal education, normal or corrected vision and hearing (conditions necessary to perform the proposed activities), a desktop or laptop computer with a 14" or larger screen and stable internet connection of more than 1 megabyte. We

excluded participants with psychopathological symptoms or those under psychopharmacological treatment, and with uncorrected visual or hearing difficulties.

Procedure

To constitute the sample, community members aged between 30 and 35 were invited to participate. A "snowball sampling" or "participant-driven sampling" was conducted (Heckathorn, 1997; Watters & Biernacki, 1989). First, a Google form was distributed via email and social networks (Facebook, Twitter, WhatsApp, Telegram and Instagram) inviting people to participate in the study, explaining the main objectives, requirements, and activities involved, the voluntary nature of participation, and the confidentiality of the results and the information obtained. Those interested were subsequently contacted and asked to sign an informed consent form as a fundamental condition for their participation in the study. In the same form, a short ad hoc survey was presented to collect data about inclusion criteria, and the Symptom Assessment-45 Questionnaire (SA-45) was presented in its Spanish version, to collect information about psychopathological symptomatology. In the next contact, a video conference meeting was arranged in which the VS task and the questionnaire were administered.

Instruments

Visual Search Task. The VS task is part of the computer-based system TAC (Introzzi & Canet-Juric, 2019) for the assessment of executive functioning (working memory, inhibitory processes, and cognitive flexibility). This task has been designed based on the CVS paradigm (Treisman

& Gelade, 1980). It consists in determining the presence or absence of a target stimulus (blue square) that is presented mixed among a variable set of distractor stimuli (red squares and blue circles). The stimuli consist of double conjunctions, which are defined by the combination of two visual characteristics: shape and color. In addition, all distractors share one of these visual features with the target, a condition that guarantees the visual interference effect and thus the engagement of the PI.

The task includes a block of 10 practice trials, followed by three blocks of 40 trials respectively. Every block contains 10 trials for condition (according to the number of distractors = 4, 8, 16, or 32). The 40 trials were randomly distributed in every block; 50% of trials in each block presented the target and the rest were absent. In each trial, the participant has to provide an answer, whether affirmative or negative, as quickly and accurately as possible, by pressing the appropriate key (“Z” key if the target is present and “M” key if the target is absent). After the answer is given, the following trial is shown.

Performance measures. Task performance is obtained through the percentage of correct responses (response accuracy) and the mean response time (RT) for each distractor condition. It is important to note that the register of RT is obtained just for correct answers. When performance is assessed through the RTs and accuracy independently, differences in performance can result from the speed of response, accuracy, or the interaction between the two factors. Therefore, to obtain a better index of performance in these tasks, a single measure combining both factors has been proposed. The inverse efficiency index (EI), results from the ratio of RT and accuracy (Christie & Klein, 1995). This index compensates for the relationship between speed and accuracy, which is why several studies have chosen to include it as one of the main performance indicators (e.g., Comesaña et al., 2019; Introzzi et al., 2020; Zamora et al., 2020). Its interpretation is simple: higher indexes (expressed in RT) suggest a worse performance on the task.

In addition, the task has a section to assess the speed of processing (SP) which is administered immediately before or independently of the per-

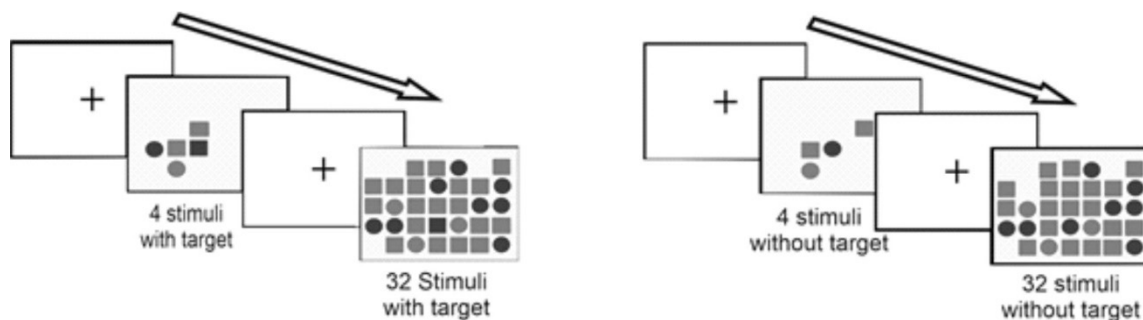


Figure 1

Visual Search. The figure (left) shows two consecutive trials. The first one with 4 distractors and the second one with 32 distractors. In both trials the target is present. The figure on the right shows two consecutive trials where the target is absent. The first with 4 and the second with 32 distractors.

ceptual inhibition task. The SP task is similar in every aspect to the inhibition task (task instructions, stimuli, presentation, and response keys) except that no distractors are presented (e.g., in this task there is only one stimulus per trial). Therefore, it provides a measure of SP but also functions as a baseline score for the computation of a PI index (Introzzi et al., 2020). This section consists of 10 practice trials and 20 assessment trials, which allow the recording of the two main performance indices: RT and accuracy. As in the conjunction search task, 50% of the trials presented the blue square and the other 50% a red square or a blue circle.

Given the absence of distractors, the task allows the intervention of other processes such as visual search and PI to be minimized, ensuring that SP can be assessed with minimal intervention from them. The literature also recommends the use of difference scores to quantify the interference produced by more difficult task conditions (Mullane et al., 2009). For computing these scores, the mean RT of trials with interference is usually subtracted from those where there is little or no interference. Without such a subtraction a high or low score could be attributed to the speed or slowness of the subject's response. In the present task, the difference between the mean RT in the no-distractor condition (SP task) and the mean RT in the 4-distractor condition was used to obtain this index. Thus, higher scores indicate less efficient interference control. Table 1 presents the description and abbreviation of the performance indices used in this study.

Ad-hoc questionnaire to assess the functioning of the Visual Search task in the synchronous remote administration mode. A short and simple ad-hoc questionnaire was designed and administered to participants to collect data on several issues related to the general functioning of the VS

task in the remote administration mode and some specific characteristics of the task.

One set of items was constructed to obtain information about the difficulty perceived regarding the actions that participants had to take, after the administration of VS task. This set of administrative procedures is normally done by the examiner in the face-to-face administration mode, but in the remote mode, it must be done by the person being assessed. Specifically, questions were asked about access to the web page on which TAC and VS tasks are inserted; access using the username and password provided by the researcher; search and selection using their full name (previously uploaded in the system). Participants recorded their answers on a Likert scale where they had to mark the degree of difficulty (low, medium, or high).

Other items were designed to explore a set of questions related to understanding the instructions, the type of activity they were asked to do and the type of response required. Therefore, the following questions were asked: *Were you able to understand what you were asked to do in the task? Did you find the practice instance of the task useful? Did it allow you to understand better what you were asked to do? Were you able to visualize the indications and text presented on the screen adequately? Were you able to distinguish the stimuli presented on the screen clearly?* To answer these questions, participants had to choose between two response options: *Yes* or *No*. In addition, they had to answer whether they found the indication or signalling regarding the keys to be pressed to give their answer easily, understandable or confusing.

To analyze perceived effort related to the different conditions of the task and the opinion regarding length or extension, four items were elaborated. Concerning the difficulty level of the task, we asked whether the difficulty increased

Table 1
Main Performance indexes of the Visual Search task.

Description of index and abbreviation	Index abbreviation	Index calculation
Medium RT in the 4-distractor condition	TR4	
Medium RT in the 8-distractor condition	TR8	
Medium RT in the 16-distractor condition	TR16	
Medium RT in the 32-distractor condition	TR32	
Response accuracy in the 4-distractor condition	PREC4	Not applicable
Response accuracy in the 8-distractor condition	PREC8	
Response accuracy in the 16-distractor condition	PREC16	
Response accuracy in the 32-distractor condition	PREC32	
Inverse efficiency index 4-distractors condition	EI4	TR4/Prec4
Inverse efficiency index 8-distractors condition	EI8	TR8/Prec8
Inverse efficiency index 16-distractors condition	EI16	TR16/Prec16
Inverse efficiency index 32-distractors condition	EI32	TR32/Prec32
Difference between medium RT in the SP Task and medium RT in the 4-distractor condition	IDif	TRmedioVP - TR4

progressively throughout the task, whether all parts of the task were of equal difficulty, whether the difficulty decreased as the task progressed, and whether some parts of the task were more difficult than others but were mixed. Concerning cognitive effort, we asked whether all parts of the task required the same cognitive effort or whether the effort was related to the number of stimuli (more stimuli, more effort; fewer stimuli, less effort). Finally, perceived accuracy and time were assessed visually. The subject had to choose between four boxes showing the four test conditions. Most subjects stated that they were more accurate in the conditions with fewer stimuli and

that the conditions with more stimuli took longer (Figure 2), which was in line with the assumptions of the baseline paradigm.

Data analysis

First, the internal consistency of the task was estimated by the split-half method, using the Spearman-Brown correction (Cohen & Swerdluk, 2009). The reliability of RT was calculated for each block. Thus, the consistency within each block was estimated and then the results were averaged to obtain the total reliability, with the

Table 2

Descriptive statistics of the main indexes of the Visual Search task.

	M	SD	Skewness	Kurtosis	Kolmogorov-Smirnov
Response accuracy with no distractors	94.69	13.06	-4.1	19.22	.35*
Response accuracy in the 4-distractor condition	97.87	2.50	-1	0.54	.30*
Response accuracy in the 8-distractor condition	97.93	3.51	-2.9	13.66	.34*
Response accuracy in the 16-distractor condition	96.90	3.47	-0.9	0.27	.25*
Response accuracy in the 32-distractor condition	92.44	7.01	-1	0.47	.19*
Medium RT with no distractor	786.46	173.73	1	3.62	.08
Medium RT in the 4-distractor condition	844.70	135.11	0.7	1.43	.05
Medium RT in the 8-distractor condition	901.44	140.73	0.3	0.16	.06
Medium RT in the 16-distractor condition	1040.67	193.25	0.9	1.57	.09*
Medium RT in the 32-distractor condition	1348.31	295.82	1	1.01	.13*
Difference in accuracy between the 4 and 32-distractor condition	5.4305	7.09	0.9	0.68	.19*
Difference in RT between the 4 and 32-distractor condition	503.61	222.33	1.2	1.18	.13*
Reverse efficiency without distractors	8.69	3.89	5	30.02	.25*
Inverse efficiency with 4 stimuli	8.62	1.33	0.6	1.24	.05
Inverse efficiency with 8 stimuli	9.21	1.43	0.3	-0.07	.07
Inverse efficiency with 16 stimuli	10.75	2.05	1	2.06	.10*
Inverse efficiency with 32 stimuli	14.61	3.09	0.9	0.98	.15*

Note. * $p < .05$

understanding that each block has the same conditions (4, 8, 16 and 32 distractors).

To obtain internal validity evidence for VS in the remote administration modality, it was analyzed whether the results allowed verifying the

presence of the two main experimental effects linked to the Conjunction Search paradigm derived from the feature integration theory (e.g., [Treisman & Sato, 1990](#)) based on which the task was designed. For this purpose, the distribution

of the data was analyzed according to centrality, skewness, and kurtosis (Table 2). When it was confirmed that in general, they do not adjust to a normal distribution, we chose to use the Friedman test for related samples, with the Wilcoxon rank test to analyze post hoc pairwise analysis to contrast the two main assumptions of the task.

Furthermore, because the change in the mode of administration (face-to-face or remote) could affect the psychometric properties (validity, reliability, norms) of the tests (Medrano & Pérez, 2018), the responses to the ad hoc questionnaire (see Instruments) were analyzed by calculating the percentages obtained for each response mode.

Ethical considerations

The implemented protocols were approved by an ethics committee (in this first version of the manuscript its name has been omitted to avoid providing data that could identify the authors) and all the established ethical guidelines were respected. Moreover, the ethical principles and guidelines outlined in the code of conduct of the American Psychological Association (APA, 2017) for the implementation of scientific research in psychology were followed. After reading the information about the project, clarifying their doubts, and signing the informed consent form, people could access the assessment instruments and participate in the study. As mentioned before, participation was voluntary and could be interrupted at any time without giving reasons. The information obtained in the tasks and questionnaires was treated confidentially to protect the identity of the participants. All personal data were associated with an alphanumeric code that was used in the database. In addition, the information was stored on servers and computers that satisfy computer security requirements. The data

collected is used only for academic and scientific purposes by National Law 25.326 on the protection of personal data.

Results

Evidence of validity

Internal validity. The task paradigm (Treisman & Gelade, 1980) is based on two fundamental premises. Analyses of these are developed below:

Presence/absence effect of the target. Medium RT was analyzed in the conditions in which the target is absent and present. These data were compared in each of four conditions (4, 8, 16, and 32) in the task. For this, Wilcoxon signed-rank test was used. The effect size was calculated for each pair by dividing the z-score obtained by the square root of the number of observations (Palant, 2007). The results show that there would be significant differences between the RTs that contain the target and those that do not, in all four conditions (Table 3). In addition, the effect size observed increases as the number of stimuli becomes larger.

Effect of the number of distractors. To test this hypothesis, the study was conducted using mean RTs, accuracy and the IE index for each condition. Friedman's test for related samples was applied, with Wilcoxon's rank test to analyze post hoc pairwise behavior. Accuracy and RT were compared for no-distractor (SP) condition and the four test conditions, finding significant differences in accuracy ($Q(4) = 74.87; p < .05$) and RTs ($Q(4) = 326.4; p < .05$). Next, the same test was performed for the inverse efficiency indices in the four conditions and the no-distractor condition, with significant differences ($Q(4) = 302.8; p < .05$). Subsequently, Wilcoxon's rank test was ap-

Table 3

Wilcoxon signed-rank test (W) for RT by presence or absence of the target.

Target	4				8			
	M	SD	W	r	M	SD	W	r
Present (RT)	826.89	145.51	-2.62*	.26	865.89	130.29	-5.34*	.54
Absent (RT)	862.34	156.43			937.28	171.3		
Target	16				32			
	M	SD	W	r	M	SD	W	r
Present (RT)	945.64	146.50	-7.81*	.79	1123.1	212.16	-8.38*	.85
Absent (RT)	1133.4	266.39			1551.4	429.1		

Note. * $p < .05$; r = effect size.

plied to observe pairwise associations (Table 4). Regarding accuracy, the Wilcoxon test showed that there would be no significant associations in the pairs with no distractors-C16 and C4-C8, i.e., the difference in medians between the results of these conditions would be due to a random chance and cannot be associated with a tendency. Accuracy, RTs, and inverse efficiency indices for all other conditions were shown to follow a trend. Additionally, the effect size was notably larger for reaction times than for accuracy.

Remote assessment access questionnaire. The questionnaire was designed to provide evidence of construct validity and consisted of a series of items related to the level of difficulty in accessing the platform and understanding the instructions, as well as the extent and perceived effort during the task.

The perceived difficulties of accessing the platform, access to the website, logging in with username and password, search, and selection of the participant by full name were specifically evaluated using an ordinal scale with three response categories (*high*, *medium*, or *low*). In general, the perceived difficulty was *low*: 91.9% for the difficulty of accessing the page; 96.5% for the login with username and password, as well as for the search for users by name and surname.

For questions aimed at obtaining comprehension information on the different instances of the task, a dichotomous scale (yes/no) was used. Most subjects responded affirmatively to questions related to the instructions (100%); the practice instance (97.7%); the indications during the test (100%); the stimuli presented (97.7%); and the indications regarding the keys to be used (97.7%). Regarding the task length, it was rat-

Table 4

Wilcoxon signed-rank test for accuracy, RT, and inverse efficiency index, according to the number of distractors.

Pairs	Accuracy		RT		Inverse efficiency index	
	W	<i>r</i>	W	<i>r</i>	W	<i>r</i>
ND – C4	-2.04**	.20	-4.68**	.47	-3.12**	.31
ND – C8	-2.36**	.23	-6.63**	.67	-5.43**	.55
ND – C16	-0.19	.01	-8.04**	.81	-7.3**	.74
ND – C32	-3.7**	.30	-8.5**	.86	-7.82**	.79
C4 - C8	-0.32	.03	-6.84**	.69	-6.34**	.64
C4 - C16	-2.49**	.25	-8.54**	.86	-8.54**	.86
C4 - C32	-6.61**	.67	-8.55**	.86	-8.55**	.86
C8 - C16	-2.29**	.23	-8.48**	.86	-8.32**	.84
C8 - C32	-6.62**	.67	-8.55**	.86	-8.55**	.86
C16 - C32	-5.78**	.58	-8.55**	.86	-8.55**	.86

Note. ** $p < .05$; r = effect size; ND = no-distractor condition.

ed by 62.8% of the participants as *short* and by 37.2% as *not too short and not too long*, on a three-choice ordinal scale, where no participant answered that the task was *too long*.

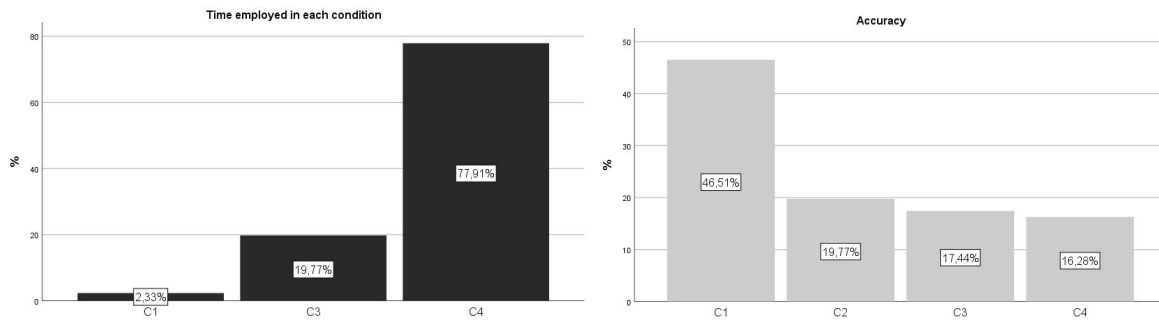
As regards perceived effort in the task, the following were assessed: perception of cognitive effort about the number of stimuli, perception of the accuracy of their response in the different conditions, and perception of the time required to complete the task under the different conditions. Regarding task progression, 64% perceived that *the level of difficulty of the task increases progressively as the task progresses* and 31.4% that *some parts of the task are more difficult than others and are mixed throughout the task* (corresponding to the correct answer in this case). Regarding cognitive effort, 70% indicated that *more stimuli re-*

quired more attention or cognitive effort (which is in line with the visual search paradigm used) and 24.4% indicated that *all parts of the task required the same attention or cognitive effort*.

Perceived accuracy and time were assessed visually. The subjects had to choose between four boxes showing the four test conditions. Most subjects stated that they were more accurate in the conditions with fewer stimuli and that the conditions with more stimuli took longer (Figure 2), in line with the premises of the visual search paradigm.

Reliability

RT reliability was estimated using a split-half method with Spearman-Brown correction.

**Figure 2**

Perception of accuracy and performance time in each visual search task condition.

For this purpose, two computations were made. First, internal consistency was calculated for all items without discriminating the block to which they belong ($r_s = .94$). Second, the internal consistency of each block was calculated. Each block contained the same number of stimuli belonging to the four conditions (4-8-16-32). All blocks exhibited a consistency similar to the total (block 1, $r_s = .88$; block 2, $r_s = .83$; block 3, $r_s = .78$). Finally, the consistency of all blocks was averaged ($r_s = .83$). The results obtained for the RTs can be interpreted as excellent for all scores, except for block 3, where substantial consistency was found (Cicchetti & Sparrow, 1981; Fleiss, 1981).

Discussion

This study aimed to obtain evidence of construct validity of the VS task (Introzzi & Canet-Juric, 2019), using a remote administration modality. The study was conducted with a sample of adults aged between 30-35 who were administered the task and an ad hoc questionnaire. Although the task has been adapted and validated in the local population, no studies of its remote administration have been recorded. To achieve this purpose, firstly, evidence was obtained on the

reliability of the task in its remote administration. Secondly, to obtain evidence of construct validity, the two main postulates of Treisman and Gelade's (1980) visual search paradigm were tested.

The first criterion of the paradigm indicates that the average RTs increase, and the percentage of correct responses decrease as the number of distractors increases. The second criterion indicates that a higher RT mean is to be expected when the target is absent compared to when it is present. The results showed that as the number of distractors increases, the RTs also increase, while the percentage of correct answers tends to decrease. Regarding the results showing differences when the target is present versus absent, some researchers have interpreted the increase in RT related to the number of distractors as evidence that the search progresses stimulus by stimulus until the target is located (Treisman, 1988; Treisman & Gelade, 1980). The observed difference in the results would indicate that on trials where the target is absent, the subject examines each item to confirm that the target is not present. In contrast, on trials where the target is present, the subject must examine on average only half of the items to locate the target (Wolfe, Cave, & Franzel, 1989). The findings of the remote task seem to be consistent with the empirical evidence so far, as in

each condition subjects gave faster responses in presence of the target than in absence of the target. Moreover, the task generated higher interference when the number of distractors was higher (see review [Richard's et al., 2022](#)). The analysis of the ad hoc questionnaire answers suggests that participants had neither difficulty accessing the platform and the task, nor in understanding the instructions and prompts during the task.

As mentioned above, it is important to have a reliable and valid set of instruments to adequately assess PI given its importance in different domains. Due to the pandemic, many tools for assessing cognitive processes -especially those based on pencil and paper- have been insufficient, given the impossible nature of face-to-face encounters. In the modern world and with IT development, we consider essential the adjustment of tests to a remote administration environment. In this respect, we found evidence of validity for the VS task which, although computerized, had no remote administration studies.

However, this study was not free of limitations. First, the generalizability of these results is limited, mainly due to the age range of the sample (30-35 years). The remote administration of the task may be more challenging in children and older adults, due to the technical handling of the test. It is expected that the sample age range will be extended in the future. Furthermore, due to the characteristics of non-probability sampling, the results are only partially generalizable.

Second, this study is solely limited to internal validity. We did not consider analyzing other types of validity or reliability because this test does not present changes in its conformation, but just in the way it is administered.

In conclusion, we consider that the VS task satisfies the validity criteria to be administered remotely. There is still a need to expand the sample groups in which it can be applied, given the

particularities of the lifespan that especially affect the EFs.

References

- American Psychological Association. (2017, 1st January). *Ethical principles of psychologists and code of conduct*. Retrieved from <https://www.apa.org/ethics/code>
- Aydmune, Y., & Introzzi, I. (2018). Inhibición: Una función ejecutiva difícil de medir. Algunas problemáticas en relación con las pruebas de inhibición informatizadas. *Psicodebate*, 18(2), 7-25. doi: [10.18682/pd.v18i2.741](https://doi.org/10.18682/pd.v18i2.741)
- Aydmune, Y. S., Introzzi, I. M., Olaechea, M. C., & López-Ramón, M. F. (2022). Inhibitory processing training during childhood: A review – Processes-based inhibitory training. In F. Alcantud-Marín, M. López-Ramón, E. Navarro-Pardo, V. Moreno-Campos & Y. Alonso-Esteban (Ed.), *Handbook of Research on Neurocognitive Development of Executive Functions and Implications for Intervention* (pp. 50-77). doi: [10.4018/978-1-7998-9075-1.ch003](https://doi.org/10.4018/978-1-7998-9075-1.ch003)
- Aydmune, Y., Introzzi, I., & Zamora, E. V. (2020). Tarea de interferencia proactiva (IP) para la medición de la inhibición cognitiva en niños de 6 a 8 años. *Revista Evaluar*, 20(3), 34-50. doi: [10.35670/1667-4545.v20.n3.31710](https://doi.org/10.35670/1667-4545.v20.n3.31710)
- Borella, E., Carretti, B., & Pelegrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities*, 43(6), 541-552. doi: [10.1177/0022219410371676](https://doi.org/10.1177/0022219410371676)
- Borella, E., & de Ribaupierre, A. (2014). The role of working memory, inhibition, and processing speed in text comprehension in children. *Learning and Individual Differences*, 34, 86-92. doi: [10.1016/j.lindif.2014.05.001](https://doi.org/10.1016/j.lindif.2014.05.001)
- Brearly, T. W., Shura, R. D., Martindale, S. L., Lazowski, R. A., Luxton, D. D., Shenal, B. V., & Rowland, J.

- A. (2017). Neuropsychological test administration by videoconference: A systematic review and meta-analysis. *Neuropsychology Review*, 27(2), 174-186. doi: [10.1007/s11065-017-9349-1](https://doi.org/10.1007/s11065-017-9349-1)
- Christ, S. E., Kester, L. E., Bodner, K. E., & Miles, J. H. (2011). Evidence for selective inhibitory impairment in individuals with autism spectrum disorder. *Neuropsychology*, 25(6), 690-701. doi: [10.1037/a0024256](https://doi.org/10.1037/a0024256)
- Christie, J., & Klein, R. (1995). Familiarity and attention: does what we know affect what we notice? *Memory & Cognition*, 23, 547-550. doi: [10.3758/BF03197256](https://doi.org/10.3758/BF03197256)
- Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*, 86(2), 127-137.
- Cohen, R. J., & Swerdlik, M. E. (2009). *Psychological Testing and Assessment: An Introduction to Tests and Measurement* (7th ed.). New York, NY: McGraw-Hill.
- Comesaña, A., Richard's, M. M., & Vido, V. (2019). Comparative analysis of the perceptual inhibition between children and older adults. *Psychology & Neuroscience*, 12(1), 65-77. doi: [10.1037/pne0000167](https://doi.org/10.1037/pne0000167)
- Darowski, E. S., Helder, E., Zacks, R. T., Hasher, L., & Hambrick, D. Z. (2008). Age-related differences in cognition: The role of distraction control. *Neuropsychology*, 22(5), 638-644. doi: [10.1037/0894-4105.22.5.638](https://doi.org/10.1037/0894-4105.22.5.638)
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive function from 4 to 13 years: Evidence from manipulations of memory, inhibitions and tasks switching. *Neuropsychologia*, 44(11), 2037-2078. doi: [10.1016/j.neuropsychologia.2006.02.006](https://doi.org/10.1016/j.neuropsychologia.2006.02.006)
- Delalande, L., Moyon, M., Tissier, C., Dorriere, V., Guillois, B., Mevell, K., ... & Borst, G. (2020). Complex and subtle structural changes in prefrontal cortex induced by inhibitory control training from childhood to adolescence. *Developmental Science*, 23(4), e12898. doi: [10.1111/desc.12898](https://doi.org/10.1111/desc.12898)
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135-168. doi: [10.1146/annurev-psych-113011-143750](https://doi.org/10.1146/annurev-psych-113011-143750)
- Diamond, A. (2016). Why improving and assessing executive functions early in life is critical. In J. A. Griffin, P. McCardle & L. S. Freund (Eds.), *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research* (pp. 11-43). Washington, DC: American Psychological Association. doi: [10.13140/RG.2.1.2644.6483](https://doi.org/10.13140/RG.2.1.2644.6483)
- Diamond, A., & Ling, D. S. (2020). Review of the evidence on, and fundamental questions about, efforts to improve executive functions, including working memory. In J. M. Novick, M. F. Bunting, M. R. Dougherty & R. W. Engle (Eds.), *Cognitive and working memory training: Perspectives from psychology, neuroscience, and human development* (pp. 143-431). doi: [10.1093/oso/9780199974467.003.0008](https://doi.org/10.1093/oso/9780199974467.003.0008)
- Drasgow, F., & Mattern, K. (2006). New tests and new items: Opportunities and issues. In D. Bartram & R. Hambleton (Eds.), *Computer-based testing and the Internet: Issues and advances* (pp. 59-75). New Jersey, NJ: John Wiley & Sons. doi: [10.1002/9780470712993.ch3](https://doi.org/10.1002/9780470712993.ch3)
- Duckworth, A. L., Taxer, J. L., Eskreis-Winkler, L., Galla, B. M., & Gross, J. J. (2019). Self-control and academic achievement. *Annual Review of Psychology*, 70(1), 373-399. doi: [10.1146/annurev-psych-010418-103230](https://doi.org/10.1146/annurev-psych-010418-103230)
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York, NY: John Wiley.
- Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, 133(1), 101-135. doi: [10.1037/0096-3445.133.1.101](https://doi.org/10.1037/0096-3445.133.1.101)
- Gandolfi, E., Viterbori, P., Traverso, L., & Usai, M. C. (2014). Inhibitory processes in toddlers: A latent-variable approach. *Frontiers in Psychology*, 5. doi: [10.3389/fpsyg.2014.00381](https://doi.org/10.3389/fpsyg.2014.00381)
- Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher &

- A. Koriat (Eds.), *Attention and Performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 653-675). Cambridge, MA: MIT Press.
- Hazlett, K. E., Figueroa, C. M., & Nielson, K. A. (2015). Executive functioning and risk for Alzheimer's disease in the cognitively intact: Family history predicts Wisconsin Card Sorting Test performance. *Neuropsychology*, 29(4), 582-591. doi: [10.1037/neu0000181](https://doi.org/10.1037/neu0000181)
- Heckathorn, D. D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems*, 44(2), 174-199. doi: [10.2307/3096941](https://doi.org/10.2307/3096941)
- Introzzi, I., & Canet-Juric, L. (2019). Tareas de Autorregulación Cognitiva [Software y manual de usuario]. Retrieved from <https://tac.com.ar/evaluacion>
- Introzzi, I. M., Canet-Juric, L., Aydmune, Y., & Stelzer, F. (2016). Perspectivas teóricas y evidencia empírica sobre la inhibición. *Revista Colombiana de Psicología*, 25(2), 351-368. doi: [10.15446/rcp.v25n2.52011](https://doi.org/10.15446/rcp.v25n2.52011)
- Introzzi, I. M., Richard's, M. M., Aydmune, Y., Zamora, E. V., Stelzer, F., García-Coni, A., ..., & Navarro-Pardo, E. (2021). Development of perceptual inhibition in adolescents, a critical period? *Symmetry*, 13(3), 457. doi: [10.3390/sym13030457](https://doi.org/10.3390/sym13030457)
- Introzzi, I., Zamora, E., Aydmune, Y., Canet Juric, L., & López, S. (2017). El rol de la inhibición en la Teoría de Integración de las Características. *Cuadernos de Neuropsicología*, 11(3), 135-150. doi: [10.7714/CNPS/11.3.208](https://doi.org/10.7714/CNPS/11.3.208)
- Introzzi, I., Zamora, E., Aydmune, Y., Richard's, M., Comeña, A., & Canet-Juric, L. (2020). The change processes in selective attention during adulthood. Inhibition or processing speed? *The Spanish Journal of Psychology*, 23, e37. doi: [10.1017/SJP.2020.41](https://doi.org/10.1017/SJP.2020.41)
- Jacob, R., & Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Review of Educational Research*, 85(4), 512-552. doi: [10.3102/0034654314561338](https://doi.org/10.3102/0034654314561338)
- Logan, G. D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189-239). San Diego, CA: Academic Press.
- Mammarella, I. C., Caviola, S., Giofrè, D., & Borella, E. (2017). Separating math from anxiety: The role of inhibitory mechanisms. *Applied Neuropsychology: Child*, 7(4), 342-353. doi: [10.1080/21622965.2017.1341836](https://doi.org/10.1080/21622965.2017.1341836)
- Marra, D. E., Hoelzle, J. B., Davis, J. J., & Schwartz, E. S. (2020). Initial changes in neuropsychologists clinical practice during the COVID-19 pandemic: A survey study. *The Clinical Neuropsychologist*, 34(7-8), 1251-1266. doi: [10.1080/13854046.2020.1800098](https://doi.org/10.1080/13854046.2020.1800098)
- Medrano, L., & Pérez, E. (2018). *Manual de psicometría y evaluación psicológica*. Córdoba, Argentina: Brujas.
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21(1), 8-14. doi: [10.1177/0963721411429458](https://doi.org/10.1177/0963721411429458)
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100. doi: [10.1006/cogp.1999.0734](https://doi.org/10.1006/cogp.1999.0734)
- Montero, I., & León, O. G. (2002). Clasificación y descripción de las metodologías de investigación en Psicología. *International Journal of Clinical and Health Psychology*, 2(3), 503-508. Retrieved from <https://www.redalyc.org/articulo.oa?id=33720308>
- Mullane, J. C., Corkum, P. V., Klein, R. M., & McLaughlin, E. (2009). Interference control in children with and without ADHD: a systematic review of Flanker and Simon task performance. *Child Neuropsychology*, 15(4), 321-342. doi: [10.1080/09297040802348028](https://doi.org/10.1080/09297040802348028)
- Oberauer, K. (2001). Removing irrelevant information from working memory: A cognitive aging study with the modified Sternberg task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(4), 948-957. doi: [10.1037/0278-7393.27.4.948](https://doi.org/10.1037/0278-7393.27.4.948)

- Pallant, J. (2007). *SPSS Survival Manual. A Step by Step Guide to Data Analysis Using SPSS for Windows* (3rd ed.). Maidenhead: Open University Press.
- Richard's, M. M., Introzzi, I., Zamora, E., & Vernucci, S. (2017a). Analysis of internal and external validity criteria for a computerized visual search task: A pilot study. *Applied Neuropsychology: Child*, 6(2), 110-119. doi: [10.1080/21622965.2015.1083433](https://doi.org/10.1080/21622965.2015.1083433)
- Richard's, M. M., López -Ramón, M. F., Zamora, E., V., & Introzzi, I. M. (2022). Development of perceptual inhibition throughout the lifespan: Theoretical models and studies. In F. Alcántud-Marín, M. F. López-Ramón, E. Navarro-Pardo, V. Moreno-Campos & Y. Alonso-Esteban (Eds.), *Handbook of research on neurocognitive development of executive functions and implications for intervention*. doi: [10.4018/978-1-7998-9075-1](https://doi.org/10.4018/978-1-7998-9075-1)
- Richard's, M., Vernucci, S., Zamora, E., Canet-Juric, L., Introzzi, I., & Guardia, J. (2017b). Contribuciones empíricas para la validez de grupos contrastados de la Batería de Tareas de Autorregulación Cognitiva (TAC). *Interdisciplinaria*, 34(1), 173-192. Retrieved from <http://www.ciipme-conicet.gov.ar/ojs/index.php?journal=interdisciplinaria>
- Stelzer, F. (2014). *Inteligencia, funciones ejecutivas y temperamento. Sus relaciones al inicio de la escolaridad básica*. (Tesis doctoral). Universidad Nacional de Mar del Plata. Retrieved from <http://m.rpsico.mdp.edu.ar/handle/123456789/2>
- Stevens, C., & Bavelier, D. (2012). The role of selective attention on academic foundations: A cognitive neuroscience perspective. *Developmental Cognitive Neuroscience*, 2, 30-48. doi: [10.1016/j.dcn.2011.11.001](https://doi.org/10.1016/j.dcn.2011.11.001)
- Stolte, M., Kroesbergen, E. H., & Van Luit, J. E. (2019). Inhibition, friend or foe? Cognitive inhibition as a moderator between mathematical ability and mathematical creativity in primary school students. *Personality and Individual Differences*, 142, 196-201. doi: [10.1016/j.paid.2018.08.024](https://doi.org/10.1016/j.paid.2018.08.024)
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology*, 40(2), 201-237. doi: [10.1080/02724988843000104](https://doi.org/10.1080/02724988843000104)
- Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12(1), 97-136. doi: [10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5)
- Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, 16(3), 459-478. doi: [10.1037/0096-1523.16.3.459](https://doi.org/10.1037/0096-1523.16.3.459)
- Vadaga, K. K., Blair, M., & Li, K. Z. H. (2016). Are age-related differences uniform across different inhibitory functions? *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 71(4), 641-649. doi: [10.1093/geronb/gbv002](https://doi.org/10.1093/geronb/gbv002)
- Zamora, E. V., Richard's, M. M., Canet-Juric, L., Aydmune, Y., & Introzzi, I. (2020). Perceptual, cognitive and response inhibition in emotional contexts in children. *Psychology & Neuroscience*, 13(3), 257-272. doi: [10.1037/pne0000202](https://doi.org/10.1037/pne0000202)
- Watters, J. K., & Biernacki, P. (1989). Targeted sampling: Options for the study of hidden populations. *Social Problems*, 36(4), 416-430. doi: [10.2307/800824](https://doi.org/10.2307/800824)
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the modified feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 419-433. doi: [10.1037/0096-1523.15.3.419](https://doi.org/10.1037/0096-1523.15.3.419)

Escala de Cansancio Emocional: Análisis psicométrico en estudiantes de posgrado en Puerto Rico

Emotional Exhaustion Scale: Psychometric Analysis in Postgraduate Students in Puerto Rico

Juan Aníbal González-Rivera *¹, Yazmín Álvarez-Alatorre¹,
Ernesto Rosario-Hernández², Vicmarie Sepúlveda-López³, Natasha Torres-Rivera³,
Taimara Ortiz-Santiago³, Miriam Tirado de Alba³, Chardlyn M. González-Malavé³

Introducción
Método
Resultados
Discusión
Referencias

1- Ponce Health Sciences University, San Juan University Center, Puerto Rico.

2- Ponce Health Sciences University, Ponce, Puerto Rico.

3- Universidad Carlos Albizu, San Juan, Puerto Rico.

Recibido: 21/03/2022 Revisado: 14/04/2022 Aceptado: 13/05/2022

Resumen

Esta investigación analiza las propiedades psicométricas de la Escala de Cansancio Emocional y confirma su estructura unidimensional. Participaron en la investigación 442 estudiantes de maestría y doctorado de 21 a 55 años ($M = 29.52$; $DE = 5.63$). Se realizaron análisis factoriales exploratorios y confirmatorios, análisis de invarianza y análisis de consistencia interna. Los análisis factoriales confirmaron la estructura unidimensional del instrumento (índices de ajuste apropiados) y los coeficientes alfa de Cronbach y omega de McDonald fueron adecuados ($\alpha = .883$; $\omega = .883$). Se concluye que la Escala de Cansancio Emocional cuenta con las propiedades psicométricas adecuadas para ser utilizada en estudiantes de posgrado de Puerto Rico y es una herramienta práctica y útil en el quehacer clínico y académico de las universidades. Se discuten las implicancias prácticas y limitaciones de los hallazgos.

Palabras clave: burnout, cansancio emocional, propiedades psicométricas, estudiantes universitarios, validez

Abstract

This research analyzes the psychometric properties of the Emotional Exhaustion Scale and confirms the unifactorial structure of the scale. 442 master's and doctoral students from 21 to 55 years old participated in the research ($M = 29.52$; $DE = 5.63$). Exploratory and confirmatory factor analysis, invariance analysis and internal consistency analysis were performed. Factor analyses confirmed the unidimensional structure of the scale (appropriate fit indexes), and Cronbach's alpha and McDonald's omega coefficients were adequate ($\alpha = .883$; $\omega = .883$). It was concluded that the Emotional Exhaustion Scale has good psychometric properties to be used in postgraduate students in Puerto Rico and is a practical and helpful tool for clinical and academic work in universities. The practical implications and limitations of the findings are discussed.

Keywords: burnout, emotional exhaustion, psychometric properties, college students, validity

*Correspondencia a: Dr. Juan Aníbal González-Rivera, 500 West Main Suite 215, Bayamón, Puerto Rico. 00961. E-mail: jagonzalez@psm.edu

Nota de autores: Juan Aníbal González-Rivera <https://orcid.org/0000-0003-0622-8308>, Yazmín Álvarez-Alatorre <https://orcid.org/0000-0001-8989-1679>, Ernesto Rosario-Hernández <https://orcid.org/0000-0002-2523-6162>, Vicmarie Sepúlveda-López <https://orcid.org/0000-0003-2426-4708>, Natasha Torres-Rivera <https://orcid.org/0000-0003-3975-4208>, Taimara Ortiz-Santiago <https://orcid.org/0000-0002-3944-4835>, Miriam Tirado de Alba <https://orcid.org/0000-0003-3898-3041>, Chardlyn M. González-Malavé <https://orcid.org/0000-0002-4286-6043>

Cómo citar este artículo: González-Rivera, J. A., Álvarez-Alatorre, Y., Rosario-Hernández, E., Sepúlveda-López, V., Torres-Rivera, N., Ortiz-Santiago, T., Tirado de Alba, M., & González-Malavé, C. M. (2022). Escala de Cansancio Emocional: Análisis psicométrico en estudiantes de posgrado en Puerto Rico. *Revista Evaluar*, 22(2), 47-63. Recuperado de <https://revistas.unc.edu.ar/index.php/revaluar>

Participaron en la edición de este artículo: Rita Hoyos, Stefano Macri, Juan Cruz Balverdi, Mercedes Leaden, Florencia Ruiz, Mónica Serppe, Alicia Molinari, Ricardo Hernández, Benjamín Casanova

Introducción

El estudio del agotamiento o cansancio emocional, conocido como *burnout*, y los contextos en los que se presenta ha sido tema de interés en ramas desde lo fisiológico hasta lo psicosocial. El término *burnout* surgió como una expresión usada en novelas y artículos publicados en la segunda mitad del siglo XX; sin embargo, ninguno tuvo la intención de extrapolarlo a definición conceptual (Fernandes-Fontes, 2020). Las raíces de este concepto se asocian a la alta velocidad con la que se han desarrollado cambios socioeconómicos y culturales que han transformado la sociedad industrial en una sociedad de economía de servicios. Sin embargo, el *burnout* no tiene una definición única, sino que puede variar según el país o el enfoque de las investigaciones, ya que este término es usado ampliamente en investigaciones internacionales (Caballero-Domínguez, Hederich-Martínez, & Palacio-Sañudo, 2010; Caballero-Domínguez, González-Gutiérrez, & Palacio-Sañudo, 2015; Dominguez-Lara, 2018; López-Osorio, Cano, & Salazar-Ospina, 2020; Manassero-Mas et al., 2003). Por lo tanto, algunos autores optan por usar *burnout* como término universal, más allá de que sea una palabra en inglés. Asimismo, cabe resaltar que el término *burnout* puede encontrarse en distintas publicaciones en español conservando el término original en inglés (Caballero et al., 2010, 2015; Dominguez-Lara, 2018; Martínez-Pérez, 2010; Rosales-Ricardo & Rosales-Paneque, 2013) o traducido en español a términos como: *síndrome de quemarse* (Rodríguez-Colón, 2016), *agotamiento emocional* (Dominguez-Lara, 2014), *desgaste emocional* (Fontana, 2011), *cansancio emocional* (González-Ramírez & Landero-Hernández, 2007; Ramos-Campos, Manga-Rodríguez, & Morán-Astorga, 2005).

La concepción de *burnout* como erosión de

un estado psicológico positivo ha ido en aumento, sin embargo, la visión de este término como estado emocional es diversa. En algunos países de Europa, como Suecia y Holanda (donde el entorno médico se caracteriza por tener sistemas de seguridad social ampliamente estructurados) el diagnóstico médico de *burnout* requiere formación previa para su evaluación y tratamiento. En otros países, el modelo médico para definir este concepto como estado emocional no se utiliza, más bien se tiene un enfoque social que lo aleja del estigma de un diagnóstico psiquiátrico (Schaufeli, Leiter, & Maslach, 2009). Es importante puntualizar que el *burnout* como un síndrome no se encuentra registrado en el Manual Diagnóstico y Estadístico de los Trastornos Mentales, 5° Edición (DSM-5; American Psychiatric Association, 2013); sin embargo, su relevancia mundial y los efectos que tiene en la salud física y mental han hecho que sea reconocido a nivel mundial e incluido en la Clasificación Internacional de Enfermedades, 10° Edición (ICD-10) con el código Z73.0 que lo define como “estado de agotamiento vital” (Bianchi, Schonfeld, & Laurent, 2015).

Múltiples autores dan crédito al psiquiatra Herbert Freudenberg (1974) por acuñar el concepto en sus estudios en los años 70 (Fernandes-Fontes, 2020; Rodríguez-Colón, 2016; Rosales-Ricardo & Rosales-Paneque, 2013; Schaufeli et al., 2009). Freudenberg (1974) describió los síntomas físicos que producen ciertos entornos laborales y que causan agotamiento emocional, pérdida de motivación y compromiso. Este autor recibió, incluso, reconocimientos de la Asociación Americana de Psicología (APA) por sus estudios sobre *burnout*, los que tuvieron más relevancia por sus relatos autobiográficos, que aumentaron su credibilidad. A raíz de sus publicaciones, se continuó investigando el desarrollo y los efectos de este estado de agotamiento emocional de relevancia cada vez mayor.

Una de las pioneras en investigar el *burnout* fue Christina Maslach, quien en 1976 publicó el artículo “Burned-Out”, dejando al descubierto sus efectos dañinos sobre distintas profesiones. En ese estudio, la autora concluyó que los profesionales que trabajan brindando servicio directo a otras personas (médicos, abogados, personal de cárceles, trabajadores sociales, psicólogos clínicos, psiquiatras, entre otros) tienden a lidiar con altos niveles de estrés y su manera de afrontarlo es a través del distanciamiento de sus propios sentimientos, reprimiendo la necesidad de expresar sus angustias. Se destaca dentro de sus hallazgos el hecho de que el *burnout* conduce al deterioro físico; el profesional que lo padece se reporta enfermo constantemente, afectado principalmente por insomnio, migraña, úlceras y otras afecciones más severas. La autora incluso encontró correlaciones con el alcoholismo, las enfermedades mentales, los problemas maritales y el suicidio. Actualmente existe un consenso general que define al *burnout* como “una respuesta al estrés crónico en el trabajo (a largo plazo y acumulativo), con consecuencias negativas a nivel individual y organizacional...” (Martínez-Pérez, 2010, p. 44).

En 1981, se desarrolló y validó el Inventario de Burnout de Maslach (MBI); con una estructura de tres dimensiones: *agotamiento emocional*, *despersonalización* y *realización personal* (Maslach & Jackson, 1981). Con el fin de expandir estas medidas al contexto de proveedores de servicios de otros grupos ocupacionales, se adaptó la escala original y se desarrolló el Inventario de Burnout de Maslach - Encuesta General (MBI-GS); con tres dimensiones: *agotamiento*, *eficacia personal* y *cinismo*. Se destaca el reemplazo de la dimensión de *despersonalización* por la de *cinismo* (indiferencia o actitud distante hacia el trabajo), y que se busca la adaptación de la dimensión *realización personal* al personal que brinda servicio ocasional o indirecto a otros ya que esta dimen-

sión está exclusivamente asociada a los servicios humanos (Maslach, Jackson, & Leiter, 1997).

Sin embargo, Bianchi et al. (2015) advierten sobre las razones por las que no debe considerarse al *burnout* como una categoría nosológica y señalan fallas metodológicas en la definición del constructo y arbitrariedades en la elaboración del MBI. Estos autores destacan cuestionamientos que se han hecho a la estructura tridimensional del constructo en la última década, por ejemplo, que el centro de las dimensiones del inventario: el *agotamiento emocional*, ha sido consistentemente relacionado con síntomas de depresión. Es por esto que los autores puntualizan sobre diferenciar y distanciar al *burnout* de una categoría diagnóstica o enfermedad particular, ya que este término se solapa con el de depresión; y afirman que la ciencia en la actualidad sugiere que el *burnout* es una forma de depresión, más que una patología diferenciada (Bianchi et al., 2015, p. 158).

Aunque se haya confirmado una correlación entre *burnout* y depresión mediante la dimensión del agotamiento emocional, no ha sido así con las otras dimensiones (Martínez-Pérez, 2010). La importancia del estudio del *burnout* radica en sus manifestaciones físicas, conductuales, psicológicas, sociales, actitudinales y cognitivas y con los costos asociados con tratar este síndrome y su posible propagación entre los empleados (Rodríguez-Colón, 2016). Existen autores que consideran que los síntomas de depresión pueden ser consecutivos al *burnout* (Manassero-Mas et al., 2003; Quintero-Idárraga & Hernández-Calle, 2021; Szigeti, Balázs, Bikfalvi, & Urbán, 2017); sin embargo, algunos síntomas ocurren en ambos estados y es necesario distinguirlos. La extenuación extrema (agotamiento emocional, cansancio físico, baja energía, que puede manifestarse en dolores y problemas gastrointestinales), el ánimo decaído y el bajo rendimiento (en el trabajo, en la rutina diaria o en el cuidado de familiares), son

síntomas que se presentan en ambas condiciones.

Por su parte, el *burnout* se distingue porque el malestar está fuertemente relacionado con las actividades laborales, profesionales o académicas y provoca, de esta manera, una alienación (distanciamiento, cinismo, desconexión e indiferencia hacia el trabajo y los colegas). Es por esto que inventarios como el MBI, originalmente creado con propósitos investigativos y no clínicos, deben usarse con cautela, acompañándolos con un adecuado cernimiento médico y evaluación psicológica. Unas vacaciones pueden servir de recomendación para aquellos con *burnout* y eso traducirse en bienestar y reducción de estrés, pero sería negligente recomendarlas a alguien con depresión, para quien la visión personal y del mundo se componen de pensamientos y sentimientos negativos, en casi todas las áreas de la vida, no únicamente en lo laboral u ocupacional (InformedHealth.org, 2020).

Para este estudio usaremos el término *agotamiento emocional*, constructo medido por [Dominguez-Lara \(2014\)](#) usando la Escala de Cansancio Emocional (ECE), de la que obtuvo datos normativos válidos, confiables y adecuados en una muestra de estudiantes de psicología de tres universidades de Lima, Perú. Este autor destaca la importancia de reconocer la relación inversa entre el agotamiento emocional y la autoeficacia académica, autoestima, compromiso e inteligencia emocional, así como las repercusiones de las exigencias académicas en la salud de los estudiantes universitarios (sentimientos de soledad, impotencia, apatía y afectación de la percepción de su futuro profesional, que provocan falta de motivación que pudiera culminar en deserción).

En Puerto Rico, el *burnout* o *agotamiento emocional* ha sido un tema poco investigado. Los efectos del *burnout* se han expuesto en estudios relacionados con profesiones como la Consejería Profesional, específicamente en ambientes acadé-

micos. Este entorno puede resultar estresante en niveles que pueden afectar el desempeño laboral, la productividad y el presentismo. Además, es poco probable que los consejeros se percaten del estado de *quemazón laboral*, ya que el exceso de trabajo les impide reflexionar sobre cantidad versus calidad de trabajo; su deterioro físico y agotamiento emocional les impide ser conscientes de sus actividades y autocuidado ([Rodríguez-Colón, 2016](#)). En profesiones relacionadas con la provisión de servicios de atención clínica, en medicina y enfermería, más del 90% de los encuestados vía correo electrónico y teléfono (al Colegio de Médicos Cirujanos de PR y la Asociación de Enfermería de Puerto Rico) reportaron niveles de agotamiento emocional de moderados a severos ([Cortina-Rodríguez & Afanador, 2020](#)). En otros países, el *burnout* ha provocado la creación de protocolos de prevención para atender el problema, por ejemplo, en Colombia ([Ministerio del Trabajo, 2015](#)), donde además algunas investigaciones han sido específicas para explorar este estado emocional en personal de salud en tiempos de la pandemia por COVID-19 ([López-Osorio et al., 2020](#)).

Escala de Cansancio Emocional (ECE)

La ECE fue desarrollada por [Ramos-Campos et al. \(2005\)](#); tiene una estructura unidimensional y está compuesta por 10 ítems que miden el *agotamiento emocional* en estudiantes universitarios. La escala obtuvo un índice de consistencia interna alfa de Cronbach de .83. Del análisis de resultados de la validación de la ECE surge la estrecha relación de variables de personalidad (específicamente inestabilidad emocional, bajo autoconcepto, baja autoestima) con el cansancio emocional en estudiantes universitarios, siendo el neuroticismo el mayor predictor de este construc-

to, que resultó ser mayor en mujeres que en varones. Las propiedades psicométricas de este instrumento han sido validadas en distintos países de Latinoamérica como México (Bonilla et al., 2009; González-Ramírez & Landero-Hernández, 2007), Argentina (Fontana, 2011), Perú (Dominguez-Lara, 2014) y Chile (Martínez-Líbano, Yeomans, & Oyanedel, 2022). Cabe señalar que, de estos estudios, los únicos que realizaron análisis factoriales confirmatorios fueron González-Ramírez y Landero-Hernández (2007), Dominguez-Lara (2014) y Martínez-Líbano et al. (2022), quienes sustentan un modelo unidimensional de la ECE.

Por otra parte, se ha medido la varianza por el error transitorio (error sistemático que refleja fluctuaciones en los estados psicológicos de los participantes como la atención y el estado de ánimo) y se ha probado su invarianza longitudinal y su confiabilidad temporal, lo que hace útil a la escala para proyectos preprueba y posprueba (Dominguez-Lara, 2018). Al presente, no existen instrumentos para medir el agotamiento emocional en estudiantes universitarios en Puerto Rico, por lo que el propósito de esta investigación será obtener evidencias de validez a partir de las puntuaciones generadas en la Escala de Cansancio Emocional.

Método

Diseño de investigación y asuntos éticos

Según las guías clasificatorias de Ato, López-García y Benavente (2013), esta investigación se encuadra dentro del modelo no experimental de tipo instrumental. El proyecto fue aprobado por el Comité para la Ética en la Investigación de la Universidad Carlos Albizu en San Juan, Puerto Rico. La recopilación de datos se realizó de forma electrónica a través de la plataforma PsychData y se reclutaron participantes mediante propaganda

en las redes sociales que los dirigía a un consentimiento informado y a la encuesta en línea. El consentimiento informado explicaba detalladamente el propósito del estudio, los criterios de inclusión, la naturaleza voluntaria del estudio, los posibles riesgos y beneficios, así como el derecho de los voluntarios a retirarse del estudio en cualquier momento.

Participantes

Participaron en la investigación 442 estudiantes de posgrado entre 21 y 55 años, con una edad media de 29.52 y una desviación estándar de 5.63. La estrategia de muestreo fue por conveniencia. Los criterios de inclusión fueron: (1) ser mayor de 21 años de edad, (2) estar matriculado en un programa de maestría o doctorado al momento de la investigación y (3) ser residente de Puerto Rico. La Tabla 1 muestra que la mayoría de los estudiantes eran mujeres (75.1%), un 65.6% cursaba estudios de maestría, un 85.1% de la muestra se encontraba entre el primer y tercer año de la maestría, un 83.5% tenía un promedio académico entre 3.50 y 4.00 en una escala de 0 a 4, y un 55.9% indicó que estudiaba diariamente entre una y tres horas.

Instrumentos

Escala de Cansancio Emocional (ECE). La escala fue desarrollada y validada originalmente por Ramos-Campos et al. (2005) y contiene 10 ítems que se puntúan de 1 a 5 (desde *Raras veces* hasta *Siempre*) considerando los últimos 12 meses de vida estudiantil. Ramos-Campos et al. (2005) reportan que los ítems fueron inspirados en la escala de cansancio emocional del Maslach Burnout Inventory (Maslach et al., 1997) y en el concepto

Tabla 1

Datos sociodemográficos de los participantes del estudio (n = 442).

Variable	f	%	Variable	f	%
<i>Género</i>			<i>Ciclo Académico que Cursa</i>		
Femenino	332	75.1	Primero	131	29.6
Masculino	108	24.4	Segundo	128	29.0
Otro	2	0.5	Tercero	73	16.5
<i>Estudia</i>			Cuarto	35	7.9
Maestría	290	65.6	Quinto o más	75	17.0
Doctorado	148	33.5			
Post Doctorado	4	0.9			
<i>Promedio Académico</i>			<i>Horas Dedicadas Diariamente a Estudiar</i>		
3.50 - 4.00	369	83.5	1-3	247	55.9
3.00 - 3.49	65	14.7	3-5	125	28.3
2.50 - 2.99	6	1.4	5 o más	70	15.8
2.00 - 2.49	2	0.5			

Nota. f = frecuencia; % = por ciento.

de *burnout* de Freudenberger (Morán-Astorga, 2003). Fontana (2011) indica que, según los resultados de un análisis de factores exploratorio, estos 10 ítems son una medida unidimensional de agotamiento emocional y obtuvieron un coeficiente de confiabilidad de .87 mediante el índice alfa de Cronbach.

Análisis de datos

Primeramente, se realizaron análisis descriptivos y del conjunto de ítems de la Escala de Cansancio Emocional con los cuales se obtuvo la media, desviación estándar, asimetría, curtosis e índice de discriminación (r_{bis}) en una muestra de 442 participantes. En lugar de analizar toda la muestra en una sola acción de análisis, se aplicó una estrategia de validación cruzada para evaluar la estabilidad de los parámetros de validez en la muestra (Browne & Cudeck, 1989; Cudeck

& Browne, 1983; Whittaker & Stapleton, 2006). De esta forma, se dividió la muestra de forma aleatoria en Muestra 1 (n_1) y Muestra 2 (n_2), la primera sirvió como muestra de calibración y la segunda como muestra de validación. De esta forma, se realizaron análisis de factores exploratorios (AFE) con la n_1 mientras que, con la n_2 , se realizó un análisis factorial confirmatorio (AFC). Por otro lado, para examinar la invarianza de medición, confiabilidad y estadística descriptiva de las puntuaciones de la Escala de Cansancio Emocional, se volvieron a juntar todos los datos en una sola muestra.

Se utilizó el programa Statistical Package for Social Science conocido como SPSS versión 28 para realizar los análisis estadísticos descriptivos, de ítems y AFE. Para el AFE, se utilizó como método de extracción ejes principales de factorización con oblimin directa como técnica de rotación oblicua. Además, se tomó como criterio de selección el que los ítems obtuvieran una carga

factorial igual o mayor a .30 en el factor al cual supuestamente pertenecían y menos de .30 en los demás factores (Kline, 1994). Por otro lado, se utilizó el modelado de ecuaciones estructurales para realizar el análisis factorial confirmatorio (AFC) con el sistema R (versión R3.6.3) y el paquete estadístico Lavaan para modelos de ecuaciones estructurales (Rosseel, 2012). En cuanto al AFC, se usó el método de estimación de mínimos cuadrados ponderados robusto y ajuste de varianza (WLSMV), el cual trata de forma robusta datos potencialmente no normales y los elementos se tratan como ordinales (Li, 2016a, 2016b). Para evaluar los resultados del AFC, se utilizaron varios índices de ajuste, Kline (2016) recomienda el uso de al menos cuatro índices de ajuste, aunque se pueden reportar más. Uno de los índices que se informa es chi cuadrado (χ^2). Este es un índice fundamental de ajuste absoluto y es básicamente el mismo que se utiliza cuando se quiere examinar la asociación entre variables nominales; sin embargo, la diferencia crucial cuando se usa como índice de ajuste en el modelo de ecuaciones estructurales es que el investigador no busca diferencias entre las matrices para respaldar que el modelo probado es representativo de los datos (Hair et al., 2018).

Dado que χ^2 es sensible al tamaño de la muestra y por tanto la probabilidad de rechazar el modelo hipotético aumenta cuando aumenta el tamaño de la muestra, se recomienda tener en cuenta otros índices (Marsh, Balla, & Hau, 1996). También se utilizó el *root mean square error of approximation* (RMSEA; Byrne, 2016; Hu & Bentler, 1999). Valores menores a .05 indican un buen ajuste del modelo, valores hasta .08 se consideran como razonables y aquellos que van desde .08 a .10 se consideran mediocres (Browne & Cudeck, 1989; MacCallum, Browne, & Sugawara, 1996). Además, se utilizó *standardized square root mean residual* (SRMR; Hu &

Bentler, 1999), que examina la diferencia promedio entre las varianzas y covarianzas pronosticadas y observadas, con base en el error estándar residual. Cuanto menor sea el SRMR, mejor será el ajuste del modelo y para considerar un modelo aceptable debe ser igual o inferior a .05, aunque hay autores que indican que valores $\leq .08$, se consideran aceptables (Hu & Bentler, 1995). Por otro lado, se utilizó el *Bentler comparative fit index* (CFI) como índice de ajuste aumentado para comparar el modelo teórico con el modelo nulo, el cual asume que las variables latentes del modelo no se correlacionan entre sí y valores mayores a .90 se consideran aceptables (Hair et al., 2018). Otro índice de ajuste incrementado utilizado fue el índice de Tucker-Lewis (TLI) que refleja la proporción en que el modelo teórico mejora el ajuste en relación con el modelo nulo (Littlewood-Zimmerman & Bernal-García, 2011; Tucker & Lewis, 1973). Valores superiores a .90 se consideran aceptables.

Para establecer los niveles de invarianza de medición de la Escala de Cansancio Emocional se siguieron las recomendaciones de Chen (2007) y Cheung y Rensvold (2002). Para establecer la invarianza métrica o débil, las diferencias del índice de ajuste para CFI y SRMR debían ser $\Delta CFI \leq -.01$ y $\Delta SRMR \geq .030$, respectivamente. Mientras que, para establecer la invarianza escalar o fuerte, las diferencias del índice de ajuste para CFI y SRMR debían ser $\Delta CFI \leq -.01$ y $\Delta SRMR \geq .010$, respectivamente. Chen (2007) encontró en un estudio de Monte Carlo que estos índices eran igualmente sensibles a todos los tipos de invarianzas. Por otro lado, el índice de χ^2 se informó, pero dado que está muy influenciado por el tamaño de la muestra (Rigdon, 1995), el mismo no se consideró como índice de ajuste para la prueba de invarianza.

Finalmente, para estimar la confiabilidad de la Escala de Cansancio Emocional se utilizaron las técnicas de alfa de Cronbach y omega de

McDonald con sus respectivos intervalos de confianza. Además, se estimó la estadística descriptiva de la escala, el error estándar de medición y el intervalo de confianza de 95%.

Resultados

En la Tabla 2 se puede apreciar los estadísticos descriptivos (e.g., media, desviación estándar, entre otros) y el índice de discriminación (r_{bis}) de los 10 ítems de la Escala de Cansancio Emocional. El ítem 8 obtuvo la media más alta y el ítem 9 obtuvo la desviación estándar más alta. Mientras que el ítem 7 obtuvo el r_{bis} más alto.

Tabla 2

Estadística descriptiva e índice de discriminación (r_{bis}) de los ítems de la Escala de Cansancio Emocional.

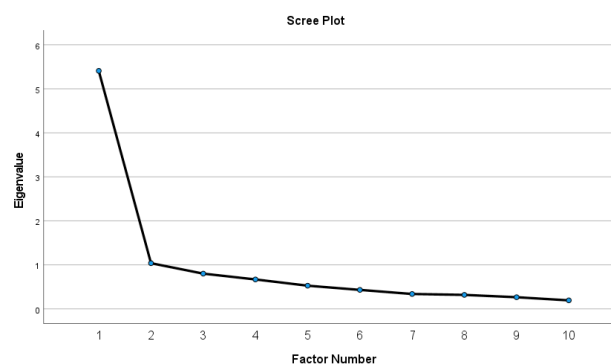
Ítem	Media	DE	Asimetría	Curtosis	r_{bis}
ítem1	3.32	1.12	-0.216	-0.490	.549
ítem2	2.48	1.24	0.384	-0.923	.536
ítem3	2.46	1.18	0.222	-0.944	.635
ítem4	2.93	1.19	-0.196	-0.901	.637
ítem5	2.70	1.22	0.066	-1.000	.718
ítem6	3.11	1.16	-0.326	-0.691	.744
ítem7	3.21	1.14	-0.318	-0.505	.773
ítem8	3.49	1.14	-0.470	-0.429	.711
ítem9	3.20	1.27	-0.160	-1.017	.700
ítem10	3.31	1.22	-0.296	-0.767	.706

Nota. n = 442.

Se realizaron AFE con la Muestra 1 usando el método de extracción de factorización de ejes principales y la técnica de rotación oblicua directa. El primer AFE obtuvo un valor adecuado en la prueba Kaiser-Meyer-Olkin ($KMO = .890$) y la prueba de Bartlett fue significativa, $\chi^2_{(45)} = 1,243.318$, $p < .05$, en cinco iteraciones. Este pri-

mer AFE produjo una solución de dos factores explicando un 49.99% de la varianza (Tabla 3). Sin embargo, el punto de inflexión del gráfico de sedimentación sugiere un solo factor (Figura 1 del primer AFE). Los ítems 2, 3, 4, 5, 6, 7 y 8, cargaron en el Factor 1; mientras que los ítems 1 y 9 cargaron en el Factor 2, y finalmente el ítem 10 obtuvo cargas factoriales cruzadas mayores a .30 en ambos factores. Por tanto, se realizó un segundo AFE sin incluir los ítems 1, 9 y 10. Este segundo AFE obtuvo un $KMO = .863$ y la prueba de Bartlett fue significativa, $\chi^2_{(21)} = 780.386$, $p < .05$, en cinco iteraciones. El punto de inflexión del gráfico de sedimentación de este segundo AFE sugiere también un factor (Figura 1 del segundo AFE). En este segundo AFE, todas las cargas factoriales fueron mayores a .50 (véase la Tabla 3).

1er AFE



2do AFE

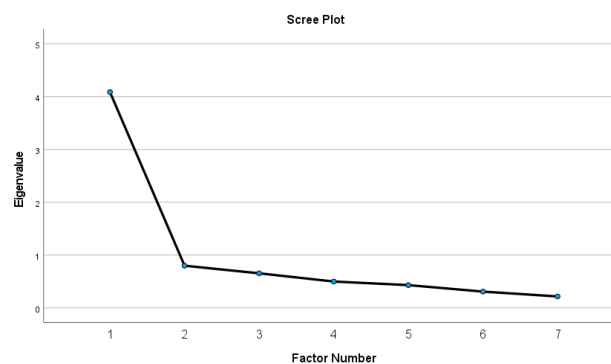


Figura 1. Gráficos de sedimentación.

Tabla 3
Análisis factorial exploratorio (AFE).

# Ítem	Aseveración	1 ^{er} AFE			2 ^{do} AFE	
		Factor		h ²	Factor	
		1	2		1	h ²
1.	Los exámenes me producen una tensión excesiva.		-.61	.483	X	X
2.	Creo que me esfuerzo mucho para lo poco que consigo.	.60		.344	.58	.331
3.	Me siento bajo de ánimo, como triste, sin motivo aparente.	.78		.483	.68	.462
4.	Hay días que no duermo bien a causa del estudio.	.62		.438	.65	.427
5.	Tengo dolor de cabeza y otras molestias que afectan a mi rendimiento académico.	.75		.530	.73	.533
6.	Hay días que noto más la fatiga y me falta energía para concentrarme.	.71		.615	.78	.615
7.	Me siento emocionalmente agotado por mis estudios.	.78		.702	.84	.708
8.	Me siento cansado al final de la jornada de estudio.	.55		.609	.74	.551
9.	Estudiar pensando en los exámenes me produce estrés.		-.987	.934	X	X
10.	Me falta tiempo y me siento desbordado por los estudios.	.507	-.321	.559	X	X
	Valor Eigen	5.00	.68		3.63	
	% Varianza Explicada	49.99	6.97		51.81	
	% Varianza Acumulada	49.99	56.96		51.81	

Nota. $n_1 = 229$, X = ítem no incluido en el análisis.

Se realizó un análisis de factores confirmatorio usando el modelo de ecuaciones estructurales, incluyendo los siete ítems que cumplieron con el criterio anterior, pero esta vez se usó la Muestra 2. Dado que la Escala de Cansancio Emocional es de tipo ordinal, se usó el estimador de mínimos cuadrados ponderados robusto y ajuste de varianza (WLSMV; Li, 2016a, 2016b). De esta forma, se procedió a examinar la estructura unidimensional de primer orden de la Escala de Cansancio Emocional con los siete ítems que cumplieron con los criterios del AFE. Así, los índices de ajuste obtenidos se pueden considerar como buenos con excepción del RMSEA, que está por encima de lo recomendado (Tabla 4). Además, todas las cargas factoriales fueron mayores a .70 (véase la

Figura 2), excepto el ítem 2, que fue mayor a .50, tal y como lo recomiendan otros autores, como por ejemplo, Kline (2016). De esta forma, se procedió a estimar la varianza media extraída de este conjunto de siete ítems que fue igual a .62, un valor que está por encima del valor umbral de .50 propuesto por otros autores (Fornell & Bookstein, 1982; Fornell & Larcker, 1981), lo cual provee evidencia que apoya la validez convergente de la Escala de Cansancio Emocional.

Luego se procedió a examinar la invarianza de medición de la Escala de Cansancio Emocional entre participantes de varios grupos, específicamente género (masculino/femenino), edad (21-30 / ≥ 31), grado de estudio (maestría/doctorado), ciclo académico de posgrado (1ro - 2do año / \geq

Tabla 4
Índices de ajustes obtenidos de la Escala de Cansancio Emocional.

χ^2	<i>gl</i>	SRMR	RMSEA (IC)	CFI	TLI
42.657*	14	.040	.098 (.065, .133)	.989	.984

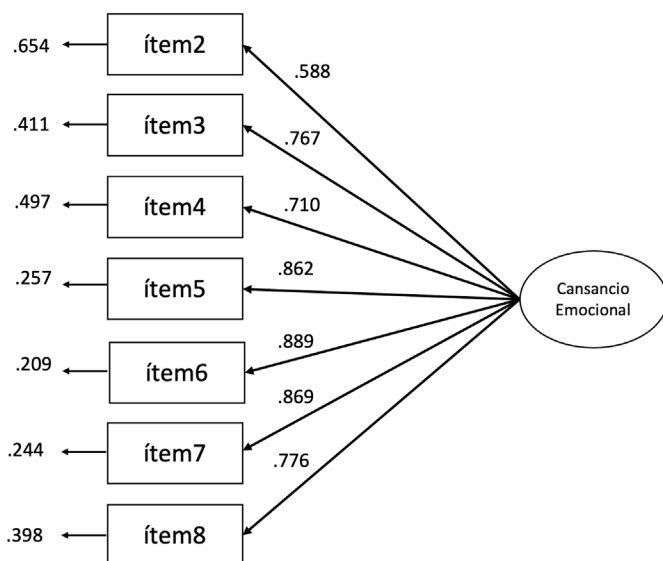


Figura 2. Modelo de medición probado de la Escala de Cansancio Emocional con sus respectivas cargas factoriales.

3er año), y horas de estudio por semana (1-3 horas / ≥ 4 horas). Se tomaron en consideración los cambios en los índices de SRMR y CFI para determinar la invarianza de medición de la Escala de Cansancio Emocional. Es importante indicar que, como el índice RMSEA puede verse afectado por un modelo con un número pequeño de grados de libertad, se decidió no utilizarlo para examinar la invarianza y utilizar el SRMR y CFI (Kenny, Kaniskan, & McCoach, 2015; Shi et al., 2021). El modelo de un factor de la Escala de Cansancio Emocional se integró en el modelo de invarianza configuracional, con el mismo pa-

trón de cargas factoriales fijas y libres, pero no se impusieron restricciones de igualdad en ningún parámetro entre los grupos. El modelo de invarianza configuracional se ajustó bien a los datos (ver Tabla 5). Este modelo configuracional se usó luego para compararlo con el modelo de invarianza de medición más restrictivo (es decir, invarianza de medición métrica o débil) que examinamos a continuación. El primer modelo más restrictivo, el modelo de invarianza débil, se ajustó bien a los datos para todos los grupos (ver Tabla 5). Los cambios de SRMR y CFI, cuando se compara el modelo de invarianza débil con el modelo de invarianza configuracional, estuvieron dentro de los valores aceptables para todas las comparaciones de grupo. El siguiente modelo restrictivo, el modelo de invarianza escalar o fuerte, también se ajustó bien a los datos (ver Tabla 5). El segundo modelo más restrictivo, que restringió las cargas factoriales y la intercepción de ítems resultó en la demostración de una invarianza fuerte. Esto indica que tanto las cargas factoriales como la intercepción de los ítems son invariantes entre los grupos comparados.

Finalmente, se estimó la confiabilidad de la Escala de Cansancio Emocional usando la técnica de alfa de Cronbach y omega de McDonald con sus respectivos intervalos de confianza. Los coeficientes de confiabilidad obtenidos fueron iguales a .883 para ambos índices. También se estimó la estadística descriptiva de las puntuaciones de la escala con estos siete ítems, tales como la media, desviación estándar, error estándar de medición y el intervalo de confianza de 95% (Tabla 6).

Tabla 5

Invarianza de medición de la Escala de Cansancio Emocional por género, edad, grado de estudio, ciclo académico y horas de estudio por semana.

Modelo	$\chi^2_{(gl)}$	SRMR	CFI	Modelo de Referencia	$\Delta\chi^2$	Δ SRMR	Δ CFI
<i>Análisis Multi-Grupo por Género (Masculino/Femenino)</i>							
1. Configural	137.240* ₍₂₈₎	.054	.977	---	---	---	---
2. Métrica	110.161* ₍₃₄₎	.055	.984	1	-27.079	+0.001	+0.007
3. Escalar	143.840* ₍₅₄₎	.054	.981	2	+33.679	-0.001	-0.003
<i>Análisis Multi-Grupo por Edad (21-30/ ≥ 31)</i>							
1. Configural	131.081* ₍₂₈₎	.051	.979	---	---	---	---
2. Métrica	113.008* ₍₃₄₎	.055	.984	1	-18.073	+0.004	+0.005
3. Escalar	167.460* ₍₅₄₎	.052	.977	2	+54.452	-0.003	-0.007
<i>Análisis Multi-Grupo por grado de estudio (Maestría/Doctorado)</i>							
1. Configural	130.008* ₍₂₈₎	.052	.980	---	---	---	---
2. Métrica	127.386* ₍₃₄₎	.058	.982	1	-2.622	+0.006	+0.002
3. Escalar	163.329* ₍₅₄₎	.053	.979	2	+35.943	-0.005	-0.003
<i>Análisis Multi-Grupo por Ciclo Académico o Nivel de Posgrado (1^{ro} - 2^{do} año / ≥ 3^{er} año)</i>							
1. Configural	156.372* ₍₂₈₎	.055	.975	---	---	---	---
2. Métrica	129.263* ₍₃₄₎	.058	.981	1	-27.109	+0.003	+0.006
3. Escalar	187.905* ₍₅₄₎	.056	.974	2	+58.642	-0.002	-0.007
<i>Análisis Multi-Grupo por Horas de Estudio por Semana (1-3 / ≥ 4)</i>							
1. Configural	135.031* ₍₂₈₎	.051	.979	---	---	---	---
2. Métrica	113.085* ₍₃₄₎	.054	.984	1	-21.946	+0.003	+0.005
3. Escalar	137.825* ₍₅₄₎	.052	.983	2	+24.740	-0.002	-0.001

Nota. n = 442, gl = grados de libertad.

Tabla 6

Confiabilidad, estadística descriptiva e intervalo de confianza de 95% de la Escala de Cansancio Emocional.

Media	DE	Rango Posible	Confiabilidad		EEM	IC de 95%
			α (IC)	ω (IC)		
20.37	6.34	1 - 35	.883 (.865, .898)	.883 (.866, .899)	2.17	4

Nota. n = 442; EEM = error estándar de medición, IC = intervalo de confianza.

Discusión

El propósito de este estudio fue obtener evidencias de validez de la Escala de Cansancio Emocional (ECE) en estudiantes de posgrado en Puerto Rico. Cabe destacar que en este país no existen instrumentos para medir cansancio o agotamiento emocional en estudiantes universitarios. El cotejo de la estructura interna de la ECE mediante análisis factoriales tanto exploratorio como confirmatorio aporta evidencia sobre la unidimensionalidad que subyace a los ítems de la escala. Estos hallazgos coinciden con los reportados en España (Ramos-Campos et al., 2005), México (Bonilla et al., 2009; González-Ramírez & Landero-Hernández, 2007), Argentina (Fontana, 2011), Perú (Dominguez-Lara, 2014) y Chile (Martínez-Líbano et al., 2022); aunque los índices de ajuste en esta muestra resultaron ser más estables, casi igualados por el estudio de Martínez-Líbano et al. (2022). Por ejemplo, encontramos marcadas diferencias en el RMSEA de los estudios que utilizaron ecuaciones estructurales para sus análisis (Dominguez-Lara, 2014; González-Ramírez & Landero-Hernández, 2007; Martínez-Líbano et al., 2022). El RMSEA debe ser inferior a .08 (Byrne, 2016), aunque son admisibles valores iguales o menores a .10 (Sánchez & Sánchez, 1998). Dos de estos estudios sobrepasaron el umbral de aceptación recomendado: RMSEA = .117 (González-Ramírez & Landero-Hernández, 2007) y RMSEA = .104 (Dominguez-Lara, 2014). Tanto el presente estudio como el de Martínez-Líbano et al. (2022) obtuvieron valores considerados como mediocres por algunos autores (Browne & Cudeck, 1989; MacCallum et al., 1996), pero aceptables por otros (Sánchez & Sánchez, 1998). No obstante, y dado que la interpretación del índice RMSEA se puede ver afectada por modelos con pocos grados de libertad, enfatizamos en los índices de CFI y SRMR como

indicadores confiables ya que no se ven afectados por los grados de libertad (Kenny et al., 2015; Shi et al., 2021).

Por otra parte, en el índice comparativo de ajuste (CFI), que compara el modelo estimado con el modelo nulo, todos los estudios obtuvieron valores iguales o superiores a .90, que se reconocen como valores aceptables. No obstante, el presente estudio fue el único que superó el umbral de .95, que suele ser clasificado como un buen ajuste (Byrne, 2016). La integración de estos hallazgos con los encontrados en los estudios antes mencionados confirma que el modelo unidimensional de la ECE descalifica al modelo nulo. En el índice Tucker-Lewis (TLI), que compara el ajuste por grado de libertad del modelo propuesto y el nulo, los únicos estudios que presentaron valores aceptables fueron el estudio de Martínez-Líbano et al. (2022), que obtuvo un valor de .93 y el presente estudio (TLI = .984), considerado como bueno. Dominguez-Lara (2014) no reportó el TLI y el estudio de González-Ramírez y Landero-Hernández (2007) no alcanzó el mínimo recomendado (TLI = .87). En síntesis, los hallazgos nos permiten concluir de manera precisa el efecto del constructo *agotamiento emocional* sobre los ítems de la escala y aportan evidencia contundente de la estructura interna unidimensional de la ECE.

En términos de la invarianza de medición de la versión con siete ítems de la Escala de Cansancio Emocional, estos parecen ser invariantes por los grupos comparados. Es decir, que se pueden hacer comparaciones por género, edad, grado que estudia, ciclo académico y horas semanales de estudio porque el agotamiento emocional parece ser interpretado de igual forma por los participantes del estudio. El establecimiento de invarianza de medición de la Escala de Cansancio Emocional es evidencia de la validez de este instrumento (Chudowsky & Behuniak, 1998).

En lo que respecta a la consistencia interna

encontrada ($\alpha = .883$; $\omega = .883$), es comparable con las reportadas en otros estudios y que fluctúan entre .83 y .90 (Bonilla et al., 2009; Dominguez-Lara, 2014; Fontana, 2011; González-Ramírez & Landero-Hernández, 2007; Martínez-Líbano et al., 2022; Ramos-Campos et al., 2005). Los coeficientes y niveles de confiabilidad alcanzados son satisfactorios en un sentido amplio, especialmente si se tiene en cuenta la interacción entre el reducido número de ítems de la escala, el tamaño de la muestra y el valor alcanzado (Ponterotto & Ruckdeschel, 2007). Debido a que los coeficientes son altos ($\geq .85$), la posibilidad de error de medición todavía se puede considerar baja. Estos niveles demuestran la utilidad de la Escala de Cansancio Emocional para aplicaciones grupales y en las que se necesitan decisiones sobre los niveles de agotamiento emocional individuales (Ponterotto & Ruckdeschel, 2007). Dada la similitud de los coeficientes α y ω , se presume que cualquier diferencia entre las cargas factoriales fue trivial (Hayes & Coutts, 2020), y no tuvo un efecto significativo en la distancia entre un coeficiente y otro. El grado de igualdad de las cargas factoriales de los ítems está frecuentemente relacionado con esta distancia, condición conocida como tau-equivalencia para validar el coeficiente (Green & Yang, 2009; Hayes & Coutts, 2020). Esta cercanía implica que la evaluación de la consistencia interna se puede hacer bien con el coeficiente, sin necesidad de modelado SEM, o bien con metodologías de modelado SEM para estimar el coeficiente (Rosario-Hernández, Rovira-Millán, & Merino-Soto, 2021).

En términos teóricos, la propuesta unidimensional de la ECE valora ciertos aspectos del agotamiento emocional en el contexto académico que influyen en su medición y se asocian directamente entre sí. Es decir, el agotamiento emocional no se manifiesta en diferentes áreas del trabajo académico aisladamente, sino que es una

valoración global de toda la experiencia estudiantil y de cómo esta produce o no fatiga emocional. Dicha valoración incluye el esfuerzo invertido versus los logros alcanzados (ítem 2), el estado de ánimo del estudiante (ítem 3), el insomnio a causa de los estudios (ítem 4), el dolor de cabeza y otras molestias fisiológicas que interfieren en el rendimiento del estudiante (ítem 5), la falta de concentración (ítem 6), el agotamiento emocional (ítem 7) y el cansancio físico (ítem 8). Tomando en consideración estos aspectos, la utilización de este instrumento contribuiría en la adaptación de políticas públicas en Puerto Rico que se atemperen a las necesidades y demandas emergentes de los estudiantes; más aún en tiempos de pandemia por COVID-19 y los retos asociados a la misma (Rosario-Rodríguez et al., 2020).

En cuanto a las implicaciones prácticas, se debe destacar que es la primera vez que se valida en Puerto Rico una escala para medir el agotamiento emocional en universitarios de posgrado. Esto toma relevancia cuando se ha evidenciado contundentemente que el agotamiento emocional conduce al deterioro físico y psicológico de las personas, y está asociado al estrés, ansiedad, depresión y alteración de la autoeficacia en el contexto académico (Dominguez-Lara, 2014). Lamentablemente, son escasas las investigaciones en Puerto Rico sobre este tema y la ECE puede contribuir a visibilizar este estado emocional en los estudiantes del país. Del mismo modo, estas investigaciones podrían promover programas de prevención y tratamiento en las universidades del país. Por otra parte, la ECE, dada su brevedad y facilidad de comprensión, podría utilizarse en los centros de ayuda al estudiantado de las universidades como medida para detectar indicadores tempranos de agotamiento emocional y poder intervenir en los mismos. En fin, proponemos la ECE como una herramienta práctica y útil en el quehacer clínico y académico de las

universidades.

Este estudio no está exento de limitaciones. Por ejemplo, la muestra fue por conveniencia y no fue aleatoria, la recopilación de datos no fue estandarizada y se realizó en línea. Todo esto puede afectar las medias del estudio y aumentar el error estándar de medición. Sin embargo, el haber dividido la muestra de forma aleatoria en una muestra de calibración y otra de validación permitió realizar una validación cruzada de nuestros hallazgos (Browne & Cudeck, 1989; Cudeck & Browne, 1983; Whittaker & Stapleton, 2006). En el futuro se deben realizar esfuerzos para reclutar una muestra representativa de la mayoría de los programas de posgrado de Puerto Rico. También, se deben realizar estudios para examinar la validez convergente, divergente y predictiva de la ECE. Por último, según encontramos en nuestra revisión de literatura, sería interesante realizar estudios empíricos que ayuden a entender la relación intrínseca del agotamiento emocional y la depresión.

Referencias

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Ato, M., López-García, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038-1059. doi: [10.6018/analesps.29.3.178511](https://doi.org/10.6018/analesps.29.3.178511)
- Bianchi, R., Schonfeld, I. S., & Laurent, E. (2015). Is it time to consider the “burnout syndrome” a distinct illness? *Frontiers in Public Health*, 3, 158. doi: [10.3389/fpubh.2015.00158](https://doi.org/10.3389/fpubh.2015.00158)
- Bonilla-Muñoz, M. P., Lira-Mendiola, G. L., Balcázar-Nava, P., Enríquez-Bielma, J. F., & Gurrola-Peña, G. M. (2009). Adaptación de la Escala de Cansancio Emocional en adolescentes mexicanos del nivel superior. *Interpsiquis*, 10. Recuperado de <http://psiqui.com/1-4314>
- Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, 24(4), 445-455. doi: [10.1207/s15327906mbr2404_4](https://doi.org/10.1207/s15327906mbr2404_4)
- Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (3rd ed.). New York, NY: Routledge. doi: [10.4324/9781315757421](https://doi.org/10.4324/9781315757421)
- Caballero-Domínguez, C., González-Gutiérrez, O., & Palacio-Sañudo, J. (2015). Relación del burnout y el engagement con depresión, ansiedad y rendimiento académico en estudiantes universitarios. *Revista Científica Salud Uninorte*, 31(1), 59-69. doi: [10.14482/sun.31.1.5085](https://doi.org/10.14482/sun.31.1.5085)
- Caballero-Domínguez, C., Hederich-Martínez, C., & Palacio-Sañudo, J. E. (2010). El burnout académico: Delimitación del síndrome y factores asociados con su aparición. *Revista Latinoamericana de Psicología*, 42(1), 131-146. Recuperado de <http://revistalatinamericanadepsicologia.konradlorenz.edu.co>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464-504. doi: [10.1080/10705510701301834](https://doi.org/10.1080/10705510701301834)
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. doi: [10.1207/S15328007SEM0902_5](https://doi.org/10.1207/S15328007SEM0902_5)
- Chudowsky, N., & Behuniak, P. (1998). Using focus groups to examine the consequential aspects of validity. *Educational Measurement: Issues and Practice*, 17(4), 28-38.
- Cortina-Rodríguez, G., & Afanador, Y. (2020). Burnout in the Clinical Personnel of Puerto Rico during the COVID-19 Pandemic. *Preprints*, 2020070451. doi: [10.20944/preprints202007.0451.v1](https://doi.org/10.20944/preprints202007.0451.v1)
- Cudeck, R., & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral*

- Research*, 18(2), 147-167. doi: [10.1207/s15327906mbr1802_2](https://doi.org/10.1207/s15327906mbr1802_2)
- Dominguez-Lara, S. (2018). Invarianza longitudinal y error transitorio de una medida de burnout académico en universitarios. *Avaliação Psicológica*, 17(3), 311-320. doi: [10.15689/ap.2018.1703.14421.04](https://doi.org/10.15689/ap.2018.1703.14421.04)
- Dominguez-Lara, S. A. (2014). Escala de Cansancio Emocional: Estructura factorial y validez de los ítems en estudiantes de una universidad privada. *Avances en Psicología*, 22(1), 89-97. doi: [10.33539/avpsicol.2014.v22n1.275](https://doi.org/10.33539/avpsicol.2014.v22n1.275)
- Fernandes-Fontes, F. (2020). Herbert J. Freudenberger e a constituição do burnout como síndrome psicopatológica. *Memorandum: Memória e História em Psicologia*, 37. doi: [10.35699/1676-1669.2020.19144](https://doi.org/10.35699/1676-1669.2020.19144)
- Fontana, S. A. (2011). Estudio preliminar de las propiedades psicométricas de la Escala de Desgaste Emocional para estudiantes universitarios. *Revista Argentina de Ciencias del Comportamiento*, 3(2), 44-48. Recuperado de <https://revistas.unc.edu.ar>
- Fornell, C., & Bookstein, F. L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research*, 19(4), 440-452. doi: [10.2307/3151718](https://doi.org/10.2307/3151718)
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. doi: [10.2307/3151312](https://doi.org/10.2307/3151312)
- González-Ramírez, M. T., & Landero-Hernández, R. (2007). Escala de Cansancio Emocional (ECE) para estudiantes universitarios: Propiedades psicométricas en una muestra de México. *Anales de Psicología*, 23(2), 253-257. Recuperado de <https://revistas.um.es/analesps>
- Green, S. B., & Yang, Y. (2009). Commentary on Coefficient Alpha: A Cautionary Tale. *Psychometrika*, 74(1), 121-135. doi: [10.1007/s11336-008-9098-4](https://doi.org/10.1007/s11336-008-9098-4)
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). *Multivariate data analysis* (8^{va} ed.). Boston, MA: Cengage.
- Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. *But... Communication Methods and Measures*, 14(1), 1-24. doi: [10.1080/19312458.2020.1718629](https://doi.org/10.1080/19312458.2020.1718629)
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. doi: [10.1080/10705519909540118](https://doi.org/10.1080/10705519909540118)
- InformedHealth.org [Internet]. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG); 2006-. Depression: What is burnout? [Updated 2020 Jun 18]. <https://www.ncbi.nlm.nih.gov/books/NBK279286>
- Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, 44(3), 486-507. doi: [10.1177/0049124114543236](https://doi.org/10.1177/0049124114543236)
- Kline, P. (1994). *An Easy Guide to Factor Analysis* (1st ed.). Oxfordshire, UK: Routledge. doi: [10.4324/9781315788135](https://doi.org/10.4324/9781315788135)
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: Guilford Press.
- Li, C.-H. (2016a). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavioral Research Methods*, 48(3), 936-949. doi: [10.3758/s13428-015-0619-7](https://doi.org/10.3758/s13428-015-0619-7)
- Li, C.-H. (2016b). The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological Methods*, 21(3), 369-387. doi: [10.1037/met0000093](https://doi.org/10.1037/met0000093)
- Littlewood-Zimmerman, H. F., & Bernal-García, E. R. (2011). *Mi primer modelamiento de ecuación estructural: LISREL*. Medellín, Colombia: Centro de Investigación en Comportamiento Organizacional (CINCEL).
- López-Osorio, E. A., Cano, C., & Salazar Ospina, V. (2020). *Caracterización del Síndrome de Burnout del personal de salud que labora en telemedicina, durante la pandemia COVID 19, en el convenio UT San Vicente*

- CES. [Tesis de posgrado, Universidad Cooperativa de Colombia]. Repositorio Institucional UCC. <https://repository.ucc.edu.co/handle/20.500.12494/20449>
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130-149. doi: [10.1037/1082-989X.1.2.130](https://doi.org/10.1037/1082-989X.1.2.130)
- Manassero-Mas, M. A., Vázquez-Alonso, A., Ferrer-Pérez, V. A., Fornés-Vives, J., & Fernández-Bennassar, M. C. (2003). *Estrés y burnout en la enseñanza*. Palma de Mallorca, España: UIB.
- Marsh, H. W., Balla, J. R., & Hau, K.-T. (1996). An evaluation of incremental fit indexes: A clarification of mathematical and empirical properties. In G. A. Marcoulides, & R. E. Schumacker (Eds.), *Advanced Structural Equation Modeling Techniques* (pp. 315-353). Mahwah, NJ: Lawrence Erlbaum.
- Martínez-Líbano, J., Yeomans, M.-M., & Oyanedel, J.-C. (2022). Psychometric Properties of the Emotional Exhaustion Scale (ECE) in Chilean Higher Education Students. *European Journal of Investigation in Health, Psychology and Education*, 12(1), 50-60. doi: [10.3390/ejihpe12010005](https://doi.org/10.3390/ejihpe12010005)
- Martínez-Pérez, A. (2010). El síndrome de Burnout. Evolución conceptual y estado actual de la cuestión. Vivat Academia. *Revista de Comunicación*, 13(112), 42-80. doi: [10.15178/va.2010.112.42-80](https://doi.org/10.15178/va.2010.112.42-80)
- Maslach, C. (1976). Burned-Out. *Human Behavior*, 5, 7-22. Recuperado de https://www.researchgate.net/publication/263847499_Burned-Out
- Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior*, 2(2), 99-113. doi: [10.1002/job.4030020205](https://doi.org/10.1002/job.4030020205)
- Maslach, C., Jackson, S. E., & Leiter, M. P. (1997). *Maslach Burnout Inventory Manual* (3rd ed.). Sunnyvale, CA: Consulting Psychologists Press.
- Ministerio del Trabajo. (2015). Síndrome de agotamiento laboral - Burnout: Protocolo de prevención y actuación. Bogotá, Colombia: Fundación Cultural Javeriana de Artes Gráficas - JAVEGRAF.
- Morán-Astorga, M. C. (2003). *Relación entre variables de personalidad y estrategias de afrontamiento del estrés laboral*. (Tesis doctoral). Recuperado de <https://buleria.unileon.es>
- Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Perceptual and Motor Skills*, 105(3), 997-1014. doi: [10.2466/pms.105.3.997-1014](https://doi.org/10.2466/pms.105.3.997-1014)
- Quintero-Idárraga, S., & Hernández-Calle, J. (2021). Síntomas de depresión asociados al síndrome de burnout y a condiciones socio laborales de docentes de colegios públicos de Envigado (Colombia). *Psicología desde el Caribe*, 38(1), 133-147. doi: [10.14482/psdc.38.1.158.724](https://doi.org/10.14482/psdc.38.1.158.724)
- Ramos-Campos, F., Manga-Rodríguez, D., & Morán-Astorga, C. (2005). Escala de Cansancio Emocional (ECE) para estudiantes universitarios: Propiedades psicométricas y asociación. *Interpsiquis*, 6, 1-9. Recuperado de <http://psiqui.com/1-2898>
- Rigdon, E. E. (1995). A necessary and sufficient identification rule for structural models estimated in practice. *Multivariate Behavioral Research*, 30(3), 359-383. doi: [10.1207/s15327906mbr3003_4](https://doi.org/10.1207/s15327906mbr3003_4)
- Rodríguez-Colón, M. (2016). El síndrome de quemarse por el trabajo y los profesionales de consejería del nivel postsecundario. *Revista Griot*, 8(1), 42-59. Recuperado de: <https://revistas.upr.edu>
- Rosales-Ricardo, Y., & Rosales-Paneque, F. R. (2013). Burnout estudiantil universitario: Conceptualización y estudio. *Salud Mental*, 36(4), 337-345. doi: [10.17711/sm.0185-3325.2013.041](https://doi.org/10.17711/sm.0185-3325.2013.041)
- Rosario-Hernández, E., Rovira-Millán, L. V., & Merino-Soto, C. (2021). Review of the internal structure, psychometric properties, and measurement invariance of the Work-Related Rumination Scale - Spanish Version. *Frontiers in Psychology*, 12. doi: [10.3389/fpsyg.2021.774472](https://doi.org/10.3389/fpsyg.2021.774472)
- Rosario-Rodríguez, A., González-Rivera, J. A., Cruz-Santos, A., & Rodríguez-Ríos, L. (2020). Demandas

- tecnológicas, académicas y psicológicas en estudiantes universitarios durante la pandemia por COVID-19. *Revista Caribeña de Psicología*, 4(2), 176-185. doi: [10.37226/rcp.v4i2.4915](https://doi.org/10.37226/rcp.v4i2.4915)
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36. doi: [10.18637/jss.v048.i02](https://doi.org/10.18637/jss.v048.i02)
- Sánchez, E., & Sánchez, M. (1998). Los modelos de estructuras de covarianza como método de validación de constructo. En V. Manzano, & M. Sánchez. (Eds.). *Investigación del Comportamiento. Innovaciones metodológicas y estrategias de docencia* (pp. 101-112). Sevilla: Instituto Psicosociológico Andaluz de Investigaciones.
- Schaufeli, W. B., Leiter, M. P., & Maslach, C. (2009). Burnout: 35 years of research and practice. *Career Development International*, 14(3), 204-220. doi: [10.1108/13620430910966406](https://doi.org/10.1108/13620430910966406)
- Shi, D., DiStefano, C., Maydeu-Olivares, A., & Lee, T. (2021). Evaluating SEM Model Fit with Small Degrees of Freedom. *Multivariate Behavioral Research*, 1-36. doi: [10.1080/00273171.2020.1868965](https://doi.org/10.1080/00273171.2020.1868965)
- Szigeti, R., Balázs, N., Bikfalvi, R., & Urbán, R. (2017). Burnout and depressive symptoms in teachers: Factor structure and construct validity of the Maslach Burnout inventory educators survey among elementary and secondary school teachers in Hungary. *Stress and Health*, 33(5), 530-539. doi: [10.1002/smi.2737](https://doi.org/10.1002/smi.2737)
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1-10. doi: [10.1007/BF02291170](https://doi.org/10.1007/BF02291170)
- Whittaker, T. A., & Stapleton, L. M. (2006). The performance of cross-validation indices used to select among competing covariance structure models under multivariate nonnormality conditions. *Multivariate Behavioral Research*, 41(3), 295-335. doi: [10.1207/s15327906mbr4103_3](https://doi.org/10.1207/s15327906mbr4103_3)

Inventory of Interpersonal Problems-32 (IIP-32): Psychometric properties and normative data in a clinical sample from Argentina

Inventario de Problemas Interpersonales-32 (IIP-32): Propiedades psicométricas y datos normativos en una muestra clínica de Argentina

Juan Martín Gómez-Penedo^{1,2*}, Javier Fernández-Álvarez^{3,1}, Mariana Maristany^{3,4}, Agustín Freiberg-Hoffmann^{1,2}

1 - Facultad de Psicología, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, Argentina.

2 - Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Ciudad Autónoma de Buenos Aires, Argentina.

3 - Asociación Aiglé, Valencia, Spain.

4 - Fundación Aiglé, Ciudad Autónoma de Buenos Aires, Argentina.

Recibido: 30/05/2022 Revisado: 07/06/2022 Aceptado: 20/06/2022

Introduction
Methods
Results
Discussion
References

Abstract

This study analyzes the psychometric properties of the Inventory of Interpersonal Problems-32 (IIP-32), the most widely used instrument to measure relational difficulties. A sample of 2128 participants completed the IIP-32 and two additional measures of interpersonal difficulties before starting a psychotherapy treatment. To evaluate reliability, we analyzed internal consistency and item homogeneity. We analyzed the construct validity of IIP-32 through a confirmatory factor analysis and the concurrent validity through correlations between the IIP-32 and other measures of interpersonal deficits. The results of the study show excellent internal consistency and homogeneity of the items in the IIP-32. Furthermore, the results show construct validity as well as concurrent validity of the instrument. In sum, the results of this paper rank the IIP-32 as a reliable and valid instrument with important clinical implications to measure interpersonal difficulties in Argentina.

Keywords: *Inventory of Interpersonal Problems, IIP-32, validity, reliability, Argentina*

Resumen

Este estudio analiza las propiedades psicométricas del Inventario de Problemas Interpersonales-32 (IIP-32), la medida más utilizada para evaluar dificultades en las relaciones interpersonales. Una muestra de 2128 participantes completaron el IIP-32 y otras medidas de dificultades interpersonales antes de comenzar un tratamiento psicoterapéutico. Para evaluar la confiabilidad se analizaron medidas de consistencia interna y de homogeneidad de ítems. Se analizó la validez del constructo mediante un análisis factorial confirmatorio y la validez concurrente mediante correlaciones entre el IIP-32 y otras medidas de problemáticas interpersonales. Los resultados muestran excelentes niveles de consistencia interna y homogeneidad de ítems. A su vez, los resultados muestran la validez de constructo y validez de concurrente del instrumento. En síntesis, los resultados del trabajo posicionan al IIP-32 como un instrumento de confiabilidad y validez para estudiar dificultades interpersonales con importantes implicancias para la práctica clínica en Argentina.

Palabras clave: *Inventario de Problemas Interpersonales, IIP-32, validez, confiabilidad, Argentina*

*Correspondence to: Juan Martín Gómez Penedo. Lavalle 2353, Buenos Aires, Argentina. Telephone: +54 9 11 6525 7053. E-mail: jmgomezpenedo@gmail.com

How to cite: Gómez-Penedo, J. M., Fernández-Álvarez, J., Maristany, M., & Freiberg-Hoffmann, A. (2022). Inventory of Interpersonal Problems-32 (IIP-32): Psychometric properties and normative data in a clinical sample from Argentina. *Revista Evaluar*, 22(2), 64-75. Retrieved from <https://revistas.unc.edu.ar/index.php/revaluar>

Participaron en la edición de este artículo: Melani Hajduczyk, María Eugenia Maiorana, Eugenia Barrionuevo, Pablo Carpintero, Stefano Macri, Florencia Ruiz, Mónica Serppe, Ricardo Hernández, Alicia Molinari.

Introduction

Interpersonal relationships have a fundamental relevance in people's daily lives (Horowitz, 2004; Lieberman, 2013; Wilson, 2012). Relational problems have been found to be associated with higher levels of psychopathology (Luo, Nuttall, Locke, & Hopwood, 2018; Segrin, 2001) and lower levels of life satisfaction (Froh et al., 2007). Meta-analytic studies have even suggested that difficulties in interpersonal relationships are a robust predictor of mortality comparable to well-established factors such as smoking and obesity (Holt-Lunstad, Smith, & Layton, 2010). For this reason, it is important to have valid and reliable instruments that enable an assessment of the degrees and nature of patients' interpersonal difficulties.

Considering that psychotherapy is by definition an interpersonal practice (Wampold & Imel, 2015), the relevance of instruments of this nature is particularly higher in the specific field of psychotherapy practice and research. On one hand, interpersonal problems represent one of the main reasons for consultation among people seeking treatment (Horowitz, 2004). On the other hand, relational difficulties have been identified as a dimension of change when evaluating the results of psychotherapy (Berghout, Zevalkink, Katzko, & de Jong, 2012; Salzer, Pincus, Winkelbach, Leichsenring, & Leibing, 2011). In other words, psychotherapy is expected to improve the way people relate to others. Finally, regarding the type of interpersonal problems that a person has, they can provide relevant information to determine the type of therapeutic process or to approach what it would be more beneficial for each patient (Gómez-Penedo, Constantino, Coyne, Westra, & Antony, 2017; Gómez-Penedo et al., 2020; Newman, Jacobson, Erickson, & Fisher, 2017; Zilcha-Mano, Muran, Eubanks, Safran, & Winston, 2018).

This provides evidence in favor of the personalization of treatments and precision in mental health (Lutz, Schwartz, Gómez-Penedo, Boyle, & Deisenhofer, 2020; Zilcha-Mano, 2021).

Although there are different instruments to assess relational difficulties, the most widely used internationally is the Inventory of Interpersonal Problems (IIP; Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988). The IIP is considered a useful dimensional instrument to measure the interpersonal problems of patients at the beginning of the therapy and to assess its results (Maristany, 2008). The first version of this instrument had 127 items distributed in six subscales (i.e., *assertive*, *sociable*, *intimate*, *submissive*, *responsible* and *controlling*). Later, Horowitz et al. (2000) conducted a revised version of the IIP, which included 64 items (IIP-64) divided equally into eight subscales (*domineering*, *intrusive*, *overly nurturing*, *exploitable*, *nonassertive*, *socially inhibited*, *cold*, and *vindictive*). The subscales of this revised version configured octants distributed in a circumplex model consistent with classical interpersonal theories (Leary, 1957; Sullivan, 1955). In this way, the eight typologies of interpersonal problems are distributed around the two basic interpersonal dimensions of *agency* and *communion* (see Figure 1). The *communion* dimension describes the extent to which a person needs to establish close relationships with others (having issues of coldness/indifference at the negative pole and issues of being too dependent at the positive pole). The *agency* dimension describes the extent to which a person needs to influence other people or be influenced by others (having issues of submission at the negative pole and issues of being too dominant at the positive pole).

The IIP-64 has been widely used in clinical research and has been translated and adapted to different languages and cultures (Stiles & Hoglend, 1994; Thomas, Brähler & Strauß, 2011;

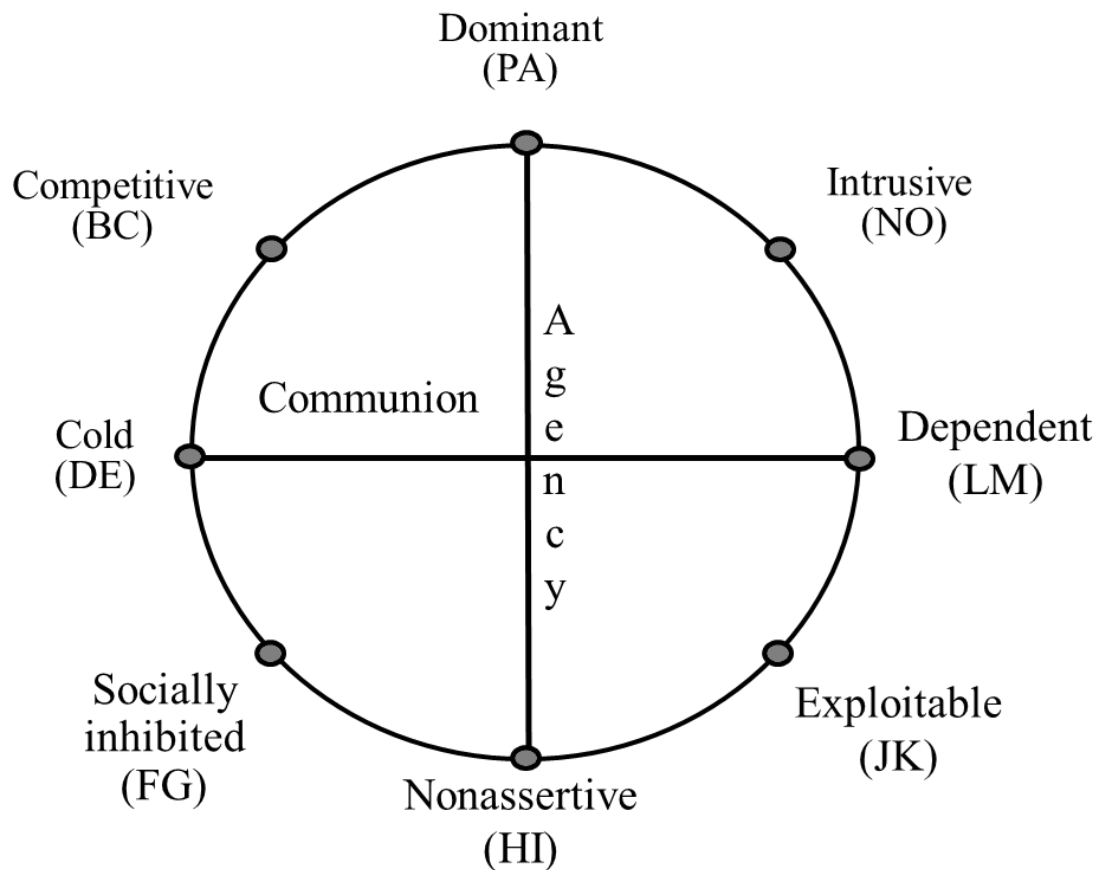


Figure 1. Subscales and dimensions of the circumplex model of the Inventory of Interpersonal Problems (Horowitz et al., 2000).

Vanheule, Desmet, & Rosseel, 2006) including the Argentinian context, where it showed adequate psychometric properties (Maristany, 2005). Despite the spread of the IIP-64 internationally, the number of items in the instrument remained a limitation for its use in clinical practice due to the burden it represented for patients. As a consequence, a shorter version of the IIP with 32 items was created (IIP-32; Barkham, Hardy, & Startup, 1996). This version can be considered ideal in clinical practice because it is brief and preserves the nature of the IIP-64 (McEvoy, Burgess, Page, Nathan, & Fursland, 2013). The IIP-32 has been adapted to countries such as Germany (Thomas et al., 2011), Norway (Vanheule et al., 2006), Italy (Lo Coco et al., 2018), Spain (Salazar, Martí,

Soriano, Beltram, & Adam, 2010) and Portugal (Faustino & Vasco, 2020), among others. However, in the literature analyzed, we could not find any study that analyzed the psychometric properties of the IIP-32 in the Latin American context.

In this framework, the aim of this paper is to analyze the psychometric properties of the IIP-32 in a clinical sample from Argentina. In particular, the levels of reliability (internal consistency and item homogeneity) and validity (construct validity and concurrent validity) of the inventory are analyzed. An instrument with adequate levels of reliability and validity would be a relevant contribution to both research and clinical practice in mental health.

Methods

Participants

The sample consisted of 2128 participants who completed the IIP-32 during the years 2009 and 2021. The mean age of the participants was 33.7 years (SD = 12.42). 62% of the participants were women and 38% were men. There was no participant diagnostic data or other clinical data.

Instruments

Inventory of Interpersonal Problems (IIP-32; Horowitz, 2000). This scale is a brief 32-item version of the Interpersonal Problems Inventory (Horowitz et al., 1988), which assesses people's difficulties in their relationships with others. The subjects of study are presented with a list of problems that people tend to have when interacting with others, being either excessive behaviors or inhibitions. Each subject has to assess whether each item represents a problem when interacting with a significant person in their life in the last two weeks. The answers run on a scale from 0 (*Not at all*) to 4 (*Very much*). The 32 items are distributed in eight subscales of interpersonal problems: *domineering, intrusive, overly nurturing, exploitable, nonassertive, socially inhibited, cold* and *vindictive*. Some of the items are: *It is difficult for me to say no to other people* or *It is difficult for me to experience a feeling of love for another person*. In addition to the subscales, the instrument helps to compute a total score of interpersonal distress. Additionally, it can calculate the two interpersonal dimensions of *agency* and *communion*, based on weighted combinations of the eight subscales. For the calculation of *agency* = .25 (domineering - nonassertive + .71[intrusive + vindictive - socially inhibited - exploitable]); *communion* = .25 (overly nurturant - cold

+ .71[intrusive - vindictive - socially inhibited + exploitable]; Ruiz et al., 2004). The IIP-64 has a Spanish adaptation in Argentina with excellent psychometric properties (Maristany, 2005).

Symptom Checklist (SCL-90; Derogatis, 1994) [Spanish version; Casullo & Pérez, 2008]. This 90-item self-report inventory measures the presence and severity of psychiatric symptomatology. The instrument has nine primary dimensions and three global indexes that represent the levels of psychological distress. Some dimensions, for example, include symptoms such as *anxious and depressive* or *psychotic and somatic*. Each item is answered on a five-point Likert-type scale ranging from 0 (*Not at all*) to 4 (*Extremely*). Especially important for this study is the fact that the SCL-90 includes an interpersonal sensitivity dimension (e.g., *Feeling critical of others* or *Feeling shy or uneasy with the opposite sex*). The SCL-90 has been adapted in Argentina and has demonstrated good levels of internal consistency for all scales and for the general index (Casullo & Pérez, 2008).

Outcome Questionnaire (OQ.45; Lambert et al., 1996) [Spanish version; von Bergen, 2002]. The OQ.45 is a 45-item self-report questionnaire that measures progress and results in psychotherapy and monitors the evolution of patients with successive measurements. The items are grouped into three subscales that describe the periodicity in experiences of distress in patients in different areas: 1) symptomatic distress, 2) interpersonal relations and 3) social role. The items are rated on a five-point Likert-type scale ranging from 1 = *Never* to 5 = *Almost always*. In this study, the subscale of interpersonal relations was used, which include items such as *I get along well with others*. The Spanish version of the instrument has shown adequate psychometric properties in the Argen-

tine context (Fernández-Álvarez, Hirsh, Maristany, & Torrente, 2005).

Procedures

Patients attending a private clinic for psychotherapeutic treatment completed the IIP-32 as part of the diagnostic and clinical assessment process. In addition, prior to their treatment, the patients completed the SCL-90 interpersonal sensitivity subscale and the OQ.45 interpersonal relations subscale. All participants signed a written consent under a confidentiality clause for the use of these data for research purposes.

Data analysis

For the data analysis, LISREL 8.8 software (Jöreskog & Sörbom, 1993) and open-source software R (R Core Team, 2021), particularly with the Hmisc (Harrell Jr, 2021) and psych (Revelle, 2017) packages, were used. Internal consistency and item homogeneity were analyzed as reliability measures of the IIP-32. The internal consistency was established by Cronbach's alpha. Since the instrument items have an ordinal measurement scale and a five-point response scale, in addition to Cronbach's alpha, the ordinal alpha is presented (Freiberg-Hoffmann, Stover, de la Iglesia, & Fernández-Liporace, 2013; Gadermann, Guhn, & Zumbo, 2012). Values between .70 and .90 are considered acceptable (Tavakol & Dennick, 2011). To analyze item homogeneity, we calculated corrected item-scale correlations. Correlations in the range from .30 to .80 are indicators of adequate levels of item homogeneity (Ratray & Jones, 2007).

To measure the validity of the instrument, construct and concurrent validity were evaluated.

For the construct validity, a confirmatory factor analysis was performed followed by an analysis of factorial invariance. For this purpose, the diagonally weighted least squares (DWLS) estimation method was employed, an alternative to the maximum likelihood method for ordinal items and large sample sizes (Kiliç & Doğan, 2021; Li, 2016; Mîndrilă, 2010). To interpret the model fit, comparative fit index (CFI), Tucker-Lewis index (TLI) and root mean square error of approximation (RMSEA) were used. Values above .90 in CFI and TLI (Schumacker & Lomax, 2016) and below .08 in RMSEA (Schumacker & Lomax, 2016) were indicators of an adequate fit of the model. Due to the nature of the instrument, a bifactor model was adjusted (Reise, 2012). Additionally, the bias of the bifactor model items in the estimation of unidimensional measures was evaluated through the percent of uncontaminated correlations (PUC) and average relative parameter bias (ARPB) statistics whose values above .80 and below 15% respectively were indicators of low bias (Davidov, Billiet, Meuleman, & Schmidt, 2018; Hammer, 2016; Rodriguez, Reise, & Haviland, 2016).

The analysis of factorial invariance was performed by segmenting the sample according to sex variable—men and women—. Three nested models with different levels of restriction (configural, metric, scalar) were analyzed. The invariance was interpreted with the CFI and RMSEA indexes whose differences should be less than .01 and .015 respectively (Putnick & Bornstein, 2016; Rial-Boubeta, Varela-Mallou, Abalo-Piñeiro, & Levy-Mangin, 2006).

For the concurrent validity study, Pearson product-moment correlations were calculated between the IIP-32 total score, the SCL-90 interpersonal sensitivity scale and the OQ.45 interpersonal relationship distress scale. For evidence of adequate concurrent validity, correlation values in the range of .50 to .85 are expected (Rial-Boubeta

et al., 2006).

The code of the study analyses and its results will be published on the open science framework page (after the anonymous peer review process), in markdown format (Allaire et al., 2022).

Results

Descriptive analysis (Sample 1)

Table 1 presents the descriptive analyses of all the items of the IIP-32. Table 2 reports de-

Table 1

Descriptive analysis and corrected item-scale correlations of the IIP-32 items.

Item	Mean	SD	Scale	<i>r</i>
Item 1	2.09	1.29	JK	.49
Item 2	1.48	1.37	FG	.72
Item 3	1.07	1.17	NO	.39
Item 4	1.32	1.33	HI	.51
Item 5	1.39	1.36	FG	.66
Item 6	1.41	1.30	HI	.49
Item 7	1.29	1.16	HI	.53
Item 8	1.36	1.28	JK	.40
Item 9	1.17	1.31	FG	.70
Item 10	1.17	1.27	DE	.63
Item 11	0.76	1.07	DE	.54
Item 12	1.61	1.30	HI	.53
Item 13	1.05	1.31	DE	.56
Item 14	0.60	1.03	BC	.59
Item 15	1.06	1.18	DE	.58
Item 16	0.89	1.15	BC	.61
Item 17	1.27	1.30	BC	.41
Item 18	0.67	1.10	BC	.56
Item 19	1.23	1.32	FG	.52
Item 20	2.01	1.26	JK	.36
Item 21	1.54	1.32	NO	.49
Item 22	0.80	1.07	PA	.52
Item 23	1.77	1.30	LM	.53
Item 24	0.82	1.09	NO	.36
Item 25	1.08	1.19	PA	.49
Item 26	1.66	1.30	LM	.60
Item 27	1.58	1.25	LM	.57
Item 28	1.00	1.17	PA	.38
Item 29	1.05	1.16	NO	.57
Item 30	0.94	1.10	PA	.68
Item 31	0.78	1.05	JK	.40
Item 32	1.92	1.27	LM	.41

Note. SD = Standard deviation. For scale references see Figure 1.

Table 2

Descriptive analyses of the IIP-32 total score and scales.

Scales	Mean	SD	Range
PA	0.94	0.81	[0; 4]
BC	0.88	0.87	[0; 4]
DE	1.02	0.94	[0; 4]
FG	1.29	1.08	[0; 4]
HI	1.40	0.94	[0; 4]
JK	1.54	0.84	[0; 4]
LM	1.73	0.95	[0; 4]
NO	1.11	0.84	[0; 4]
Total	1.24	0.56	[0; 3.31]

Table 3

Cronbach's alpha and ordinal alpha for IIP-32 scales.

Scales	Cronbach's alpha	Ordinal alpha
PA	.70	.76
BC	.74	.81
DE	.77	.83
FG	.82	.86
HI	.73	.77
JK	.63	.68
LM	.73	.77
NO	.67	.72
TOTAL	.87	.97

scriptive analyses of the scales and the total score of the IIP-32.

Reliability

Internal consistency. The items of the IIP-32 presented a Cronbach's alpha of .88. Due to the ordinal level of measurement of the items, the ordinal alpha of the IIP-32 items was computed with a value of .90. Table 3 presents the Cronbach's alphas and ordinal alphas at the level of the eight scales of the IIP-32.

Item homogeneity. Table 1 shows the item-scale correlations. None of the items of the IIP-32 presented correlations lower than .30 or higher than .80 with the items of its scale. The items with greatest item-total adjusted correlations of each scale were the following: item 30 (PA, $r = .68$), item 14 (BC, $r = .59$), item 10 (SD, $r = .63$), item

Table 4

Parameters and coefficients of determination of the confirmatory factor model.

Scale - Items	λ Subscale	Scale - Items	λ Total	R^2
PA – Item 22	.732	T – Item 22	.301	.626
PA – Item 25	.511	T – Item 25	.313	.359
PA – Item 28	.370	T – Item 28	.305	.230
PA – Item 30	.730	T – Item 30	.320	.635
BC – Item 14	.612	T – Item 14	.595	.650
BC – Item 16	.603	T – Item 16	.491	.605
BC – Item 17	.509	T – Item 17	.196	.297
BC – Item 18	.543	T – Item 18	.530	.575
DE – Item 10	.776	T – Item 10	.481	.833
DE – Item 11	.222	T – Item 11	.719	.567
DE – Item 13	.477	T – Item 13	.497	.475
DE – Item 15	.264	T – Item 15	.702	.562
FG – Item 2	.665	T – Item 2	.566	.762
FG – Item 5	.603	T – Item 5	.535	.650
FG – Item 9	.601	T – Item 9	.616	.741
FG – Item 19	.259	T – Item 19	.651	.491
HI – Item 4	.363	T – Item 4	.604	.497
HI – Item 6	.212	T – Item 6	.602	.407
HI – Item 7	.269	T – Item 7	.610	.445
HI – Item 12	.395	T – Item 12	.581	.493
JK – Item 1	.588	T – Item 1	.483	.579
JK – Item 8	.275	T – Item 8	.511	.337
JK – Item 20	.355	T – Item 20	.306	.220
JK – Item 31	.280	T – Item 31	.598	.436
LM – Item 23	.592	T – Item 23	.394	.506
LM – Item 26	.691	T – Item 26	.308	.572
LM – Item 27	.683	T – Item 27	.267	.537
LM – Item 32	.448	T – Item 32	.240	.259
NO – Item 3	.435	T – Item 3	.421	.366
NO – Item 21	.733	T – Item 21	.084	.545
NO – Item 24	.408	T – Item 24	.284	.247
NO – Item 29	.745	T – Item 29	.298	.645

2 (FG, $r = .72$), item 7 (HI, $r = .53$), item 1 (JK, $r = .49$), item 26 (LM, $r = .60$) and item 29 (NO, $r = .59$). The average of these eight aggregated items correlated strongly with the total IIP-32 score ($r = .91$).

Validity

Construct validity. A bifactor model was tested. Table 4 shows the parameters of the confirmatory

factor model and the coefficient of determination for each item. The model verified an adequate fit to the empirical data with CFI indices of .93, TLI of .92 and RMSEA [CI90%] of .076 [.074, .078]. Acceptable PUC indices of .90 % and ARPB of 12% were obtained.

The factorial invariance of the bifactor model was tested by segmenting the sample according to sex. The metric equivalence of the model between men and women was verified (Table 5).

Table 5

Factorial invariance of the bifactor model segmenting the sample by sex.

	CFI	Δ CFI	RMSEA	Δ RMSEA
Configural	.941	-	.075 [.073, .077]	-
Metric	.938	.003	.074 [.072, .076]	.001
Scalar	.935	.006	.075 [.073, .077]	0

Concurrent validity. Pearson correlations showed significant direct associations of the IIP-32 with the OQ.45 interpersonal relationship distress scale ($r = .59, p < .001$) and with the SCL-90 interpersonal sensitivity scale ($r = .62, p < .001$). In both cases the correlations demonstrated a large effect size in the associations of the IIP-32 and the other measures of interpersonal distress (Dominguez-Lara, 2018).

Discussion

The aim of this study was to analyze the psychometric properties of the IIP-32 in a sample of patients from Argentina. For this purpose, these patients completed a version of the IIP in the diagnostic evaluation process, prior to initiating psychotherapeutic treatment in a private clinical center. The results of this study show adequate levels of reliability and validity of the instrument in the Argentinian context.

Internal consistency and item homogeneity were used to evaluate reliability. The instrument showed an excellent level of internal consistency in its total score, with a Cronbach's alpha of .88 and an ordinal alpha of .90 and adequate levels of internal consistency in six of the eight dimensions, with alphas in the range of .70 to .90 (Tavakol & Dennick, 2011). The *intrusive* (NO; alpha = .67) and *exploitable* (JK; alpha = .63) scales presented alphas below the expected range.

In addition, adequate levels of item homogeneity were observed in the instrument, with all items presenting corrected item-scale correlations in the suggested range of .30 to .80 (Ratray & Jones, 2007). The fact that the correlations are greater than .30 implies that there is some degree of minimal association between the items, while correlations below .80 suggest that the items are not redundant.

Both construct and concurrent validity of the IIP-32 were analyzed. The confirmatory factor analysis demonstrated adequate construct validity of the IIP-32 in the Argentine context, with all goodness-of-fit measures within the parameters suggested in the literature. The factor model obtained CFI (.93) and TLI (.92), values above .90 (Schumacker & Lomax, 2016), while the RMSEA index (.07) was below .08 (Schumacker & Lomax, 2016). Acceptable PIC and ARPB values were also obtained, which indicated low item bias in the bifactor model (Davidov et al., 2018). The analysis of factorial invariance verified that the structure of the instrument remains metrically equivalent in both male and female samples. In the analysis of the items, 17 of the 32 items contributed at least 50% of their variability to the explanation of the latent variables of the model (Kline, 2005). This would indicate that most of the items that make up the model have a high degree of representativeness to assess the construct of interpersonal problems.

Pearson's correlations between the IIP-32 total score and the subscales of interpersonal relationship distress of the OQ.45 ($r = .59$) and interpersonal sensitivity of the SCL-90 ($r = .62$) demonstrated concurrent validity between IIP-32 scores and instruments that measure close constructs. By being in the range of .50 to .85, the correlations indicate that the constructs are associated, but not identical (Rial-Boubeta et al., 2006).

In summary, the results of this work present the IIP-32 as a reliable and valid instrument for the evaluation of interpersonal problems in patients at the beginning of their psychotherapy treatment. The IIP-32 encourages a global evaluation of the degree of difficulty that people have in establishing and maintaining relationships and also helps to perform a thorough analysis of the type of difficulties that an individual may have in

a relationship. The characterization of these difficulties has great clinical relevance when evaluating patients, designing treatments and analyzing their results. Unlike the original IIP (127 items) and the IIP-64 (64 items), the IIP-32 is a shorter version that involves less time for patients to complete. Therefore, the use of the IIP-32 would improve the measurements of interpersonal problems in clinical practice and its use for research (Gómez-Penedo et al., 2021).

It should be noted that this study has a number of limitations that will need to be addressed in future research. First, the instruments used in the validation of the IIP-32 were entirely self-report measures. Consequently, the correlations between measurements increase because they come from the same source. Future research should include measures that are completed by other sources (outside observers, therapists, acquaintances, etc.) to explore the concurrent validity of the IIP-32. These measures would provide detailed and complementary information about the interpersonal problems of the participants, providing valuable information in the study of the subject. Second, as it was previously mentioned, there was no information on the diagnosis of the participants. This limits the type of analyses used, since the results obtained cannot be compared according to the diagnosis of the participants. For example, significant differences could be found depending on whether patients are diagnosed with personality disorders or depressive disorders. For this reason, future research could be favored by incorporating, through a unified diagnostic criterion, systematized information about the diagnosis of the participants. It should be noted as well that in this study the sample used was made up of participants from a clinical population, which affects the generalization of the results. Therefore, a future investigation should include participants from both a clinical and non-clinical population.

Therefore, evaluation procedures could be used to analyze differences between groups. Finally, no repeated measures of the instrument were taken. Future research could study the instrument's sensitivity to change as well as detect and create criteria for clinically significant change that could be used in monitoring and feedback systems for therapists.

The results of this study place the IIP-32 as a reliable and valid instrument for the Argentinian context, therefore, it is a relevant resource for clinical studies and research in psychotherapy.

References

- Allaire, J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., & Iannone, R. (2022). *rmarkdown: Dynamic documents for R. R package version 2.14*. Recuperado de <https://github.com/rstudio/rmarkdown>
- Barkham, M., Hardy, G. E., & Startup, M. (1996). The IIP-32: A short version of the Inventory of Interpersonal Problems. *British Journal of Clinical Psychology*, 35(1), 21-35. doi: 10.1111/j.2044-8260.1996.tb01159.x
- Berghout, C. C., Zevalkink, J., Katzko, M. W., & de Jong, J. T. (2012). Changes in symptoms and interpersonal problems during the first 2 years of long-term psychoanalytic psychotherapy and psychoanalysis. *Psychology and Psychotherapy: Theory, Research and Practice*, 85(2), 203-219. doi: 10.1111/j.2044-8341.2011.02022.x
- Casullo, M., & Pérez, M. (2008). *El inventario de síntomas SCL-90-R de L. Derogatis*. Universidad de Buenos Aires. Recuperado de <https://www.psi.uba.ar>
- Davidov, E., Billiet, J., Meuleman, B., & Schmidt, P. (2018). *Cross-Cultural Analysis: Methods and Applications (2nd ed.)*. Routledge. doi: 10.4324/9781315537078
- Derogatis, L. (1994). *SCL-90-R*. Minneapolis, MN: NCS.
- Dominguez-Lara, S. (2018). Magnitud del efecto, una guía

- rápida. *Educación Médica*, 19(4), 251-254. doi: [10.1016/j.edumed.2017.07.002](https://doi.org/10.1016/j.edumed.2017.07.002)
- Faustino, B., & Vasco, A. B. (2020). Factor structure and convergent validity of the Portuguese version of the Inventory of Interpersonal Problems-32. *Journal of Relationships Research*, 11. doi: [10.1017/jrr.2020.18](https://doi.org/10.1017/jrr.2020.18)
- Fernández-Álvarez, H., Hirsh, H., Maristany, M., & Torrente, F. (2005). Propiedades psicométricas del OQ-45.2 en la Argentina: Un estudio piloto. Poster presentado en 4° Congreso Mundial de Psicoterapia, Buenos Aires, Argentina.
- Freiberg-Hoffmann, A., Stover, J. B., De la Iglesia, G., & Fernández Liporace, M. (2013). Correlaciones policóricas y tetracóricas en estudios factoriales exploratorios y confirmatorios. *Ciencias Psicológicas*, 7(2), 151-164. doi: [10.22235/cp.v7i1.1057](https://doi.org/10.22235/cp.v7i1.1057)
- Froh, J. J., Fives, C. J., Fuller, J. R., Jacofsky, M. D., Terjesen, M. D., & Yurkewicz, C. (2007). Interpersonal relationships and irrationality as predictors of life satisfaction. *The Journal of Positive Psychology*, 2(1), 29-39. doi: [10.1080/17439760601069051](https://doi.org/10.1080/17439760601069051)
- Gadernann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research, and Evaluation*, 17(1), 3-13. doi: [10.7275/n560-j767](https://doi.org/10.7275/n560-j767)
- Gómez-Penedo, J. M., Areas, M. A., Manubens, R., Babl, A. M., Challú, L., Juan, S., ... Grosse-Holtforth, M. (2021). Propiedades psicométricas del Hopkins Symptom Checklist (HSCL-11) en Argentina: Un instrumento para monitoreo y feedback en psicoterapia. *Revista Evaluar*, 21(2), 33-47. doi: [10.35670/1667-4545.v21.n2.34393](https://doi.org/10.35670/1667-4545.v21.n2.34393)
- Gómez-Penedo, J. M., Babl, A., Krieger, T., Heinonen, E., & Flückiger, C. (2020). Interpersonal agency as predictor of the within-patient alliance effects on depression severity. *Journal of Consulting and Clinical Psychology*, 88(4), 338-349. doi: [10.1037/ccp0000475](https://doi.org/10.1037/ccp0000475)
- Gómez-Penedo, J. M., Constantino, M. J., Coyne, A. E., Westra, H. A., & Antony, M. M. (2017). Markers for context-responsiveness: Client baseline interpersonal problems moderate the efficacy of two psychotherapies for generalized anxiety disorder. *Journal of Consulting and Clinical Psychology*, 85(10), 1000-1011. doi: [10.1037/ccp0000233](https://doi.org/10.1037/ccp0000233)
- Hammer, J. H. (2016). Percent of Uncontaminated Correlations (PUC) Calculator: A Microsoft Excel-based tool to calculate the PUC statistic. Recuperado de <http://drjosephhammer.com>
- Harrell Jr, F. (2021). Hmisc: Harrell miscellaneous. R package version 4.6-0. Recuperado de <https://cran.r-project.org>
- Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: A meta-analytic review. *PLoS Medicine*, 7(7), e1000316. Recuperado de <https://journals.plos.org>
- Horowitz, L. M. (2004). *Interpersonal foundations of psychopathology*. American Psychological Association. Recuperado de <https://www.apa.org>
- Horowitz, L. M., Alden, L. E., Wiggins, J. S., & Pincus, A. L. (2000). *IIP, Inventory of Interpersonal Problems Manual*. San Antonio: The Psychological Corporation. Recuperado de <https://fetzer.org>
- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of Interpersonal Problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology*, 56(6), 885-892. doi: [10.1037//0022-006x.56.6.885](https://doi.org/10.1037//0022-006x.56.6.885)
- Jöreskog, K., & Sörbom, D. (1993). LISREL 8 [Computer software]. Chicago, IL: Scientific Software International, Inc. Recuperado de <https://ssicentral.com>
- Kiliç, A. F., & Doğan, N. (2021). Comparison of confirmatory factor analysis estimation methods on mixed-format data. *International Journal of Assessment Tools in Education*, 8(1), 21-37. Recuperado de <https://files.eric.ed.gov>
- Kline, R. (2005). *Principles and Practice of Structural equation modeling (2nd ed.)*. Guilford Press.
- Lambert, M. J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar,

- S. C. (1996). Reliability and validity of the Outcome Questionnaire. *Clinical Psychology & Psychotherapy: An International Journal of Theory and Practice*, 3(4), 249-258. doi: [10.1002/\(SICI\)1099-0879\(199612\)3:4<249::AID-CPP106>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1099-0879(199612)3:4<249::AID-CPP106>3.0.CO;2-S)
- Leary, T. (1957). *Interpersonal diagnosis of personality; a functional theory and methodology for personality evaluation*. Ronald Press.
- Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936-949. doi: [10.3758/s13428-015-0619-7](https://doi.org/10.3758/s13428-015-0619-7)
- Lieberman, M. D. (2013). *Social: Why our brains are wired to connect*. OUP Oxford.
- Lo Coco, G., Mannino, G., Salerno, L., Oieni, V., Di Fratello, C., Profita, G., & Gullo, S. (2018). The Italian version of the inventory of interpersonal problems (IIP-32): Psychometric properties and factor structure in clinical and non-clinical groups. *Frontiers in Psychology*, 9, 341. doi: [10.3389/fpsyg.2018.00341](https://doi.org/10.3389/fpsyg.2018.00341)
- Luo, X., Nuttall, A. K., Locke, K. D., & Hopwood, C. J. (2018). Dynamic longitudinal relations between binge eating symptoms and severity and style of interpersonal problems. *Journal of Abnormal Psychology*, 127(1), 30-42. doi: [10.1037/abn0000321](https://doi.org/10.1037/abn0000321)
- Lutz, W., Schwartz, B., Gómez-Penedo, J. M., Boyle, K., & Deisenhofer, A. K. (2020). Working towards the development and implementation of precision mental healthcare: An example. *Administration and Policy in Mental Health and Mental Health Services Research*, 47(5), 856-861. doi: [10.1007/s10488-020-01053-y](https://doi.org/10.1007/s10488-020-01053-y)
- Maristany, M. (2005). *Problemas en las relaciones interpersonales y trastornos de la personalidad*. Universidad de Belgrano. Recuperado de <http://repositorio.ub.edu.ar>
- Maristany, M. (2008). Diagnóstico y evaluación de las relaciones interpersonales y sus perturbaciones. *Revista Argentina de Clínica Psicológica*, 17(1), 19-36. Recuperado de <http://www.redalyc.org/articulo.oa?id=281921796002>
- McEvoy, P. M., Burgess, M. M., Page, A. C., Nathan, P., & Fursland, A. (2013). Interpersonal problems across anxiety, depression, and eating disorders: A transdiagnostic examination. *British Journal of Clinical Psychology*, 52(2), 129-147. doi: [10.1111/bjc.12005](https://doi.org/10.1111/bjc.12005)
- Míndrilă, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society*, 1(1), 60-66. Recuperado de <https://infonomics-society.org>
- Newman, M. G., Jacobson, N. C., Erickson, T. M., & Fisher, A. J. (2017). Interpersonal problems predict differential response to cognitive versus behavioral treatment in a randomized controlled trial. *Behavior Therapy*, 48(1), 56-68. doi: [10.1016/j.beth.2016.05.005](https://doi.org/10.1016/j.beth.2016.05.005)
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71-90. doi: [10.1016/j.dr.2016.06.004](https://doi.org/10.1016/j.dr.2016.06.004)
- R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Recuperado de <https://www.r-project.org>
- Rattray, J., & Jones, M. C. (2007). Essential elements of questionnaire design and development. *Journal of Clinical Nursing*, 16(2), 234-243. doi: [10.1111/j.1365-2702.2006.01573.x](https://doi.org/10.1111/j.1365-2702.2006.01573.x)
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667-696. doi: [10.1080/00273171.2012.715555](https://doi.org/10.1080/00273171.2012.715555)
- Revelle, W. R. (2017). psych: Procedures for personality and psychological research. Northwestern University. Recuperado de <https://www.scholars.northwestern.edu>
- Rial-Boubeta, A., Varela-Mallou, J., Abalo-Piñeiro, J., & Levy-Mangin, J. (2006). El análisis factorial confirmatorio. En J. P. Lévy Mangin (Dir.) & J. Varela Mallou (Coord.), *Modelización con estructuras de*

- covarianzas en ciencias sociales: temas esenciales, avanzados y aportaciones especiales* (pp. 119-154). Netbiblo, S. L.
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98(3), 223-237. doi: [10.1080/00223891.2015.1089249](https://doi.org/10.1080/00223891.2015.1089249)
- Ruiz, M. A., Pincus, A. L., Borkovec, T. D., Echemendia, R. J., Castonguay, L. G., & Ragusea, S. A. (2004). Validity of the Inventory of Interpersonal Problems for predicting treatment outcome: An investigation with the Pennsylvania Practice Research Network. *Journal of Personality Assessment*, 83(3), 213-222. doi: [10.1207/s15327752jpa8303_05](https://doi.org/10.1207/s15327752jpa8303_05)
- Salazar, J., Martí, V., Soriano, S., Beltran, M., & Adam, A. (2010). Validity of the Spanish version of the Inventory of Interpersonal Problems and its use for screening personality disorders in clinical practice. *Journal of Personality Disorders*, 24(4), 499-515. doi: [10.1521/pedi.2010.24.4.499](https://doi.org/10.1521/pedi.2010.24.4.499)
- Salzer, S., Pincus, A. L., Winkelbach, C., Leichsenring, F., & Leibing, E. (2011). Interpersonal subtypes and change of interpersonal problems in the treatment of patients with generalized anxiety disorder: A pilot study. *Psychotherapy*, 48(3), 304-310. doi: [10.1037/a0022013](https://doi.org/10.1037/a0022013)
- Schumacker, E., & Lomax, G. (2016). *A Beginner's Guide to Structural Equation Modelling (4th ed.)*. Routledge.
- Segrin, C. (2001). *Interpersonal processes in psychological problems*. Guilford Press.
- Stiles, T., & Hoglend, P. (1994). IIP-64-C - norsk oversettelse [IIP- 64-C - Norwegian translation].
- Sullivan, H. (1955). *The interpersonal Theory of Psychiatry*. London: Tavistock.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55. doi: [10.5116/ijme.4dfb.8dfd](https://doi.org/10.5116/ijme.4dfb.8dfd)
- Thomas, A., Brähler, E., & Strauß, B. (2011). IIP-32: Entwicklung, validierung und normierung einer kurzform des inventars zur erfassung interpersonaler probleme. *Diagnostica*, 57(2), 68-83. doi: [10.1026/0012-1924/a000034](https://doi.org/10.1026/0012-1924/a000034)
- Vanheule, S., Desmet, M., & Rosseel, Y. (2006). The factorial structure of the Dutch translation of the inventory of interpersonal problems: a test of the long and short versions. *Psychological Assessment*, 18(1), 112-117. doi: [10.1037/1040-3590.18.1.112](https://doi.org/10.1037/1040-3590.18.1.112)
- Von Bergen, A., & de la Parra, G. (2002). OQ-45.2, Cuestionario para evaluación de resultados y evolución en psicoterapia: Adaptación, validación e indicaciones para su aplicación e interpretación. *Terapia Psicológica*, 20(2), 161-176. Recuperado de <https://pesquisa.bvsalud.org/portal/resource/pt/lil-389263?lang=es>
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. Routledge.
- Wilson, E. O. (2012). *The social conquest of earth*. WW Norton & Company.
- Zilcha-Mano, S. (2021). Toward personalized psychotherapy: The importance of the trait-like/state-like distinction for understanding therapeutic change. *American Psychologist*, 76(3), 516-528. doi: [10.1037/amp0000629](https://doi.org/10.1037/amp0000629)
- Zilcha-Mano, S., Muran, J. C., Eubanks, C. F., Safran, J. D., & Winston, A. (2018). Not just a non-specific factor: Moderators of the effect of within-and between-clients alliance on outcome in CBT. *Cognitive Therapy and Research*, 42(2), 146-158. doi: [10.1007/s10608-017-9866-5](https://doi.org/10.1007/s10608-017-9866-5)