

TEORÍA DE LA RESPUESTA AL ÍTEM APLICADA A UNA ESCALA DE SOLEDAD: COMPARACIÓN ENTRE MODELOS

Item response theory applied to a loneliness scale:

Comparison between models

Sofía Auné y María Silvia Galibert

sofiaaune177@hotmail.com / sofiaaune@psi.uba.ar

UBA y Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT)¹

Eje temático: Evaluación psicológica y técnicas y fundamentos de la exploración
psicológica.

Resumen

La Escala de Soledad de Buenos Aires (ESBA; Auné y otros, 2019) es un nuevo instrumento construido en la Argentina para la medición psicológica del aislamiento social autopercebido, también denominado soledad. La ESBA está compuesta por siete ítems que se responden mediante una escala Likert de cuatro opciones desde “nada de acuerdo” hasta “totalmente de acuerdo”. Sofía Auné, Facundo Abal y Horacio Attorresi, (2019) estudiaron las propiedades psicométricas del instrumento -como su estructura factorial y su consistencia interna- en el contexto de la teoría clásica de los test. Asimismo, obtuvieron evidencias de validez convergente por medio de la correlación de los puntajes de los participantes en la ESBA con los puntajes en la versión argentina de la Escala de Soledad de Los Angeles, California (UCLA-LS; Russell y otros, 1980; Sacchi & Richaud de Minzi, 1997) y con preguntas de autopercepción de la soledad. Sería útil, además, modelizar la ESBA con la teoría de la respuesta al ítem (TRI), ya que esto permitirá evaluar la calidad psicométrica de la escala mediante otros métodos de análisis más potentes. Habitualmente, la elección del modelo a aplicar de la TRI se ha realizado con base en la preferencia o la habilidad de un equipo de investigación (e.g. Edelen

& Reeve, 2007), cuando es de conocimiento estadístico que ciertos modelos pueden tener más ajuste que otros a datos empíricos específicos. Actualmente existen métodos objetivos para comparar el ajuste relativo entre modelos de la TRI (e.g. De Ayala, 2009; Toland, 2013) que permiten seleccionar el modelo más adecuado para ser aplicado a un conjunto de datos concretos.

Objetivo

Comparar el ajuste a las respuestas a la ESBA de los modelos más utilizados en la actualidad para respuesta politómica ordenada y determinar cuál es el más adecuado para calibrar las respuestas a sus ítems.

Metodología

Se utilizó una muestra por accesibilidad de 427 participantes (55% mujeres), residentes del Área Metropolitana de Buenos Aires, cuya media de edad fue de 45.2 años con un desvío estándar de 11 (mínimo = 18 años, máximo = 73 años). Mediante el software IRTPRO 4.2 se aplicaron a los datos, sucesivamente: el modelo de respuesta graduada (MRG; Samejima, 1969, 2016), el MRG con restricción de igualdad en el parámetro a (*reduced GR model*, MRGR), el modelo generalizado de crédito parcial (MGCP; Muraki, 1992) y el modelo de crédito parcial (MCP; Masters, 1982, 2016). Para evaluar el ajuste relativo, se calcularon el criterio de información de Akaike (*Akaike information criterion*, AIC), el criterio de información bayesiano (*Bayesian information criterion*, BIC), el estadístico M_2 (Maydeu Olivares & Joe, 2005, 2006) y el índice RMSEA asociado. Menores valores de AIC, BIC, estadístico M_2 y RMSEA indican un mejor ajuste (Toland, 2013). Además, se calculó el estadístico R^2_{Δ} (Haberman, 1978), que informa en qué porcentaje el modelo completo (i.e. MRG, MGCP) aumenta la explicación de las respuestas a los ítems con respecto a otro modelo con restricciones (i.e. MRGR y MCP).

Resultados

El MRG fue el modelo con mejor ajuste ($AIC = 5883.57$, $BIC = 5997.16$, $M_2 = 326.29$, $RMSEA = .04$), seguido del MGCP ($AIC = 5922.85$, $BIC = 6036.44$, $M_2 = 413.51$, $RMSEA = .05$). El R^2_{Δ} entre el MRG y el MRGR fue de $= .0145$, señalando por consiguiente un aumento en la explicación de los datos por parte del MRG del 1.45%. De forma consistente, el R^2_{Δ} entre el MGCP y el MCP fue de $= .0122$.

Discusión y conclusiones

Los cuatro modelos ajustaron globalmente las respuestas a la ESBA. Sin embargo, existieron diferencias de grado que permiten concluir que en este caso el MRG es el modelo más adecuado para la calibración de los ítems. Se destaca la importancia de estos procedimientos y la necesidad de su generalización.

Palabras clave: soledad; teoría de la respuesta al ítem; modelo de respuesta graduada.

Notas

1- Esta investigación fue financiada con dos subsidios de la Universidad de Buenos Aires, UBACyT 2018 código 20020170100200BA y UBACyT 2018 código 20020170200001BA y con el de la Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) PICT-2017-3226.

Abstract

The Buenos Aires Loneliness Scale (BALS; Auné *et al.*, 2019) is a new instrument designed in Argentina for the psychological measurement of self-perceived social isolation, also called loneliness. BALS is composed of seven items

that are answered using a 4-option Likert response format from “completely disagree” to “totally agree”. Sofía Auné, Facundo Abal and Horacio Attorresi (2019) studied the psychometric properties, such as its factorial structure and its internal consistency, of the instrument in the context of the classical test theory. Likewise, they obtained evidence of convergent validity through the correlation of the scores of the participants in the BALS with the scores in the Argentine version of the Los Angeles, California, Loneliness Scale (UCLA-LS; Russell *et al.* 1980; Sacchi & Richaud de Minzi, 1997) and with questions of self-perception of loneliness. It would also be useful to model the BALS with the item response theory (IRT), since this will allow the evaluation of the psychometric quality of the scale through other more powerful methods of analysis. Usually, the choice of the IRT model to be applied has been made based on the preference or skill of a research team (e.g. Edelen & Reeve, 2007), when it is statistically known that some models may have more adjustment than others to specific empirical data. Currently, there are objective methods to compare the relative adjustment between models of the IRT (e.g. De Ayala, 2009; Toland, 2013) that allow selecting the most appropriate model to be applied to a set of specific data.

Objective

To compare the adjustment to the BALS responses of the most commonly used models for orderly polytomous response and to determine which one is the most appropriate to calibrate the responses to the items.

Methodology

Data was provided by a convenience sample of 427 participants (55% women), living in the Metropolitan Area of Buenos Aires, whose average age was 45.2 years old with a standard deviation of 11 (minimum = 18 years old, maximum = 73 years old). Through the IRTPRO 4.2 software the following models were

successively applied to the data: the graded response model (GRM; Samejima, 1969, 2016), the GRM with equality restriction in parameter a (reduced GR model, GRMR), the generalized partial credit model (GPCM; Muraki, 1992) and the partial credit model (PCM; Masters, 1982, 2016). To assess the relative adjustment, the Akaike information criterion (AIC), the Bayesian information criterion (BIC), the M_2 statistic (Maydeu Olivares & Joe, 2005, 2006) and the associated RMSEA index were calculated. Lower values of AIC, BIC, M_2 statistic and RMSEA indicate a better adjustment (Toland, 2013). In addition, the R^2_{Δ} statistic was calculated (Haberman, 1978), it reports in what percentage the complete model (i.e. MRG, MGCP) increases the explanation of the responses to the items in relation to another model with restrictions (i.e. MRGR and MCP).

Results

The GRM was the model with the best adjustment (AIC = 5883.57, BIC = 5997.16, M_2 = 326.29, RMSEA = .04), followed by MGCP (AIC = 5922.85, BIC = 6036.44, M_2 = 413.51, RMSEA = .05). The R^2_{Δ} between the GRM and the RGRM was = .0145, therefore indicating an increase in the explanation of the data by the GRM of 1.45%. Consistently, the R^2_{Δ} between the MGCP and the MCP was = .0122.

Discussion and conclusions

The four models adjusted globally to the GRM responses. However, there were differences in grade that allow us to conclude that the GRM is the most suitable model for the calibration of the items in this case. The importance of these procedures and the need for their generalization are highlighted.

Keywords: loneliness; item response theory; graded response model.

Referencias bibliográficas

- Auné, S. E., Abal, F. J. P. & Attorresi, H. F. (2019). "Construction and psychometric properties of the Loneliness Scale in adults". *International Journal of Psychological Research*, 12(2), pp. 82-90. Disponible en <https://doi.org/10.21500/20112084.4257>
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. Nueva York: Guilford.
- Edelen. M. O., & Reeve, B. B. (2007). "Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement". *Qual Life Res.*, 16(1), pp. 5-18. Disponible en <http://dx.doi.org/10.1007/s11136-007-9198-0>
- Haberman, S. J. (1978). *Analysis of qualitative data: Vol. 1: Introductory topics*. Nueva York: Academic Press.
- Masters, G. N. (1982). "A Rasch model for partial credit scoring". *Psychometrika*, 47(2), pp. 149-174.
- Masters, G. N. (2016). "Partial Credit Model". En W. J. van der Linden (ed.). *Handbook of Item Response Theory, Volume 1: Models* (pp. 109-126). Boca Raton: Chapman & Hall/CRC.
- Maydeu Olivares, A. & Joe, H. (2006). "Limited information goodness-of-fit testing in multidimensional contingency tables". *Psychometrika*, 71, pp. 713-732. Disponible en <http://dx.doi.org/10.1007/s11336-005-1295-9>
- Maydeu Olivares, A. & Joe, H. (2005). "Limited and full information estimation and testing in 2n contingency tables: A unified framework". *Journal of the American Statistical Association*, 100, pp.1009-1020 Disponible en <http://dx.doi.org/10.1198/016214504000002069>
- Muraki, E. (1992). "A generalized partial credit model: Application of an EM algorithm". *Applied Psychological Measurement*, 16, pp. 159-176. Disponible en <https://doi.org/10.1002/j.2333-8504.1992.tb01436.x>

- Russell, D., Peplau, L. & Cutrona, C. (1980). "The revised UCLA-LS Loneliness Scale: Concurrent and discriminant validity evidence". *Journal of Personality and Social Psychology*, 39(3), pp. 472-480.
- Sacchi, C. & Richaud de Minzi, M. C. (1997). "La Escala Revisada de Soledad de UCLA-LS: Una adaptación argentina". *Rev. Argent. Clín. Psicol*, 6(1), pp. 43-53.
- Samejima, F. (2016). "Graded response models". En W. J. van der Linden (ed.), *Handbook of Item Response Theory. Volume 1: Models*, pp. 123-136). Boca Ratón: Chapman and Hall/CRC.
- Samejima, F. (1969). "Estimation of latent ability using a response pattern of graded scores". *Psychometrika monograph supplement*, 17(4), p. 2. Disponible en <http://dx.doi.org/10.1002/j.2333-8504.1968.tb00153.x>
- Toland, M. (2013). "Practical guide to conducting an item response theory analysis". *The Journal of Early Adolescence*, 34(1), pp. 120-151. Disponible en <https://doi.org/10.1177/0272431613511332>