Estructura interna y evidencias de validez de la Escala Baptista de Depresión para Adultos en población adulta de Argentina

Makilim Nunes Baptista; Thiago Francisco Pereira Soares; Sebastian Urquijo; Macarena Verónica del Valle

Cómo citar este artículo:
Nunes Baptista, M., Pereira Soares, T., F., Urquijo, S., & del Valle, M. V. (2021). Estructura interna y evidencias de validez de la Escala Baptista de Depresión para Adultos en población adulta de Argentina. *Acta Colombiana de Psicología*, *24*(1), 32-46. https://www.doi.org/10.14718/ACP.2021.24.1.4

Recibido, febrero 05/2019; Concepto de evaluación, mayo 05/2019; Aceptado, octubre 03/2020

Makilim Nunes Baptista[1]
ORCID: https://orcid.org/0000-0001-6519-254X
Universidade São Francisco, São Paulo, Brasil.

Thiago Francisco Pereira Soares
ORCID: https://orcid.org/0000-0001-6681-462X
Universidade São Francisco, São Paulo, Brasil.

Sebastian Urquijo
ORCID: https://orcid.org/0000-0002-8315-9329
Instituto de Psicología Básica, Aplicada y Tecnología (IPSIBAT), Universidad Nacional de Mar del Plata (UNMDP) y Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Mar del Plata, Argentina.

Macarena Verónica del Valle
ORCID: https://orcid.org/0000-0003-3549-7224
Instituto de Psicología Básica, Aplicada y Tecnología (IPSIBAT), Universidad Nacional de Mar del Plata (UNMDP) y Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Mar del Plata, Argentina.

Resumen

El objetivo de este estudio fue analizar la estructura factorial de la adaptación al español rioplatense de la Escala Baptista de Depresión para Adultos (EBADEP-A) en población argentina, y aportar evidencias de validez sobre la misma. Para evaluar la estructura interna de la escala se realizaron cálculos por medio de análisis factoriales exploratorios y confirmatorios, así como de la teoría de respuesta al ítem; mientras que, para brindar evidencias de validez convergente y concurrente, se aplicaron correlaciones, regresiones y análisis de modelos de ecuaciones estructurales (SEM) con medidas relacionadas. Los resultados sugieren que la versión final del instrumento podría reducirse de 45 a 23 ítems con adecuados índices de ajuste, con un modelo unidimensional, y con cargas factoriales satisfactorias tanto en el análisis exploratorio como en el confirmatorio. Los ítems no revelaron discrepancias en las habilidades estimadas de los participantes, y presentaron ajustes estructurales adecuados y opciones de respuestas claras. Respecto a la validez, la escala presentó asociaciones con variables externas —personalidad y bienestar psicológico—; y el análisis SEM demostró una fuerte relación entre el EBADEP-A y el *Beck Depression Inventory-II*, así como una predicción fuerte del primero sobre el segundo (r^2 = .65). Se concluye que el factor general del EBADEP-A representa adecuadamente los fenómenos psicológicos subyacentes a la depresión, y que la versión final de la escala presenta propiedades psicométricas adecuadas para la población argentina.
Palabras clave: psicometría, depresión, evaluación, salud mental.

[1] R. Waldemar César da Silveira, 105, Jardim Cura D'ars, Campinas, SP. Tel.: 08007278855. Correo electrónico: makilim01@gmail.com

Internal structure and evidence of validity of the Baptist Adult Depression Scale in the adult population of Argentina

Abstract

The aim of this study was to analyze the factorial structure of the adaptation to Rio de la Plata´s Spanish of the Baptist Scale of Depression for Adults (EBADEP-A) in the Argentine population, and to provide evidence of its validity. In order to evaluate the internal structure of the scale, calculations were made by means of exploratory and confirmatory factorial analysis, as well as the Item Response Theory. Also, in order to provide evidence of convergent and concurrent validity, correlations, regressions and analysis of structural equation models (SEM) with related measures were applied. The results suggest that the final version of the instrument could be reduced from 45 to 23 items with adequate adjustment indexes, with a one-dimensional model, and with satisfactory factor loads in both the exploratory and confirmatory analyses. The items revealed no discrepancies in the estimated skills of the participants, and presented adequate structural adjustments and clear response options. Regarding the validity, the scale presented associations with external variables - personality and psychological well-being-; and the SEM analysis demonstrated a strong relationship between the EBADEP-A and the *Beck Depression Inventory-II*, as well as a strong prediction of the first over the second (r2 = .65). It is concluded that the general factor of EBADEP-A adequately represents the psychological phenomena underlying depression, and that the final version of the scale presents adequate psychometric properties for the Argentine population.
Keywords: psychometrics, depression, evaluation, mental health.

Introducción

La depresión es actualmente la principal causa de discapacidad en todo el mundo. De hecho, la prevalencia global de la depresión y de los síntomas depresivos ha aumentado en las últimas décadas (Vos et al., 2016), pues la proporción de la población global con depresión en 2015, según la Organización Mundial de la Salud (World Health Organization [WHO], 2017), fue estimada en 4.4 %, y en la Argentina en 4.7 %, siendo más común en mujeres que en hombres. Recientemente, Wang et al. (2017) presentaron una revisión sistemática de la literatura científica sobre el tema, en la que incluyeron 83 estudios transversales con 41 344 pacientes ambulatorios. Como resultados, informaron, por una parte, que la prevalencia de depresión y de síntomas depresivos en dicha población alcanzaba el 27 %, con un mínimo de 17 % y un máximo de 53 % —dependiendo de la especialidad médica analizada—; y, por otra, que los países en desarrollo presentaron mayores índices de depresión que los países desarrollados. Con estos resultados, los autores resaltan la importancia de desarrollar estrategias de evaluación eficaces para la identificación temprana de esta psicopatología.

Ahora bien, la depresión está asociada a la funcionalidad de la persona en su vida diaria, debido a que esta conlleva una disminución de calidad de vida, un aumento de sobrecarga física y cognitiva, y un mayor riesgo de mortalidad (Bromet et al., 2011). Incluso, se ha establecido como un determinante significativo de la calidad de vida y supervivencia, pues representa aproximadamente el 50 % de las consultas psiquiátricas y el 12 % de todas las admisiones hospitalarias (Kuo et al., 2015). No obstante, a pesar de su importancia, se estima que alrededor del 50 % de los pacientes deprimidos no son identificados correctamente en las evaluaciones y diagnósticos rutinarios, y que solo un porcentaje limitado de los casos recibe el tratamiento adecuado (Demyttenaere et al., 2004; Vilagut et al., 2016). Respecto a esto último, la Organización Mundial de la Salud (OMS, 2018) informó que, a pesar de que existen tratamientos eficaces para la depresión, más de la mitad de los afectados en todo el mundo —y más del 90 % en muchos países— no recibe esos tratamientos, debido, en parte, a la evaluación errónea y a la falta de diagnósticos correctos.

Considerando esta información, la evaluación adecuada y sistemática de la depresión supone un importante medio para mejorar la detección temprana, el tratamiento y el seguimiento del progreso de los pacientes (Jeon et al., 2017; Pignone et al., 2002). Además, analizar la sintomatología depresiva también resulta importante, tanto en contextos de atención primaria como en la población en general, para monitorear la prevalencia de enfermedades mentales (Williams et al., 2002) y para desarrollar e implementar intervenciones. Por lo tanto, resulta clara la necesidad creciente de contar con instrumentos de evaluación —breves y fiables— que permitan discriminar a las personas con depresión de aquellas que no presentan síntomas depresivos (Vilagut et al., 2016).

Los diversos instrumentos que evalúan y rastrean la depresión poseen una amplia gama de características definitorias muy diferentes entre sí, pues presentan variaciones respecto a su teoría de referencia, su dimensionalidad, su forma de medida —escalas de respuesta tipo Likert vs.

escalas dicotómicas—, el número de ítems, la especificidad del ciclo de vida, o el formato de aplicación, entre otras (Baptista, 2018a; Baptista et al., 2012).

Respecto al contenido de los ítems de las diferentes medidas de depresión, Calil y Pires (1998) realizaron comparaciones entre distintas escalas —la Escala de Hamilton, la Escala de Evaluación de la Depresión de Montgomery-Asberg, la Escala de Evaluación de la Melancolía de Bech-Rafaelsen, el Inventario de Depresión de Beck, la Escala de Autoevaluación de la Depresión de Zung, el Inventario de Autoevaluación de Wakefield y la Escala de Evaluación de Depresión de Carrol— y hallaron que de un 8 a un 30 % de los ítems evalúa la categoría humor, un 18 a un 35 % indaga sobre síntomas vegetativos, entre un 0 y un 18 % evalúa síntomas motores, entre un 0 y un 9 % explora síntomas sociales, entre un 0 y un 52 % la mayor variación encontrada— indaga síntomas cognitivos, de un 0 a un 17 % evalúa síntomas de ansiedad, y, por último, entre un 0 y un 8 % mide síntomas de irritabilidad.

Por otra parte, Baptista y Borges (2016) realizaron una revisión respecto los instrumentos de depresión —en niños/adolescentes y adultos— utilizados entre 2005 y 2015 en Brasil, y encontraron que, de 24 instrumentos, solo 12 de ellos eran específicos para evaluar depresión. Entre los más citados, se encontraban el Inventario de Depresión de Beck (BDI-II), la Hospital Anxiety and Depression Scale, la Hamilton Depression Rating Scale, el Children Depression Inventory, y la Escala Baptista de Depresión Versión Adulto (EBADEP-A).

En particular, la EBADEP-A es una escala de autoinforme desarrollada en Brasil que tiene por objetivo evaluar la intensidad de depresión en adolescentes y adultos (Baptista, 2012), que ha sido utilizada en varios estudios, y que posee diversas versiones (p. ej., Cremasco & Baptista, 2018) elaboradas teniendo como base el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM), la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud (CIE), y la Teoría Cognitiva de la Depresión (Baptista, 2012).

Respecto a su estructura, la escala cuenta con 45 ítems —divididos en dos frases por ítem— que el sujeto debe responder en una escala con formato de respuesta tipo Likert de cuatro opciones. Específicamente, el 28.3 % de los ítems responde a aspectos sociales —incapacidad, disminución de la productividad, dependencia, evitación social, etc.—; el 24.5 %, al humor —humor deprimido, llanto, anhedonia, etc.—; el 22.6 %, a la dimensión cognitiva —desamparo, indecisión, desesperanza, autocrítica exacerbada, disminución de la concentración, ideación suicida y culpa—; el 15.1 %, a aspectos vegetativo-somáticos —fatiga, alteración del sueño, del apetito, del peso y pérdida de la libido—; el 3.8 %, a aspectos motores —agitación, lentitud e irritabilidad—; el 3.8 %, a irritabilidad; y el 1.9 %, a ansiedad —hipocondría—. No obstante, independientemente de esta distribución, la EBADEP-A presenta una estructura unidimensional, con buenos indicadores de ajuste *infit* y *outfit* para el modelo, y muy buen nivel de confiabilidad (α = .95; Baptista, 2012).

Son diversos los estudios de validez y confiabilidad realizados sobre la EBADEP-A que aportan evidencia sobre su utilidad en la evaluación de la depresión. Por ejemplo, Baptista et al. (2017) realizaron una calibración de la escala y evaluaron los patrones de transferencias entre esta y el BDI-II; Bighetti et al. (2014) presentaron evidencias de validez basadas en la relación entre la EBADEP-A y el Big Five; Gomes y Baptista (2014) presentaron las puntuaciones de corte del EBADEP-A por medio de procedimientos de transferencia de normas con base en el Center for Epidemiologic Studies-Depression Scale; Baptista et al. (2013) analizaron la capacidad diagnóstica de la EBADEP-A; y Baptista et al. (2012) presentaron evidencias de validez convergente entre la EBADEP-A y el BDI-II, y evaluaron la estabilidad temporal de la escala por medio de un análisis test-retest luego de un período de un mes.

Similarmente, también existen estudios sobre otras de las versiones de la escala que dan cuenta de la utilidad, validez y confiabilidad para ser aplicada en distintas poblaciones. Por ejemplo, Baptista y Cremasco (2013), en un estudio de validez de la Escala Baptista de Depresión Versión Infanto-Juvenil, analizaron las relaciones con variables externas, y encontraron correlaciones negativas —entre moderadas y altas— con el Inventario de Percepción de Soporte Familiar y la Escala de Percepción de Soporte Social; mientras que, con el objeto de buscar evidencias de validez de la escala EBADEP en su versión para adultos mayores, Coutinho et al. (2016) realizaron un estudio en el que encontraron correlaciones positivas fuertes con el resultado en la Escala de Depresión Geriátrica, y negativas con el Mini Examen de Estado Mental.

En resumen, el EBADEP-A es una escala que ha sido utilizada en diferentes contextos y posee variada evidencia sobre su validez, lo que da cuenta de su eficacia para la evaluación de la depresión (Baptista & Gomes, 2011; Baptista et al., 2012; Bighetti et al., 2014). No obstante, aún no se han realizado evaluaciones sobre su funcionamiento en población hispanoparlante.

Teniendo esto en cuenta, el presente estudio tuvo como objetivo, por una parte, analizar la estructura interna del EBADEP-A en población argentina, y, por otra, buscar evidencias de validez con base en algunas variables externas. Para esto, se analizaron los resultados obtenidos con una adaptación del instrumento por medio del análisis de la teoría de respuesta al ítem (TRI), de los modelos de ecuaciones

estructurales (SEM), y de un análisis de relaciones con variables externas con el fin de reducir la escala para una versión argentina.

Método

Tipo de estudio

Se realizó un estudio de tipo instrumental según la clasificación de Montero y León (2007).

Participantes

Se empleó un muestreo no probabilístico y por conveniencia, con el que se obtuvo una muestra compuesta por 854 participantes argentinos, con edades comprendidas entre los 17 y los 66 años (*ME* = 23.81; *DE* = 6.39). La mayoría de los participantes era de género femenino (*n* = 591; 69.2 %) y, en cuanto al estado civil, 762 (89.2 %) eran solteros.

Para llevar a cabo los distintos análisis, el total de la muestra se dividió aleatoriamente en dos grupos mediante el comando de selección aleatoria del *software* SPSS (versión 23). El primero, compuesto por 433 casos, se utilizó para el análisis de la estructura interna y de validez con variables externas; y el segundo, de 421 participantes, se utilizó para realizar el análisis factorial confirmatorio (AFC). En la Tabla 1 se presentan los datos sociodemográficos de la muestra.

Finalmente, es importante mencionar que todos los participantes eran estudiantes universitarios, que la muestra total se utilizó para realizar las correlaciones y los análisis de regresión, y que las franjas etarias se dividieron según los criterios establecidos por la Organización Mundial de la Salud (WHO, 2014).

Tabla 1.

Datos sociodemográficos

Datos sociodemográficos	f	%
Franja etaria		
Jóvenes (17-24 años)	623	73
Jóvenes adultos (25-44 años)	209	24.4
Adultos (45-59 años)	20	2.2
Adultos mayores (60-66 años)	2	0.4
Género		
Masculino	23	30.8
Femenino	591	69.2
Estado civil		
Soltero	762	89.2
Casado	22	2.6
Viudo	9	1.1
Separado/Divorciado	1	0.1
Segundo matrimonio	29	3.4
No responde	31	3.7

Instrumentos

Adaptación al español rioplatense de la Escala Baptista de Depresión (versión Adultos)

La EBADEP-A (Baptista, 2012) es un instrumento autoadministrado que evalúa sintomatología depresiva a través de 45 ítems —divididos en dos frases por ítem— que el sujeto debe responder en una escala con formato de respuesta tipo Likert de cuatro opciones. Las respuestas se puntúan de 0 a 3, con un puntaje total de mínimo 0 y máximo 135, en donde, a menor puntuación total, menor nivel de sintomatología. La escala puede ser aplicada en una amplia franja etaria, que varía desde los 17 a los 81 años.

Para la adaptación, se realizó una traducción no literal del portugués al español rioplatense, basada en la adaptación de significados. El traductor inicial fue un investigador, psicólogo, hispanoparlante nativo, con certificaciones universitarias de dominio del portugués brasileño coloquial y científico. Se buscó que el producto final del proceso de adaptación consiguiera, con respecto a la prueba original, el máximo nivel de equivalencia lingüística, cultural, conceptual y métrica posible. Para ello, se siguieron las directrices del International Test Commission (ITC) (Muñiz et al., 2013).

Luego, se solicitó a tres investigadores —psicólogos, dos brasileños nativos y un hispanoparlante nativo residente en Brasil— con conocimiento específico de la depresión, una retrotraducción al portugués brasilero, con lo que se obtuvo una equivalencia del 95 %. La traducción inicial y las retrotraducciones fueron revisadas por un grupo de investigadores argentinos y brasileños, con lo que se pudo realizar la corrección lingüística y la adecuación práctica de los términos definitivos.

A continuación, se realizó una prueba piloto, con una pequeña muestra (*n* = 15) de pacientes psicológicos, docentes y estudiantes universitarios, amas de casa, y empleados administrativos, para analizar, estudiar y corregir aspectos relacionados con la adaptación. Con esto, se observaron las reacciones de las personas que realizaron la prueba, y se comprobó que los ítems y las instrucciones fueran correctamente comprendidos. Se corrigieron algunas expresiones idiomáticas, como "Me resulta difícil tomar decisiones" vs. "Me resulta difícil decidir", o "Por las mañanas me siento agotado" vs. "A la mañana ya me siento agotado"; se registró el tiempo promedio necesario para la ejecución del cuestionario —entre 3 y 5 minutos—; se recogió información sobre errores de contenido y de formato; y se modificó el diseño antes de pasar a la fase de colecta de datos definitiva.

Como se mencionó, en estudios previos la escala ha presentado buenas evidencias de validez y confiabilidad

(Baptista & Gomes, 2011; Baptista et al., 2012; Urquijo et al., 2017); y, en esta misma línea, en este estudio el índice de confiabilidad (alfa de Cronbach) para la escala completa fue de .944.

Inventario de Depresión de Beck

Se administró la adaptación al español (Sanz et al., 2005; Sanz & Vázquez, 2011) del Inventario de Depresión de Beck-II (BDI-II) (Beck et al., 1996), una escala de autoinforme compuesta por 21 reactivos indicativos de síntomas depresivos tales como tristeza, llanto, pérdida de placer o culpa, que se puede aplicar de forma individual o colectiva, y tanto en contextos clínicos como no clínicos. Cada reactivo está compuesto por cuatro frases indicativas de distintos niveles de intensidad de cada síntoma, y las respuestas se codifican de 0 a 3, por lo que la puntuación total, unidimensional, oscila entre los 0 y los 63 puntos.

Este instrumento es uno de los más ampliamente difundidos para evaluar la sintomatología depresiva y la presencia de trastornos depresivos; y cuenta con adecuados índices de confiabilidad (α = .89, Sanz et al., 2005) y validez (p. ej., Beltrán BDI-II, 2012; Sanz & Vázquez, 1998). En este estudio, el alfa de Cronbach para la escala completa fue de .863.

Escala de Bienestar Psicológico

Se utilizó la adaptación al español (Díaz et al., 2006) de la Escala de Bienestar Psicológico de Ryff (EBP; Ryff, 1989), la cual se compone de 29 ítems con un formato de respuesta tipo Likert de seis opciones de respuesta, que van de 1 ("Totalmente en desacuerdo") a 6 ("Totalmente de acuerdo"). Los reactivos se dividen en las seis dimensiones del bienestar psicológico propuestas por Ryff —autoaceptación, relacionamiento positivo, autonomía, dominio del entorno, satisfacción vital, y crecimiento personal—, y el instrumento es de autoinforme, breve y de fácil aplicación. La versión adaptada obtuvo adecuados índices de ajuste (Díaz et al., 2006) y ya ha sido utilizada en la población bajo estudio (Del Valle et al., 2015), en la cual se replicó la estructura de seis factores original. El coeficiente alfa de Cronbach en este estudio fue de .753.

Listado de Adjetivos para Evaluar la Personalidad

El Listado de Adjetivos para Evaluar la Personalidad (AEP; Ledesma et al., 2011; Sánchez & Ledesma, 2013) es un instrumento desarrollado en Argentina a partir del modelo Big Five (McCrae & Costa, 2008). En lugar de utilizar un formato tradicional de frases, el AEP está compuesto por un listado de 67 adjetivos que representan los cinco rasgos de personalidad: extraversión, apertura a la experiencia, responsabilidad, neuroticismo y amabilidad.

En esta escala, los participantes contestan a cada reactivo en función de cuánto los describe ese adjetivo, en una escala de tipo Likert que va de 1 ("No me describe en absoluto") a 5 ("Me describe tal como soy"). Se trata de un instrumento autoadministrable, que requiere de aproximadamente diez minutos para ser completado, y que ha presentado propiedades psicométricas adecuadas (Ledesma et al., 2011). En nuestro estudio, el alfa de Cronbach para la escala completa fue de .771.

Procedimiento

La investigación se encuentra enmarcada dentro de un proyecto general aprobado por la Secretaría de Investigación y Posgrado de la Universidad Nacional de Mar del Plata (15/H246). También, se enmarca en un proyecto general de la Universidade São Francisco-SP, que fue aprobado por el Comité de Ética de la Universidade São Francisco-SP (N. º 2.522.435). Los datos fueron recolectados en la Universidad Nacional de Mar del Plata, y los instrumentos fueron administrados en aulas de forma colectiva a pequeños grupos, bajo la supervisión de un investigador formado. Cada sujeto fue debidamente informado de los objetivos y características del estudio, así como de la confidencialidad de los datos recabados; tras esto, se solicitó la firma de un consentimiento informado. Finalmente, se aseguró que se respetaran las leyes y normas locales, regionales, nacionales e internacionales para estudios con sujetos humanos.

Análisis de los datos

En primer lugar, se analizó la aplicabilidad del análisis factorial exploratorio (AFE) mediante la prueba KMO y el índice de esfericidad de Bartlett. Posteriormente, se procedió a realizar el AFE con la primera parte de la muestra (n = 433), mediante un análisis paralelo (Horn, 1965) y el método de mínimos cuadrados no ponderados, robusto frente a variables ordinales (ULS; Lloret-Segura et al., 2014).

La rotación posteriormente aplicada fue la Varimax, ante la existencia de un único factor. Se retuvieron los reactivos con cargas factoriales iguales o superiores a .60 con el fin de disminuir el número de ítems de la escala y mantener solo aquellos que presentaran mayor correlación con el factor general (Sousa et al., 2015). A continuación, se evaluó la comunalidad de los ítems teniendo en cuenta que esta representa la proporción de varianza para cada variable incluida en el análisis —que es explicada por el factor extraído— (Figueiredo-Filho & Silva Júnior, 2010; Hair et al., 2007; Primi, 2012). Estos análisis se realizaron con el *software* MPlus 7.0.

Posteriormente, se hizo un análisis de los ítems mediante el modelo de Rasch (Rasch, 1960), un modelo de TRI que proporciona una escala intervalar que permite interpretar las distancias, en términos de medidas, de los rasgos latentes de las respuestas de los diferentes sujetos. La unidimensionalidad de los ítems se evaluó a través de la varianza explicada, —siendo aceptable una varianza explicada > 30 %— (Vélez et al., 2016). Además, como indicadores de ajuste, se emplearon los índices *infit* y *outfit*, siguiendo los criterios de Linacre (2011).

También, se evaluó la curva de información de la prueba generada a partir del modelo de Rasch, pues esta ilustra cuál región del rasgo latente es evaluada con mayor precisión por los ítems de la escala, lo que constituye un análisis en profundidad de la precisión de los instrumentos para cada nivel específico del rasgo latente (Embretson & Reise, 2013). Asimismo, se evaluaron las categorías de respuesta del factor general del EBADEP-A (Abraham & Chapelle, 1992).

Para estos análisis se utilizó el programa Winsteps 4.0.1 y se siguieron los criterios de Linacre (2011), quien establece un intervalo de 0.50 a 1.50 para considerar un buen ajuste de la dificultad de los ítems.

Tras estos análisis, se confeccionaron dos mapas de ítems: el primero corresponde a los ítems del EBADEP-A; y el segundo incluye el conjunto de los ítems del EBADEP-A y del BDI-II en el programa Winsteps 4.0.1. Específicamente, el símbolo "‡" en el mapa representa seis sujetos, la letra "M" —localizada tanto en el lado de los ítems como en el de las personas— representa la media, "S" significa desvío estándar, "T" representa dos desviaciones estándar, y el punto "." representa un número entre 1 y 5 personas.

El funcionamiento diferencial de los ítems (DIF) en función del género de los participantes se evaluó con el método de Múltiples Causas y Múltiples Indicadores (MIMIC), por medio del MPlus 7.0. Para esto, se utilizó como parámetro un $p < .05$ en cada ítem. En particular, la propuesta del análisis DIF es que las respuestas a los ítems de un instrumento reflejen la dimensión subyacente —el rasgo latente— que el mismo pretende medir, y supone que un cuestionario es aceptable cuando los ítems correspondientes no favorecen la adhesión al ítem en ninguno de los dos géneros (Michaelides, 2008).

Con la segunda parte de la muestra ($n = 421$) se realizó un análisis factorial confirmatorio (AFC), el cual tuvo como hipótesis la estructura de base encontrada inicialmente en el AFE (MacCallum et al., 1996). El uso del AFC para examinar la validez factorial e identificar niveles aceptables de adecuación de un modelo es una técnica confiable y comúnmente utilizada en psicología (Hopwood & Donnellan, 2010).

Los ítems encontrados en el AFE se fijaron en el factor; y el método de estimación utilizado fue el ULS, siguiendo los criterios ya presentados en el AFE. Los índices de ajuste considerados fueron el χ^2, el RMSEA, el CFI, el TLI, y el SRMR.

Luego del AFC, se realizó una evaluación de la validez mediante la asociación del conjunto de los ítems con variables externas (EBP y AEP) mediante correlaciones de Pearson y una regresión lineal múltiple —método *enter*—, tomando al EBADEP-A como variable criterio y a los demás instrumentos como variables predictoras (VIF < 1.62).

Por último, como una medida de validez concurrente (Chatterji & Levine, 2008) y utilizando a todos los sujetos de la muestra ($n = 853$), se realizó un análisis SEM, fijando los ítems del EBADEP-A, y explicando este último al BDI-II con sus ítems fijados. La intención de esto fue conseguir un coeficiente beta superior o igual a .60, y que el EBADEP-A explicara un porcentaje significativo del BDI-II (Hair et al., 2014).

Resultados

El índice de esfericidad de Bartlett (16231.5; $p < .01$) y el estadístico KMO (.94) sugirieron la aplicabilidad del AFE, de modo que se realizaron sucesivos AFE para llegar finalmente a un total de 23 ítems. En la Tabla 2 se presentan los resultados finales de este procedimiento.

Tabla 2.

Análisis factorial exploratorio

Test Chi-cuadrado de ajuste del modelo = 1393.978
Grados de libertad = 230
RMSEA = .05
CFI = .91
TLI = .90
SRMR = .08

Ítems originales	Carga factorial	h^2
2	.61	.37
3	.69	.48
4	.69	.48
5	.62	.38
8	.62	.38
11	.69	.48
13	.82	.67
18	.80	.64
19	.80	.64
23	.72	.52
24	.78	.61
25	.77	.59
27	.78	.61
28	.82	.67
30	.83	.69
32	.65	.42
33	.81	.66
34	.65	.42
35	.64	.41
36	.78	.61
37	.65	.42
38	.78	.61
39	.84	.70

Después de retener, mediante el AFE, solo los ítems con cargas ≥ .60, la escala quedó constituida por 23 ítems con cargas factoriales entre .61 y .86, y con buenas comunalidades —que varían de .37 a .74—. La varianza explicada total fue de 40.28 %. Después, se realizó el análisis del modelo logístico de un parámetro —modelo de Rasch— del EBADEP-A con el conjunto de los ítems encontrados en el AFE. La Tabla 3 muestra el análisis de los ítems de la escala.

Tabla 3.
Análisis de teoría de respuesta al ítem de los ítems del EBADEP-A

Ítem de la escala original	*Infit*	*Outfit*
2	1.08	1.26
3	1.06	0.91
4	1.45	1.43
5	0.99	1.07
8	1.00	0.92
11	1.09	0.98
13	1.29	1.15
18	0.67	0.61
19	0.72	0.82
23	0.80	0.76
24	1.07	0.92
25	0.97	0.90
27	1.14	1.02
28	0.74	1.09
30	0.93	0.98
32	1.06	0.93
33	0.86	1.30
34	0.68	0.72
35	1.00	1.03
36	0.89	0.78
37	1.30	1.37
38	1.16	1.09
39	0.61	0.54

Como se observa en la Tabla 3, ningún ítem presentó índices de *infit* u *outfit* por fuera de los límites críticos verificados, lo que sugiere que las opciones de respuesta de la escala Likert están dentro de los patrones esperados, y que consideran la habilidad estimada de la muestra argentina.

Por otra parte, se verificó la curva de información del test por medio de TRI. En la Figura 1 se presenta la representación gráfica de la curva para el intervalo de –17.0 a 17.0 *logits* del único factor del instrumento, así como la región del rasgo latente que es evaluada con mayor precisión por los ítems de la escala. En general, se encontró una tendencia de los ítems del EBADEP-A a ser más informativos y presentar mayor nivel de adhesión a la porción media del rasgo latente de la muestra, lo que indica un buen funcionamiento de los ítems para evaluar el rasgo latente en cuestión.

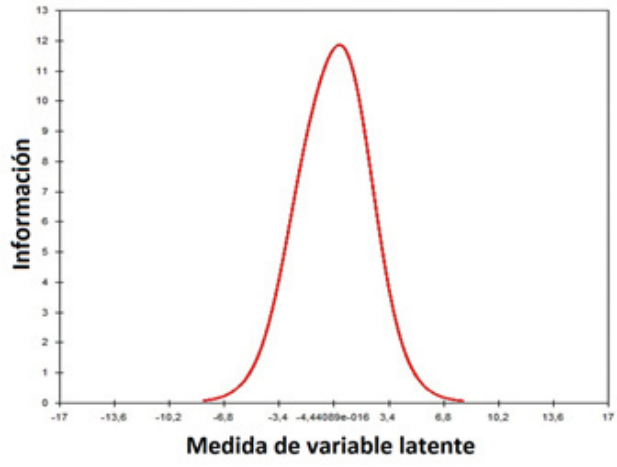

Figura 1.
Curva de información (CI) del EBADEP-A

De igual forma, en la Figura 2 se observa una representación exhaustiva de la adhesión de respuesta de los participantes para cada opción de la escala Likert de 0 a 3 puntos, en la que se encuentra que la mayor parte de los ítems presentó una adhesión a las opciones de repuesta 1 y 2.

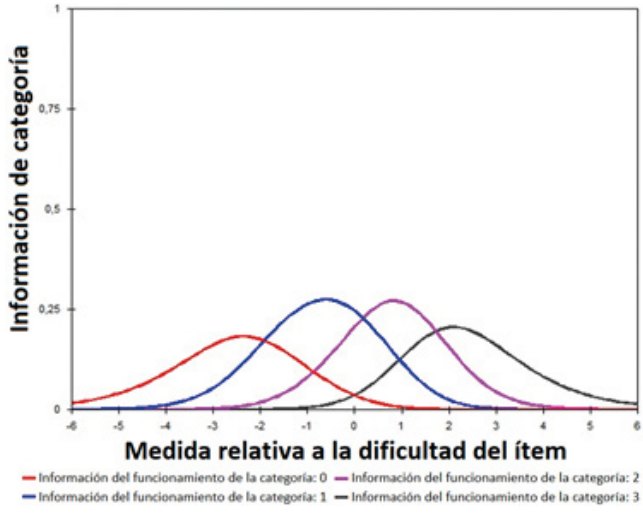

Figura 2.
Categorías de respuestas EBADEP-A

Posteriormente, como se puede observar en el mapa de ítems del desempeño en la escala EBADEP-A que aparece en la Figura 3, también se analizó la distribución de los ítems en función de las habilidades de la muestra argentina. En general, teniendo en cuenta el parámetro de dificultad encontrado por medio del modelo de Rasch, los ítems con mayores índices de dificultad fueron el 28, "Mi vida es buena/Mi vida está cada vez peor", el 30, "Creo en mí/No creo más en mí", y el 38, "Me siento con disposición/Ando más cansado"; mientras que el ítem con mayor facilidad

fue el 37, "No me siento nervioso/Cualquier cosa me deja nervioso", aunque, de manera general, se observa una fácil adhesión a los ítems por la mayor parte de la muestra.

Asimismo, se observan ítems de contenidos y descriptores diferentes en el mismo nivel de dificultad, y que la media de la muestra está muy por debajo de la media de los ítems; hecho esperado, dado que la muestra del estudio no fue clínica, y congruente con lo representado previamente en la Figura 1, donde se mostró que existe una mayor adhesión a la porción media del rasgo latente de la muestra.

Ahora bien, en el mapa de ítems y personas representado en la Figura 4 se muestra la relación del nivel de habilidad y dificultad encontrada para los dos test escogidos, el EBADEP-A y el BDI-II. En general, se puede observar que la media de los ítems fue mayor que la media de las personas, lo que indica que los test fueron considerados como difíciles para la muestra. Nuevamente, y similar a lo mencionado para la Figura 3, dado que se trata de escalas con finalidad diagnóstica que han sido aplicadas a población adulta no clínica, es esperable que el nivel de dificultad de los ítems fuera mayor a la media de la muestra. Además, por tratarse de instrumentos de escala tipo Likert, se puede decir que los participantes tienden a responder con mayor frecuencia a las opciones cuyo sentido indica poca sintomatología depresiva. También, si se considera la dificultad de los ítems de ambos instrumentos, se puede observar un balanceo en la distribución.

A continuación, se realizó un análisis DIF según el género de los participantes (véase Tabla 4), utilizando las comparaciones de las medias entre los grupos masculino y femenino, y verificando si el residuo resultante de la

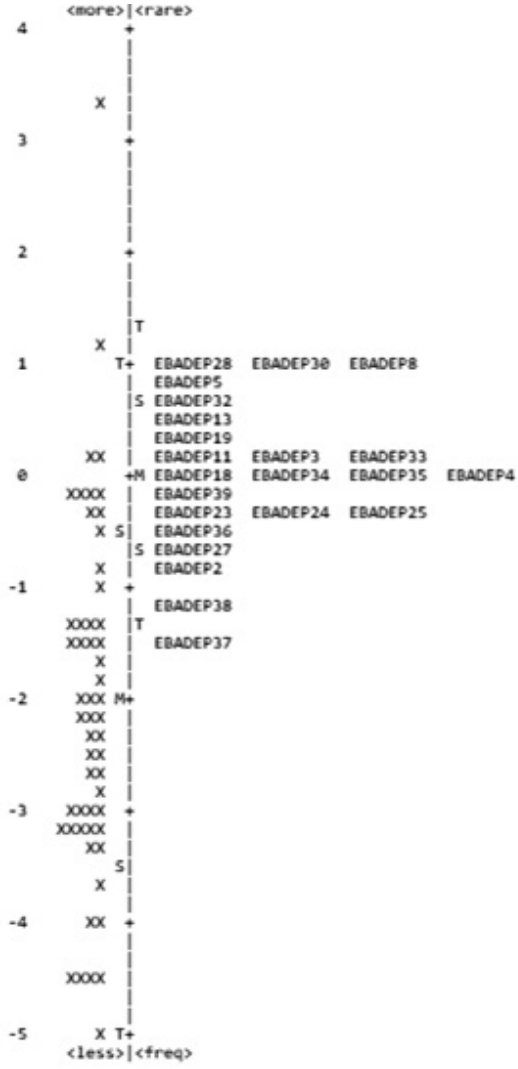

Figura 3.
Mapa de ítems EBADEP-A

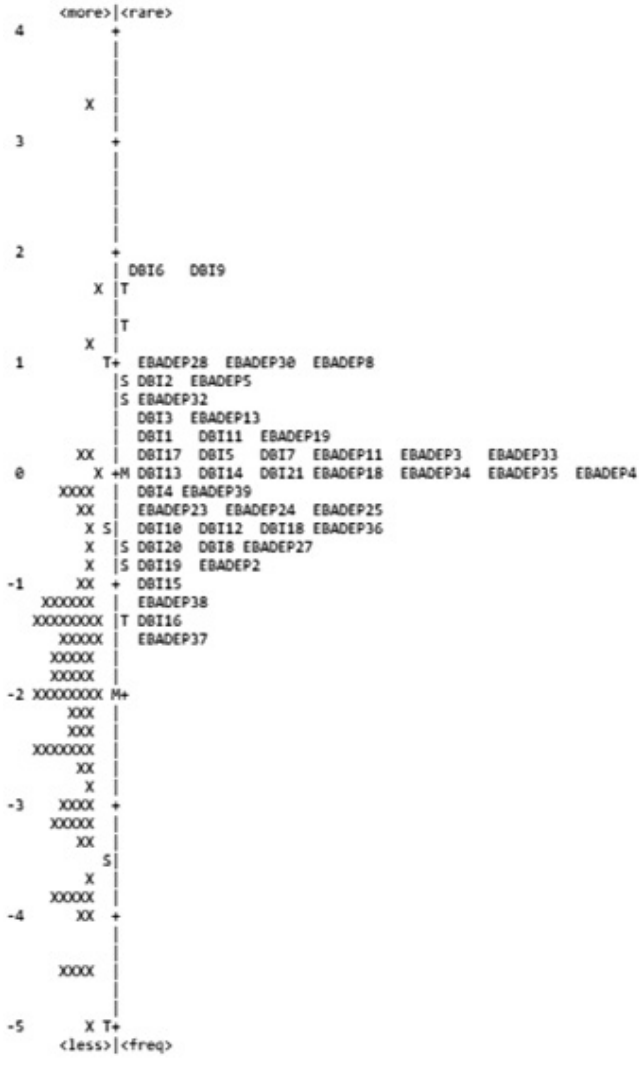

Figura 4.
Mapa de ítems EBADEP-A y BDI-II

comparación entre las respuestas podría ser atribuido al azar. Como se puede observar en la Tabla 4, solo dos ítems presentaron DIF —aunque con magnitudes pequeñas—: el 6, "No he llorado/He llorado mucho", y el 19 "Creo que las cosas van bien en mi vida/Creo que actualmente nada va bien en mi vida"; lo que demuestra que la escala, de modo general, no favorece a ninguno de los dos géneros.

Tabla 4.

DIF de los ítems y valores de comparación de medias de los grupos por género en MIMIC

Ítem	DIF	*p*	N. º del ítem de la escala original
1	.120	.260	2
2	–.137	.397	3
3	–.163	.297	4
4	–.239	.117	5
5	–.234	.118	8
6	–.381	.023	11
7	–.180	.292	13
8	–.145	.357	18
9	–.186	.223	19
10	–.321	.059	23
11	–.225	.235	24
12	–.036	.669	25
13	.060	.716	27
14	–.200	.242	28
15	–.252	.166	30
16	–.165	.309	32
17	–.220	.182	33
18	–.181	.233	34
19	–.391	.016	35
20	.044	.784	36
21	.177	.250	37
22	.152	.385	38
23	.030	.853	39

Luego, se realizó el AFC, cuyos resultados se presentan en la Tabla 5. Como se puede observar, el modelo testeado mostró buen ajuste a los datos, pues las cargas factoriales variaron de .60 a .82, y la comunalidad fue de .36 a .67.

A continuación, se analizó la relación con variables externas, y los resultados mostraron correlaciones positivas con el BDI-II (r = .720) y con la dimensión neuroticismo del AEP (r = .546). Todas las demás correlaciones fueron negativas: EBP - Relacionamiento positivo (r = –.529), EBP - Autonomía (r = –.422), AEP - Extraversión (r = –.338), AEP - Amabilidad (r = –.220), AEP - Responsabilidad (r = –.320), y AEP - Apertura a la experiencia (r = –.102). Estas correlaciones fueron significativas (p < .01), y los valores oscilaron entre bajos y medios.

Posteriormente, se realizó el análisis de regresión lineal múltiple, cuyos resultados se muestran en la Tabla 6. Como se

Tabla 5.

AFC del EBADEP-A en muestra argentina

Test Chi-cuadrado de ajuste del modelo = 1300.043		
Grados de libertad = 230		
RMSEA = .05		
CFI = .91		
TLI = .90		
SRMR = .08		
Ítems	Carga factorial	h^2
2	.68	.46
3	.60	.36
4	.68	.46
5	.61	.37
8	.77	.59
11	.60	.36
13	.82	.67
18	.81	.66
19	.60	.36
23	.70	.49
24	.77	.59
25	.71	.50
27	.80	.64
28	.80	.64
30	.76	.53
32	.60	.36
33	.82	.67
34	.76	.58
35	.70	.49
36	.72	.52
37	.61	.37
38	.74	.55
39	.82	.67

puede observar, según el análisis de los Beta, las correlaciones con EBP - Relacionamiento positivo, AEP - Responsabilidad y AEP - Neuroticismo se mantuvieron, y la mayor contribución al modelo fue la de la dimensión AEP - Neuroticismo (r = .289). En general, el modelo explicó el 44 % de los síntomas depresivos expresados por la muestra (EBADEP-A).

Por último, se realizó un modelo de ecuaciones estructurales (SEM) con el fin de determinar la relación entre las variables latentes del EBADEP-A y el BDI-II (véase Figura 5). En general, el modelo presentó adecuados índices de ajustes ($\chi 2$/df = 2866.35/901; RMSEA = .05; CFI = .93; TLI = .91; SRMR = .06), y las cargas factoriales del EBADEP-A variaron entre los .42 y .74 puntos, mientras que las del BDI-II lo hicieron de .24 a .64 puntos. Además, es importante destacar que dos ítems de este último instrumento presentaron cargas por debajo de .30; que el beta de regresión entre las variables latentes fue de .81; y que el EBADEP-A fue capaz de predecir un 65 % del BDI-II, siendo este un porcentaje considerable.

Tabla 6.
Regresión lineal múltiple del EBADEP-A *y demás instrumentos como variables independientes*

	Error estándar	β estandarizados	*t*	*p*
EBP - Relacionamiento positivo	.668	–.248	–4.01	.000
EBP - Autonomía	.769	–.062	–0.99	.323
AEP - Extraversión	.097	–.084	–1.34	.179
AEP - Amabilidad	.092	–.154	–0.95	.343
AEP - Responsabilidad	.085	–.211	–3.63	.000
AEP - Neuroticismo	.068	.289	4.67	.000
AEP - Apertura a la experiencia	.103	–.044	–0.79	.430

Nota. r^2 = .453; r^2 corregida = .436

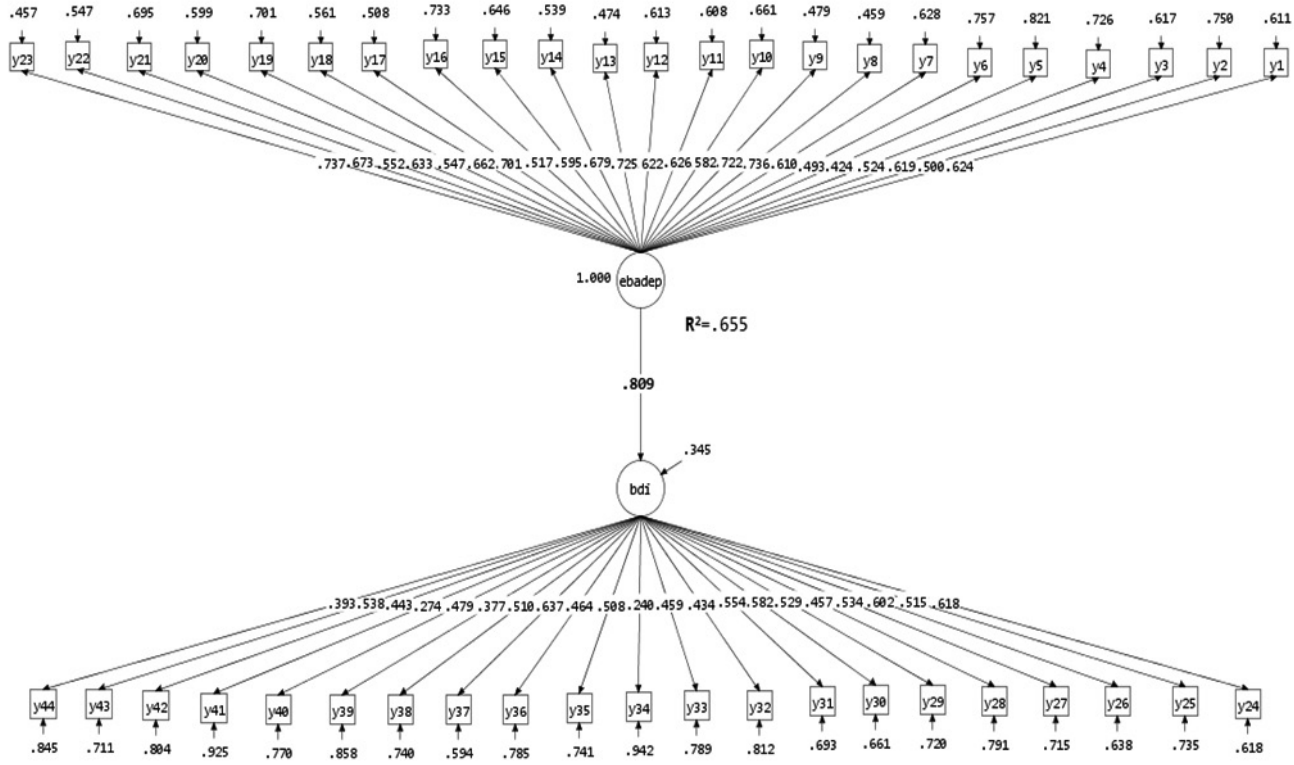

Figura 5.
SEM variables latentes

Discusión

Una vez realizados los análisis pertinentes, para la versión final del instrumento se retuvo un total de 23 ítems de los 45 iniciales de la prueba original. Esta depuración es un proceso común en psicometría (Wieland et al., 2017), pues permite a los investigadores mejorar las propiedades psicométricas de las escalas, sean estas novedosas o preexistentes. Además, la abreviación (*brief*) de los instrumentos posee una utilidad práctica: siempre y cuando se mantenga la validez y confiabilidad de las escalas, las versiones breves son más rápidas de aplicar, por lo general son más sencillas de interpretar, y evitan el cansancio de quienes las responden (Leary, 1983). Incluso, diversos autores han optado por publicar conjuntamente el desarrollo de la escala y su versión abreviada (p. ej., Tangney et al., 2004).

En este caso, la depuración de la escala mantuvo la unidimensionalidad del constructo, algo que ya había sido reportado en estudios previos sobre la escala (p. ej., Baptista, 2012; Baptista & Cremasco, 2013). De hecho,

esta estructura de la depresión es similar a la reportada por otras medidas de la literatura (p. ej., Beck et al., 1996; Sanz et al., 2005).

A su vez, el AFE demostró que los ítems retenidos poseían buenas cargas factoriales e índices de comunalidad satisfactorios, lo que puede entenderse como buenas representaciones de los ítems del instrumento para la muestra (Sousa et al., 2015). De esta forma, el factor del EBADEP-A se resume como la representación estadística de los fenómenos psicológicos subyacentes a la depresión (Jeon et al., 2017). Tales resultados ya se habían encontrado en el estudio original del EBADEP-A con una muestra brasileira (Baptista, 2012).

Por otra parte, los análisis de la TRI demostraron que los ítems del EBADEP-A no revelan discrepancias en las habilidades estimadas de los participantes, pues se encontraron ajustes estructurales adecuados. Respecto a los índices *infit* y *outfit,* todos los ítems fueron adecuados, lo que indica opciones de respuestas claras; es decir, que quienes respondieron el instrumento comprendieron la frase del ítem y consiguieron discriminar las opciones de respuesta de la escala Likert (Abraham, & Chapelle, 1992). Estos resultados replican lo reportado previamente sobre el EBADEP-A (Baptista, 2012) y sobre otras versiones de la escala (Baptista, 2018b).

En detalle, la versión final de la adaptación al español rioplatense del EBADEP-A quedó conformada por un 33 % de ítems que evalúan el humor, un 22 % de ítems que evalúan síntomas sociales, un 19 % de ítems que evalúan síntomas cognitivos, un 12 % de ítems que evalúan síntomas vegetativos/somáticos, un 7 % de ítems que evalúan síntomas motores, y un 7 % de ítems que evalúan síntomas de irritabilidad. Respecto a esto, las categorías de síntomas cognitivos, motores y de irritabilidad presentan porcentuales semejantes a los encontrados por Calil y Pires (1998), lo que indica que la proporción de ítems finales para estas categorías resulta similar a la implementada por otros instrumentos. Esto refuerza la validez de contenido de la escala, puesto que indica que la misma incluye, en la caracterización de la depresión, aquellos aspectos que suelen ser considerados como formativos del trastorno (Thompson, 1989). Las demás categorías muestran un porcentual diferente, aunque similar al reportado en versiones previas de la escala (p. ej., Baptista, 2012).

Respecto al mapa de ítems y personas del EBADEP-A y el BDI-II, el análisis mostró que los ítems de los dos instrumentos se distribuían de modo equivalente en cuanto a la dificultad de adhesión, resultado esperado para los dos instrumentos —dado que evalúan el mismo constructo— (Baptista, 2012; Beck et al., 1996), pero diferente al encontrado en la construcción del instrumento original (Baptista, 2012), en el que la adhesión a los ítems del BDI-II resultó más difícil. Asimismo, se observó que los ítems de los dos instrumentos priorizaron de manera semejante los diversos niveles de habilidad. De esta forma, el EBADEP-A en la muestra argentina, similar a lo encontrado para el BDI-II, presenta una adecuada distribución, de manera que se puede afirmar que contempla diversos niveles de habilidad.

Ahora, acerca del análisis DIF, puede observarse un equilibrio en lo que se refiere a los sesgos observados entre los géneros, lo que descarta la posibilidad de que en alguno de ambos exista un favorecimiento hacia la adhesión de los ítems que evalúan depresión (Abraham & Chapelle, 1992). Y, respecto al AFC, este análisis presentó buenos índices de ajuste, buenas cargas factoriales y comunalidades entre los ítems, lo que demuestra que el EBADEP-A posee propiedades psicométricas adecuadas para la muestra bajo estudio, y confirma el modelo testeado (Hopwood & Donnellan, 2010; MacCallum et al., 1996).

En cuanto a las correlaciones y el modelo de regresión, los resultados muestran una relación negativa entre los síntomas depresivos presentados por la muestra y la mayor parte de los constructos evaluados. Específicamente, respecto al bienestar psicológico (EBP), era esperable encontrar una relación inversa, es decir que niveles más altos de síntomas depresivos se asociaron niveles más bajos de bienestar, lo cual apunta en la misma dirección que otros estudios previos (p. ej., Okbay et al., 2016; Proyer et al., 2014; Rogers et al., 2008; Ryff, 2013; Siddaway et al., 2017).

Asimismo, respecto a los rasgos de la personalidad del modelo del Big Five —apertura a la experiencia, amabilidad, extraversión y responsabilidad—, los resultados demostraron relaciones inversas con la sintomatología depresiva, lo que ha sido discutido en diversos estudios (Bartholomeu et al., 2010; McCrae & Costa, 2008; Soto & John, 2017). En particular, el rasgo de neuroticismo presentó una relación directa con los síntomas depresivos, resultado también encontrado en el metaanálisis de Smith et al. (2016), donde se reporta la existencia de evidencias que apuntan a que las personas con altos niveles de neuroticismo resultan más vulnerables a sufrir síntomas depresivos. Estos resultados son además corroborados cuando en la regresión múltiple el neuroticismo presenta la mayor contribución entre los cinco rasgos ($\beta = .289$) dentro del modelo que predice síntomas depresivos (EBADEP-A).

Finalmente, el análisis SEM demostró que dos ítems del BDI-II no obtuvieron cargas factoriales $\geq .30$, aunque esto no compromete al modelo analizado (Hair et al., 2014), pues el beta obtenido respecto a la magnitud del efecto encontrado entre las variables latentes evidencia una alta relación entre estas; y el r^2 muestra que el EBADEP-A explica un alto porcentaje del BDI-II —valor que ya era esperado,

por tratarse de instrumentos que evalúan el mismo rasgo latente—. Esta relación alta entre los dos instrumentos ya había sido reportada por Baptista et al. (2012), aunque estos hallazgos se realizaron en una muestra brasileña, y por medio de correlaciones y no análisis de regresión. Estos resultados, sumados a los índices de ajustes del modelo, suponen evidencias respecto de la validez concurrente del EBADEP-A (Chatterji & Levine, 2008). Teniendo todo lo anterior en cuenta, la adaptación del instrumento bajo estudio se ha mostrado adecuada para la evaluación de síntomas depresivos en una muestra argentina.

Por último, en cuanto a las limitaciones de este estudio, estas refieren a la necesidad de trabajar con una muestra clínica que confirme los hallazgos, así como la sugerencia de administrar el EBADEP-A en una muestra argentina con mayor diversidad. Además, se hacen necesarias otras investigaciones que indaguen respecto a las propiedades psicométricas del instrumento para la muestra argentina, por ejemplo, con análisis de la invariancia de los parámetros en diferentes grupos poblacionales, lo que debe ser explorado en estudios futuros sobre el funcionamiento diferencial de la escala.

Asimismo, respecto al aporte del estudio, es importante recordar que las prevalencias globales de la depresión son preocupantes (WHO, 2017), en tanto que se trata de un trastorno psicológico que tiende a provocar serias discapacidades en las personas que la padecen (OMS, 2001), y a deteriorar en gran medida la calidad de vida (Bromet et al., 2011). Dado que muchas veces el trastorno no es identificado correctamente y que muchos de los afectados no reciben la atención profesional necesaria (Demyttenaere et al., 2004; OMS, 2018), el presente estudio supone un aporte a la evaluación adecuada y sistemática de la depresión, para favorecer su detección temprana, la administración de tratamientos pertinentes, y su respectivo seguimiento (Jeon et al., 2017; Pignone et al., 2002).

Referencias

Abraham, R. G., & Chapelle, C. A. (1992). The meaning of Cloze test scores: An item difficulty perspective. *The Modern Language Journal, 76*(4), 468-479. https://doi.org/10.2307/330047

Baptista, M. N. (2012). *Escala Baptista de Depressão (Versão Adulto) – EBDEP-A*. Vetor.

Baptista, M. N. (2018a). Avaliando "Depressões": dos critérios diagnósticos às escalas psicométricas. *Revista Avaliação Psicológica, 17*(3), 301-310. https://doi.org/10.15689/ap.2018.1703.14265.03

Baptista, M. N. (2018b). *Manual EBADEP-IJ-Escala Baptista de Depressão–infanto-juvenil.* Hogrefe.

Baptista, M. N., & Borges, L. (2016). Revisão integrativa de instrumentos de depressão em crianças/adolescentes e adultos na população brasileira. *Avaliação Psicológica*, *15*, 19-32. https://doi.org/10.15689/ap.2016.15ee.03

Baptista, M. N., & Cremasco, G. da S. (2013). Propriedades psicométricas da escala baptista de depressão infanto-juvenil (EBADEP-IJ). *Arquivos Brasileiros de Psicologia, Rio de Janeiro*, *65*(2), 198-213. https://www.redalyc.org/articulo.oa?id=229028926004

Baptista, M. N., & Gomes, J. O. (2011). Escala Baptista de Depressão (Versão Adulto) - EBADEP-A: evidências de validade de construto e de critério. *Psico-USF*, *16*(2), 151-161. https://doi.org/10.1590/S1413-82712011000200004

Baptista, M. N., Cardoso, H. F., & Gomes, J. O. (2012). Escala Baptista de Depressão (Versão Adulto) – EBADEP-A: validade convergente e estabilidade temporal. *Psico-USF, 17*(3), 407-416. https://doi.org/10.1590/S1413-82712012000300007

Baptista, M. N., Gomes, J. O., & Carneiro, A. M. (2013). Exploratory Study of the Diagnostic Abilities of the Baptista Depression Scale Adult Version (EBADEP-A). *Paidéia (Ribeirão Preto)*, *23*(56), 301-311. https://doi.org/10.1590/1982-43272356201304

Baptista, M. N., Primi, R., Carvalho, L. de F., Oliveira, J. G., & Elhai, J. D. (2017). Constructing a common scale between tests of depression: The use of item response theory for transferring of norms from the BDI to EBADEP-A. *Universitas Psychologica*, *16*(2), 1-11. https://doi.org/10.11144/Javeriana.upsy16-2.ccsb

Bartholomeu, D., Machado, A. A., Spigato, F., Bartholomeu, L. L., Cozza, H. F., & Montiel, J. M. (2010). Traços de personalidade, ansiedade e depressão em jogadores de futebol. *Revista Brasileira de Psicologia Do Esporte*, *3*(4), 98-114. http://pepsic.bvsalud.org/pdf/rbpe/v3n1/v3n1a07.pdf

Beck, A. T., Steer, R. A., Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. Psychological Corporation.

Beltrán, M. D., Freyre, M. Á., & Hernández-Guzmán, L. (2012). El Inventario de Depresión de Beck: Su validez en población adolescente. *Terapia Psicológica*, *30*(1), 5-13. https://doi.org/10.4067/S0718-48082012000100001

Bighetti, C. A., Alves, G. A. S., & Baptista, M. N. (2014). Escala Baptista de Depressão (EBADEP-A): evidências de validade com o Big Five. *Avaliação Psicológica, 13*(19), 29-36. http://pepsic.bvsalud.org/pdf/avp/v13n1/v13n1a05.pdf

Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., Girolamo, G., Kessler, R. C. (2011). Cross-national epidemiology of DSM-IV major depressive episode. *BMC Medicine*, *9*(90), 2-16. https://doi.org/10.1186/1741-7015-9-90

Calil, H. M., & Pires, M. L. N. (1998). Aspectos gerais das escalas de avaliação de depressão. *Revista de Psiquiatria*

Cínica, *25*(5), 240-244. https://pesquisa.bvsalud.org/portal/resource/pt/lil-228050?lang=es

Chatterji, A. K, & Levine, D. I. (2008). *Imitate or Differentiate? Evaluating the validity of corporate social responsibility ratings. UC Berkeley: Center for Responsible Business*. https://escholarship.org/uc/item/3sz7k7jc

Coutinho, F. L., Hamdan, A. C., & Baptista, M. N. (2016). Escala Baptista de Depressão para Idosos – EBADEP-ID: evidências de validade. *Perspectivas en Psicología*, *13*(2), 1-9. https://dialnet.unirioja.es/servlet/articulo?codigo=5759612

Cremasco, S., & Baptista, M. N. (2018). Análise de estrutura interna da Escala Baptista de Depressão Versão Hospital-Ambulatório (EBADEP-HOSP-AMB). *Interação em Psicologia*, *22*, 144-154. https://doi.org/10.5380/psi.v22i2.55458

Del Valle, M. V., Hormaechea, F., & Urquijo, S. (2015). El bienestar psicológico: diferencias según sexo en estudiantes universitarios y diferencias con población general. *Revista Argentina de Ciencias del Comportamiento*, *7*(3), 6-13. https://core.ac.uk/download/pdf/334390358.pdf

Demyttenaere, K., Bruffaerts, R., Posada-Villa, J., Gasquet, I., Kovess, V., Lepine, J., … Kikkawa, T. (2004). Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *Jama*, *291*(21), 2581-2590. https://doi.org/10.1001/jama.291.21.2581

Díaz, D., Rodríguez-Carvajal, R., Blanco, A., Moreno-Jiménez, B., Gallardo, I., Valle C., & van Dierendonck, D. (2006). Adaptación española de las escalas de bienestar psicológico de Ryff. *Psicothema, 18*(3), 572-577. http://www.psicothema.com/psicothema.asp?id=3255

Embretson, S. E., & Reise, S. P. (2013). *Item response theory for psychologist*. Lawrence Erlbaums.

Figueiredo-Filho, D. B., & Silva Júnior, J. A. D. (2010). Visão além do alcance: uma introdução à análise fatorial. *Opinião pública*, *16*(1), 160-185. https://doi.org/10.1590/S0104-62762010000100007

Gomes, J., & Baptista, M. (2014). Normalization procedure for the Baptista Depression Scale-adult version (EBADEP-A): transferring of norms. *Avances en Psicología Latinoamericana*, *32*(3), 419-432. https://doi.org/10.12804/apl32.03.2014.02

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2007). *Análise multivariada de dados* (6. ª ed). Bookman.

Hair, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review*, *26*(2), 106-121. https://doi.org/10.1108/EBR-10-2013-0128

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review*, *14*(3), 332-346. https://doi.org/10.1177/1088868310361240

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

Jeon, S. W., Han, C., Ko, Y. H., Yoon, S. Y., Pae, C. U., Choi, J., Chon, Y.,Kim, J., Yoon, H., Ko, S., Patkar, A., & Zimmerman, M. (2017). Measurement-based treatment of residual symptoms using clinically useful depression outcome scale: Korean validation study. *Clinical Psychopharmacology and Neuroscience*, *15*(1), 28-34. https://doi.org/10.9758/cpn.2017.15.1.28

Kuo, D. C., Tran, M., Shah, A. A., & Matorin, A. (2015). Depression and the Suicidal Patient. *Emergency Medicine Clinics of North America*, *33*(4), 765-778. https://doi.org/10.1016/j.emc.2015.07.005

Ledesma, R., Sánchez, R., & Díaz, C. (2011). Adjective Checklist to Assess the Big Five Personality Factors in the Argentine Population. *Journal of Personality Assessment, 93*(1), 46-55. https://doi.org/10.1080/00223891.2010.513708

Leary, M. R. (1983). A Brief Version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin, 9*(3), 371-375. https://doi.org/10.1177/0146167283093007

Linacre, J. M. (2011). *A user's guide to Winsteps, Program Manual 3.74.0.* Winsteps.

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicologia, 30*(3), 1151-1169. https://doi.org/10.6018/analesps.30.3.199361

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modelling. *Psychological Methods*, *1*(2), 130-149. https://doi.org/10.1037//1082-989X.1.2.130

McCrae, R. R., & Costa, P. T. (2008). The Five-Factor Theory of personality. En O. P. John, R. W. Robins & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3. ª ed., pp. 159-181). Guilford Press.

Michaelides, M. P. (2008). An Illustration of a Mantel-Haenszel Procedure to Flag. *Practical Assessment, Research & Evaluation*, *13*(7), 1-16. https://gnosis.library.ucy.ac.cy/handle/7/37475

Montero, I., & León, O. G. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology, 7*(3), 847-862. https://www.redalyc.org/articulo.oa?id=33770318

Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, *25*(2), 151-157. https://doi.org/10.7334/psicothema2013.24

Okbay, A., Baselmans, B. M. L., De Neve, J. E., Turley, P., Nivard, M. G., Fontana, M. A., … Cesarini, D. (2016). Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics*, *48*(6), 624-633. https://doi.org/10.1038/ng.3552

Organización Mundial de la Salud (OMS). (2001). Informe sobre la salud en el mundo 2001. Salud mental: nuevos conocimientos, nuevas esperanzas. https://www.who.int/whr/2001/en/whr01_es.pdf

Organización Mundial de la Salud (OMS). (2018). Depresión. http://www.who.int/es/news-room/fact-sheets/detail/depression

Pignone, M. P., Gaynes, B. N., Rushton, J. L., Burchell, C. M., Orleans, C. T., Mulrow, C. D., & Lohr, K. N. (2002). Screening for depression in adults: a summary of the evidence for the US Preventive Services Task Force. *Annals of Internal Medicine*, *136*(10), 765-776. https://doi.org/10.7326/0003-4819-136-10-200205210-00013

Primi, R. (2012). Psicometria: fundamentos matemáticos da Teoria Clássica dos Testes. *Avaliação Psicológica*, *11*(2), 297-307. http://pepsic.bvsalud.org/pdf/avp/v11n2/v11n2a15.pdf

Proyer, R. T., Gander, F., Wellenzohn, S., & Ruch, W. (2014). Positive psychology interventions in people aged 50-79 years: Long-term effects of placebo-controlled online interventions on well-being and depression. *Aging and Mental Health*, *18*(8), 997-1005. https://doi.org/10.1080/13607863.2014.899978

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests* (Expanded ed.). University of Chicago Press.

Rogers, C. H., Floyd, F. J., Seltzer, M. M., Greenberg, J., & Hong, J. (2008). Long-term effects of the death of a child on parents' adjustment in midlife. *Journal of Family Psychology*, *22*(2), 203-211. https://doi.org/10.1037/0893-3200.22.2.203

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, *57*(6), 1069-1081. https://psycnet.apa.org/doi/10.1037/0022-3514.57.6.1069

Ryff, C. D. (2013). Psychological well-being revisited: Advances in the science and practice of eudaimonia. *Psychotherapy and Psychosomatics*, *83*(1), 10-28. https://doi.org/10.1159/000353263

Sánchez, R. O., & Ledesma, R. D. (2013). Listado de Adjetivos para Evaluar Personalidad: Propiedades y normas para una población argentina. *Revista Argentina de Clínica Psicológica*, *22*(2), 147-161. http://hdl.handle.net/11336/25481

Sanz, J., García-Vera, M. P., Espinosa, R., Fortún, M., & Vázquez, C. (2005). Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 3. Propiedades psicométricas en pacientes con trastornos psicológicos. *Clínica y Salud, 16*(2), 121-142. https://journals.copmadrid.org/clysa/art/37bc2f75bf1bcfe8450a1a41c200364c

Sanz, J., & Vázquez, C. (1998). Fiabilidad, validez y datos normativos del Inventario para la Depresión de Beck. *Psicothema, 10*(2), 303-318. http://www.psicothema.com/psicothema.asp?id=167

Sanz, J., & Vázquez, C. (2011). *Adaptación española del Inventario para Depresión de Beck-II (BDI-II). Manual*. Pearson.

Siddaway, A. P., Wood, A. M., & Taylor, P. J. (2017). The Center for Epidemiologic Studies-Depression (CES-D) scale measures a continuum from well-being to depression: Testing two key predictions of positive clinical psychology. *Journal of Affective Disorders*, *213*, 180-186. https://doi.org/10.1016/j.jad.2017.02.015

Smith, M. M., Sherry, S. B., Rnic, K., Saklofske, D. H., Enns, M., & Gralnick, T. (2016). Are Perfectionism Dimensions Vulnerability Factors for Depressive Symptoms After Controlling for Neuroticism? A Meta-analysis of 10 Longitudinal Studies. *European Journal of Personality*, *30*(2), 201-212. https://doi.org/10.1002/per.2053

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117-143. https://doi.org/10.1037/pspp0000096

Sousa, L. M., Marques-Vieira, C. M., Carvalho, M. L., Veludo, F., & José, H. M. (2015). Fidelidade e validade na construção e adequação de instrumentos de medida. *Enformação*, *5*, 25-32. https://www.researchgate.net/publication/303142007_Fidelidade_e_validade_na_construcao_e_adequacao_de_instrumentos_de_medida

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, *72*(2), 271-324. https://doi.org/10.1111/j.0022-3506.2004.00263.x

Thompson, C. (1989). *Affective disorders.* En C. Thompson (Ed.), *The instruments of psychiatric research* (pp. 87-126). John Wiley & Sons.

Urquijo, S., del Valle, M., & Baptista, M. (2017). *Evaluación de la depresión. Adaptación de la Escala Baptista de depresión para adultos –EBADEP-A-*. En XVI Reunión Nacional y V Congreso Internacional de la Asociación Argentina de Ciencias del Comportamiento (AACC), San Luis, Argentina.

Vélez, C. M., Ramírez, A. C., Arias, A. C., & Eslava-Schmalbach, J. H. (2016). Validación por modelo de Rasch del Cuestionario de Calidad de Vida (PedsQL 4.0®) en niños y adolescentes colombianos. *Revista Colombiana de Psiquiatría*, *45*(3), 186-193. https://doi.org/10.1016/j.rcp.2015.12.002

Vilagut, G., Forero, C. G., Barbaglia, G., & Alonso, J. (2016). Screening for Depression in the General Population with the Center for Epidemiologic Studies Depression (CES-D): A Systematic Review with Meta-Analysis. *PLOS ONE, 11*(5), e0155431. https://doi.org/10.1371/journal.pone.0155431

Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., … Coggeshall, M. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic

analysis for the Global Burden of Disease Study 2015. *The Lancet*, *388*(10053), 1545-1602. https://doi.org/10.1016/S0140-6736(16)31678-6

Wang, J., Wu, X., Lai, W., Long, E., Zhang, X., Li, W., … Wang, D. (2017). Prevalence of depression and depressive symptoms among outpatients: a systematic review and meta-analysis. *BMJ Open*, *7*(8), 1-15. https://doi.org/10.1136/bmjopen-2017-017173

Wieland, A., Durach, C. F., Kembro, J., & Treiblmaier, H. (2017). Statistical and judgmental criteria for scale purification. *Supply Chain Management*, *22*(4), 321-328. https://doi.org/10.1108/SCM-07-2016-0230

Williams, Jr., J., Pignone, M., Ramírez, G., & Stellato, C. P. (2002). Identifying depression in primary care: a literature synthesis of case-finding instruments. *General Hospital Psychiatry*, *24*(4), 225-237. https://doi.org/10.1016/S0163-8343(02)00195-0

World Health Organization (2017). *Depression and other common mental disorders: global health estimates*. WHO Document Production Services.

World Health Organization (2014). Health for the world's adolescents: A second chance in the second decade. http://public.tableausoftware.com/profile/digitalteam#!/vizhome/shared/3JW3RBSZ3