

Escala de Desesperanza de Beck (BHS): ventajas de una administración adaptativa

Beck Hopelessness Scale (BHS): advantages of an adaptive administration



Facundo Juan Pablo **Abal**
Juan Franco **Sanchez Gonzalez**
Gabriela Susana **Lozzia**
Horacio Félix **Attorresi**

Creativa Images

Photo By/Foto:

Rip
14¹

Volumen 14 #1 ene-abr
14 Años

Revista Iberoamericana de
Psicología

ISSN-I: 2027-1786 | e-ISSN: 2500-6517

Publicación Cuatrimestral

ID: 10.33881/2027-1786.RIP.14107

Title: Beck Hopelessness Scale (BHS):
advantages of an adaptive
administration

Título: Escala de Desesperanza de
Beck (BHS): ventajas de una
administración adaptativa

Alt Title / Título alternativo:

[en]: Beck Hopelessness Scale (BHS):
advantages of an adaptive
administration

[es]: Escala de Desesperanza de
Beck (BHS): ventajas de una
administración adaptativa

Author (s) / Autor (es):

Abal, Sanchez Gonzalez, Lozzia & Attorresi

Keywords / Palabras Clave:

[en]: Hopelessness, Beck Hopelessness
Scale, Computerized adaptive test,
Two-parameter logistic model, Item
Response Theory, Psychometry.

[es]: Desesperanza, Escala de
Desesperanza de Beck, Test
adaptativo informatizado, Modelo
logístico de dos parámetros, Teoría
de respuesta al ítem, Psicometría.

Proyecto / Project:

Universidad de Buenos Aires (UBACyT2018
20020170100200BA y UBACyT 2020
20020190200156BA) y Agencia
Nacional de Promoción Científica y
Tecnológica (PICT-2017-3226)

Submitted: 2020-08-25

Accepted: 2020-10-05

Resumen

La escala de Desesperanza de Beck (BHS) es usada con frecuencia como prueba de tamizaje para la detección de riesgo suicida y/o depresión aun cuando, por su extensión de 20 ítems, resulta poco práctica. En este trabajo se analiza si una estrategia de administración adaptativa podría abreviar el tiempo de aplicación de la BHS. Participaron 783 individuos de población general (50.9% mujeres). Se seleccionó aleatoriamente un 70% de los casos para calibrar los ítems con el Modelo Logístico de 2 parámetros de la Teoría de Respuesta al Ítem. Se eliminaron dos elementos que presentaron funcionamiento inadecuado. El 30% restante de la muestra se utilizó para simular una administración adaptativa de los 18 ítems calibrados. Se compararon dos modalidades de interrupción: a) al administrar 9 ítems y b) al alcanzar un error de estimación ≤ 0.35 o administrar 9 ítems (criterio mixto). Bajo ambas condiciones se registraron correlaciones de .95 con el nivel de Desesperanza estimado a partir de los 18 ítems. No obstante, la interrupción basada en el criterio mixto no mostró un impacto adicional en la eficiencia de la medida. Al igual que la versión de 18 ítems, las aplicaciones adaptativas estimaron con más precisión los niveles elevados del rasgo. La medición adaptativa no afectó las evidencias de validez al estudiar la asociación del rasgo con facetas del Neuroticismo y dimensiones sintomatológicas. Se concluye que una administración adaptativa de 9 ítems puede abreviar considerablemente la BHS sin perjudicar la validez y confiabilidad de la medida.

Abstract

The Beck Hopelessness Scale (BHS) is frequently used as screening tool for suicide risk and / or depression even though, due to its length of 20 items, it is not very practical. This work analyses whether an adaptive administration strategy could shorten the application time of the BHS. Seven hundred eighty-three individuals from the general population (50.9% women) participated. Seventy percent of the cases were randomly selected to calibrate the items according to the two-parameter logistic model of the Item Response Theory. Two items that presented inadequate functioning were eliminated. The remaining 30% of the sample was used to simulate an adaptive administration from the 18 calibrated items. Two stopping rules were compared: a) administration of 9 items and b) estimation error ≤ 0.35 or administration of 9 items (mixed criterion). Under both conditions, correlations of .95 were registered with the levels of Hopelessness estimated from the 18-items version. However, the stopping based on the mixed criterion did not show an additional impact on the efficiency of the measure. Like the 18-item version, adaptive administrations estimated with more reliability the high trait levels. Adaptive administrations did not vary the evidence of validity when studying the association of the trait with facets of Neuroticism and symptomatic dimensions. It is concluded that an adaptive administration of 9 items can significantly shorten the BHS without compromising the validity and reliability of the measure.

Citar como:

Abal, F. J., Sanchez Gonzalez, J. F., Lozzia, G. S., & Attorresi, H. F. (2021). Escala de Desesperanza de Beck (BHS): ventajas de una administración adaptativa. *Revista Iberoamericana de Psicología*, 14 (1), 71-82. Obtenido de: <https://reviberopsicologia.iber.edu.co/article/view/1940>

Dr. Facundo Juan Pablo **Abal**

AutorID: [34972825400](https://orcid.org/0000-0001-7023-5380)

Research ID: [AAI-5650-2020](https://orcid.org/0000-0001-7023-5380)

ORCID: [http://orcid.org/0000-0001-7023-5380](https://orcid.org/0000-0001-7023-5380)

Source | Filiación:

Consejo Nacional de Investigaciones Científicas y Técnicas y Facultad de Psicología de la Universidad de Buenos Aires

BIO:

Dr. en Psicología. Profesor Adjunto en la Cátedra II de Estadística. Investigador Adjunto Consejo Nacional de Investigaciones Científicas y Técnicas.

Juan Franco **Sanchez Gonzalez**, Lic

AutorID: =

Research ID: [AAI-9908-2020](https://orcid.org/0000-0003-1993-4981)

ORCID: <https://orcid.org/0000-0003-1993-4981>

Source | Filiación:

Facultad de Psicología - Universidad de Buenos Aires

BIO:

Ayudante de Trabajos Prácticos de Segundo en la Cátedra I de Estadística. Investigador en la Facultad de Psicología de la Universidad de Buenos Aires.

Dra. Gabriela Susana **Lozzia**

AutorID: [14071676700](https://orcid.org/0000-0001-7753-6303)

Research ID: [AAI-7441-2020](https://orcid.org/0000-0001-7753-6303)

ORCID: <https://orcid.org/0000-0001-7753-6303>

Source | Filiación:

Facultad de Psicología - Universidad de Buenos Aires

BIO:

Profesora Adjunta de Estadística e Investigadora en la Facultad de Psicología de la Universidad de Buenos Aires.

Horacio Félix **Attorresi**, Lic

AutorID: =

Research ID: [ABH-6888-2020](https://orcid.org/0000-0002-3027-1069)

ORCID: [http://orcid.org/0000-0002-3027-1069](https://orcid.org/0000-0002-3027-1069)

Source | Filiación:

Facultad de Psicología - Universidad de Buenos Aires

BIO:

Profesor Titular Consulto de la Universidad de Buenos Aires. Investigador en la Facultad de Psicología de la Universidad de Buenos Aires.

Escala de Desesperanza de Beck (BHS): ventajas de una administración adaptativa

Beck Hopelessness Scale (BHS): advantages of an adaptive administration

Facundo Juan Pablo **Abal**
Juan Franco **Sanchez Gonzalez**
Gabriela Susana **Lozzia**
Horacio Félix **Attorresi**

La Desesperanza configura un sistema de expectativas negativas que tiene una persona sobre sí misma y sobre su futuro tanto inmediato como lejano (**Beck, Weissman, Lester y Trexler, 1974**). En esencia, este constructo se compone de dos elementos básicos: a) expectativas desfavorables sobre los posibles logros y resultados futuros y b) expectativas negativas acerca de la posibilidad de cambiar o revertir ese futuro. Desde sus orígenes como constructo, la Desesperanza ha ocupado un rol importante en el estudio de los trastornos depresivos (**Beck & Steer, 1989; Kliem, Lohmann, Mößle & Brähler, 2018; Pompili et al., 2013**), estrés postraumático (**Boffa, King, Turecki, & Schmidt, 2018**), esquizofrenia (**Aloba, Esan & Alimi, 2017; Kao, Liu & Lu, 2012**) y, particularmente, en condiciones clínicas como las conductas suicidas y/o autolesivas (**McMillan, Gillbody, Beresford & Neilly, 2007; Steeg et al., 2016; Wolfe et al., 2017**). Su impacto también ha sido analizado en diversas enfermedades físicas como cardiopatías (Dunn et al., 2014), diabetes (Innamorati et al., 2013), VIH (Flores & Almanza, 2013) y cáncer (Nissim et al., 2010; Parpa, Tsilika, Galanos, Nikoloudi, & Mystakidou, 2019; Spangenberg et al., 2016).

El estudio objetivo de la Desesperanza se inició con la Beck Hopelessness Scale (BHS) construida por Beck et al. (**1974**), quienes por primera vez entendieron este constructo como continuo cuantitativo capaz de discriminar diferencias individuales. La solidez psicométrica de esta escala y su contribución en el campo de clínico y de la salud propiciaron su adaptación en diversas culturas (**e.g. Aloba, Awe, Adelola, Olatunji & Aloba, 2018; Bouvard, Charles, Guerin & Aimard, 1992; Hermosillo-de la Torre, Méndez-Sánchez & Gonzalez-Betanzos, 2020; Kim et al., 2015; Kliem et al., 2018; Mikulic, Cassullo, Crespi & Marconi, 2009; Perczel Forintos, Sallai & Rózsa, 2001; Tanaka, Sakamoto, Ono, Fujihara & Kitamura, 1998; Yip & Cheung, 2006**).

Aunque en la mayoría de las investigaciones han involucrado a grupos clínicos, también son numerosos los trabajos que analizan la aplicación de la BHS en muestras no-clínicas tanto de población general (Illiceto & Fino, 2015; Kocalevent et al., 2016; Satorres et al., 2016) como de estudiantes (Alamo, Baader, Antúnez, Bagladi & Bejer, 2019; Aloba et al., 2018; Lester, 2013; Quiñonez Tapia, Méndez Luevano & Castañeda Camey, 2019). Los objetivos principales de estos estudios han sido conocer la estructura factorial del constructo en población con mayoría de personas sin manifestaciones clínicas y validar este instrumento como tamizaje para la detección temprana de personas con riesgo suicida y/o depresión.

A pesar de su uso extendido, es dable reconocer que el formato de la BHS tiene aspectos poco deseables cuando se pretende utilizar con fines de tamizaje o estudios epidemiológicos a gran escala, en donde se buscan tiempos de administración reducidos. Aplicar los 20 ítems que componen la BHS puede resultar demasiado extenso si se considera que sólo aporta información sobre un único constructo. Además, un cuestionario extenso puede resultar una tarea exigente para una persona con niveles de Desesperanza elevados como los que se pretenden detectar con el tamizaje. En esta línea, numerosos autores han considerado la necesidad de abreviar la BHS seleccionando tres (Bouvard et al., 1992; Tanaka et al., 1998) o cuatro ítems (Aish & Wasserman, 2001; Perczel Forintos et al., 2013; Yip & Cheung, 2006) de la escala original. Incluso Aish y Wasserman (2001) sugirieron que sólo uno de los ítems (7. Mi futuro parece oscuro) podría ser suficiente para tener una medida rápida de Desesperanza. Sin embargo, para mantener los estándares de calidad psicométrica las versiones breves debieron complejizar el formato de respuesta dicotómico (Verdadero/Falso) que tiene la BHS agregando más alternativas.

En las últimas décadas la Psicometría ha tenido importantes desarrollos ligados a la consolidación de la Teoría de Respuesta al Ítem (TRI) como perspectiva psicométrica (Muñiz, 2018). Aunque las primeras aplicaciones de la TRI se realizaron en el análisis de pruebas de habilidades (e.g. Gutiérrez Rincón & Escobar Melo, 2013) los avances teóricos y tecnológicos han permitido extender las nociones fundamentales a instrumentos que miden comportamientos típicos (Reise & Revicki, 2015). En este contexto, los Tests Adaptativos Informatizados (TAIs) se destacan por ofrecer un área de investigación relevante con importantes implicancias prácticas. Un TAI permite la aplicación progresiva de los ítems de un instrumento o banco considerando el nivel de rasgo estimado para el sujeto luego de cada respuesta. Así, se configura un algoritmo adaptativo que selecciona los ítems más informativos para medir el rasgo de cada persona, por lo que se consigue una secuencia de administración flexible y óptima. La administración de los ítems prosigue hasta que se alcanza un criterio de parada especificado por el diseño adaptativo; ya sea por administrar un número fijo de ítems inferior a la prueba completa (criterio longitud fija), por alcanzar un valor prefijado de error estándar en la estimación del rasgo (criterio longitud variable) o una combinación de ambas estrategias (criterio mixto) (Magis, Yan & von Davier, 2017; Olea & Ponsoda, 2013).

En esta línea se han implementado con éxito tests adaptativos de variables personológicas (e.g. Abal, Auné & Attorresi, 2019; Abal, Lozzia, Auné & Attorresi, 2020; Achtyes et al., 2015; Makransky, Mortensen & Glas, 2013; Nieto et al., 2017; Pedrosa-García, 2015) y del contexto clínico (Beiser, Vu & Gibbons, 2016; Choi, Schalet, Cook, & Cella, 2014; Devine et al., 2016; Fonseca-Pedrero, Menéndez, Paino, Lemos-Giráldez & Muñiz, 2013; Gibbons, Weiss, Frank, &

Kupfer, 2016; Magnée, de Beurs, Terluin & Verhaak, 2017; Stochl, Böhnke, Pickett & Croudace, 2016; Walter, 2010) con la finalidad de reducir los tiempos de administración y hacer más eficiente la evaluación sin comprometer la precisión o validez (Chang, 2015). Los TAIs permiten ajustar la administración de la prueba a las características del evaluado sin perder la objetividad en la medida ni la comparabilidad de personas que responden distintos subconjuntos de ítems del instrumento (Magis et al., 2017; Thompson & Weiss, 2011).

En este trabajo se propone examinar si la implementación de una estrategia de administración adaptativa para la escala BHS podría suponer una ventaja al ser comparada con la administración tradicional en formato fijo de su versión original. En consecuencia, se simuló administraciones adaptativas que permiten analizar empíricamente la posibilidad de reducir la cantidad de ítems aplicados proporcionando una estimación de la Desesperanza con igual o mayor nivel de precisión que la prueba completa. Esto es, los TAI no sólo permiten reducir los tiempos de administración sino que además, dado que seleccionan los ítems más informativos para cada sujeto, ofrecen una medida más precisa. Un TAI para la medición de la Desesperanza puede resultar particularmente útil cuando el instrumento se aplica con fines de tamizaje. Esto garantizaría la detección de niveles elevados de Desesperanza invirtiendo una menor cantidad de tiempo.

Método

Diseño

La investigación realizada responde a las características de un estudio instrumental (Ato, López, & Benavente, 2013). Se trata de un estudio de simulación que compara la eficiencia de diferentes criterios de interrupción de una administración adaptativa por sobre la aplicación tradicional de la BHS, considerando las posibles implicancias en las propiedades psicométricas del TAI (Olea & Ponsoda, 2013).

Participantes

Colaboraron 783 individuos de población general residentes en el área metropolitana de Buenos Aires, Argentina. Se realizó un muestreo no probabilístico por accesibilidad. El 50.9% de los participantes fueron mujeres y el 49.1% fueron varones. La edad osciló entre 18 y 83 años con una media de 35.3 (DE = 13.9). En cuanto al estado civil, el 50,4% de los individuos eran solteros, el 46.9% eran casados, el 1.9% eran divorciados y el 0.8% eran viudos. El 15.2% consignó estar empleado en una jornada de medio tiempo mientras que un 58.1% manifestó trabajar tiempo completo. El resto de los participantes refirió realizar exclusivamente tareas del hogar (8.2%), ser jubilados (4.9%) y estudiantes (9.1%). Sólo un 4.5% estaba desempleado. La mayoría de los participantes (70%) se autodescribió con un nivel socioeconómico medio y se observó un porcentaje similar entre quienes se identificaron con un nivel medio-bajo (15.3%) y medio-alto (14.7%). Con respecto al nivel de estudios, un 9.9% estuvo por debajo del nivel secundario y un 61.9% lo completó. El resto refirió haber alcanzado un nivel terciario (10.3%) o universitario (17.9%).

Los criterios de inclusión para los participantes de este estudio fueron: ser argentinos, mayores de 18 años y residir en el área metropolitana de Buenos Aires. Fueron excluidas las personas no dispuestas o incapacitadas para firmar el consentimiento informado.

Instrumentos

Escala de Desesperanza de Beck (**Beck et al., 1974**). Se utilizó la versión adaptada lingüística y conceptualmente por Mikulic et al. (**2009**). El instrumento se compone de 20 ítems dicotómicos (Verdadero/Falso), 11 de los cuales están redactados en el sentido del rasgo (ítems positivos) y 9 en sentido inverso. La adaptación lingüística local se efectuó usando la técnica de backward translation. Los adaptadores de la BHS reportaron evidencias de validez concurrente ($r = .43$) utilizando como criterio externo la adaptación argentina del Inventario de Depresión de Beck (BDI-II). En lo que respecta a la consistencia interna, obtuvieron un Alfa de Cronbach = **.78**, el cual puede ser interpretado como aceptable.

Banco de ítems de Neuroticismo (**BIN, Abal, et al., 2019**). El cuestionario operacionaliza las seis facetas del Neuroticismo según el modelo de McCrae y Costa (**2010**). Se compone de 54 ítems con formato de respuesta de cuatro opciones. Cuenta con evidencias de validez basadas en la estructura interna y análisis de ítems con TRI. Los coeficientes de consistencia interna Alfa de Cronbach obtenidos con la muestra del presente estudio fueron aceptables para todas las facetas: Ansiedad (**Alfa = .81**), Hostilidad (**Alfa = .72**), Depresión (**Alfa = .82**), Autoconciencia (**Alfa = .81**), Impulsividad (**Alfa = .77**) y Vulnerabilidad (**Alfa = .73**).

Inventario de síntomas SCL-90-R (**Derogatis, 1994**). Es un auto-reporte compuesto por 90 ítems que evalúan patrones de síntomas psicológicos con el fin de detectar personas en riesgo. Los elementos presentan un formato de respuesta politómica en la que el sujeto responde si el síntoma le ha preocupado o molestado nada, muy poco, poco, bastante o mucho durante los últimos siete días. Los ítems se agrupan para posibilitar la medición de nueve dimensiones clínicas y tres índices globales (Índice de Severidad Global, Total de Síntomas Positivos e Índice de Malestar Positivo). Los estudios de la adaptación argentina de este inventario muestran suficientes evidencias de validez y confiabilidad (**Sánchez & Ledesma, 2009**). Para el presente estudio, el análisis de la consistencia interna del inventario completo mostró un Alfa de Cronbach de .96. Para las dimensiones clínicas, este coeficiente mostró valores adecuados: Somatización (**Alfa = .79**), Obsesiones/Compulsiones (**Alfa = .82**), Sensitividad Interpersonal (**Alfa = .80**), Depresión (**Alfa = .86**), Ansiedad (**Alfa = .83**), Hostilidad (**Alfa = .78**), Ansiedad Fóbica (**Alfa = .77**), Ideación Paranoide (**Alfa = .78**) y Psicoticismo (**Alfa = .79**).

Procedimiento

Se han contemplado las normativas profesionales internacionales para la adaptación y la validación de tests aplicados en la práctica clínica (**American Educational Research Association, American Psychological Association y National Council on Measurement in Education, 2014; American Psychological Association, 2010**) y, más específicamente, en ámbito de investigación psicológica (**International Test Commission, 2014**). El protocolo examinado y fue aprobado por el Comité de Conductas Responsables de la Facultad de Psicología

de la Universidad de Buenos Aires, sede en donde se lleva adelante el proyecto de investigación del presente estudio (**UBACyT2020 20020190200156BA**).

Los participantes respondieron el protocolo en formato lápiz y papel de manera individual y sin tiempo límite. En todos los casos, fueron evaluados por profesionales debidamente entrenados. Antes de comenzar la administración se brindó información sobre los propósitos de la investigación y los procedimientos de evaluación que serían aplicados. Se leyó en voz alta el consentimiento informado ofreciendo una copia al evaluado para que pueda seguir su lectura y lo firme en el caso de estar de acuerdo. El consentimiento explicitaba que: a) la participación era voluntaria y que el individuo podría abstenerse de realizar la evaluación o incluso abandonarla cuando así lo deseara y b) sería resguardada la identidad del participante y se tomarían los recaudos necesarios para garantizar la confidencialidad de sus respuestas a los cuestionarios.

Análisis estadístico y psicométrico

Se dividió a los participantes en dos grupos de manera aleatoria. Las respuestas de un 70% de ellos ($n = 548$) fueron empleadas para calibrar los ítems según el modelo logístico de dos parámetros (ML2p) mientras que el resto de los individuos ($n = 235$) fueron considerados para analizar la eficiencia de la administración adaptativa.

Supuestos del ML2p. Ante la sospecha de potenciales violaciones al supuesto de independencia local de los ítems se estudió empíricamente este requisito del ML2p con el estadístico χ^2_{LD} de Chen y Thissen (**1997**). Se consideraron localmente independiente los pares de ítems con $\chi^2_{LD} > 10$ (**Reise & Rodríguez, 2016**). También se realizó un Análisis Factorial Confirmatorio (AFC) con el programa Mplus (**Muthén & Muthén, 2010**) a fin de verificar el supuesto de unidimensionalidad. Se estimaron los parámetros con el método robusto de mínimos cuadrados ponderados (Weighted Least Squares Mean and Variance Adjusted, WLSMV) sobre la base de la matriz de correlaciones tetracóricas. Para analizar el ajuste se aplicaron los criterios recomendados por Byrne (**2012**): CFI y TLI $\geq .90$ y RMSEA $\leq .05$.

Estimación y Ajuste del ML2p. La aplicación de la TRI se efectuó operando el programa IRTPRO 4.2 (**Cai, Thissen, & du Toit, 2011**). Se estimaron los parámetros de los ítems y los sujetos según el ML2p con el método de Máxima Verosimilitud Marginal. El ajuste del modelo a los datos se estudió a nivel del ítem con el estadístico $S-\chi^2$ de Kang y Chen (2011). Se consideró el criterio de Toland (**2013**) que establece un $p > .01$ para aceptar un ajuste adecuado. También se examinó el Funcionamiento Diferencial de los Ítems (DIF) en función del género a partir del test de Wald modificado (**Cai et al., 2011**).

Estudio de administración adaptativa. Se aplicó el software Fistar-D (**Choi, 2018**) para realizar simulaciones de administraciones adaptativas con los ítems calibrados en la fase anterior. El programa utiliza la matriz de datos de los sujetos apartados exclusivamente para este estudio ($n=235$) y escoge progresivamente aquellas respuestas correspondientes a los ítems que se le presentarían al evaluado si estuviera respondiendo un TAI.

Se estableció a la media del rasgo como una estimación razonable de θ para iniciar la administración. Se programó una selección aleatoria entre los cinco ítems más cercanos a la media para garantizar una mayor representatividad del contenido muestreado. Luego de la respuesta a cada ítem, se efectuó una estimación provisional de θ con el método bayesiano de estimación esperada a posteriori (EAP) usando a la normal estándar como distribución a priori. La selección progresiva de los ítems se realizó por el método de máxima información de Fisher (Magis et al., 2017).

Con el error de estimación obtenido para cada evaluado se pudo estudiar la precisión de la medida con la BHS en su formato tradicional y compararla con dos variantes adaptativas generadas por distintos criterios de parada: a) al administrar adaptativamente el 50% de los ítems calibrados (longitud fija) y b) al conseguirse un nivel de error en la estimación del rasgo ≤ 0.35 o aplicarse la mitad de los ítems (criterio mixto). Ambos criterios fueron ensayados pensando en las ventajas que traería su implementación con fines de tamizaje. El primero persigue reducir a la mitad el tiempo de administración para todos los individuos evaluados. El segundo criterio también persigue economizar la cantidad de ítems utilizados, pero alcanzando una precisión óptima (un error de 0.35 equivale a una confiabilidad clásica de .88) en la estimación de niveles de Desesperanza elevados. Bajo esta modalidad los sujetos con expectativas más desesperanzadoras, y por ende con mayor riesgo, podrían finalizar la aplicación antes de responder a la mitad de la escala por alcanzar un grado de precisión satisfactorio en la estimación de su nivel de rasgo.

Siguiendo la recomendación de Olea y Ponsoda (2013), se obtuvo el error de estimación medio para la versión con criterio de parada de longitud fija; mientras que para el de criterio mixto se obtuvo como indicador de su eficiencia la cantidad media de reactivos que se requirió para alcanzar el nivel de error prefijado. Con el fin de analizar si los distintos criterios de parada ofrecen una estimación del rasgo similar a la que se obtendría si la persona respondiera a todos los ítems, se correlacionaron los niveles θ estimados bajo cada condición de parada con los estimados al contestar el instrumento completo. Además, se correlacionaron los errores de estimación de los niveles del rasgo obtenidos mediante el TAI de longitud fija con los errores de estimación obtenidos con la BHS en su formato tradicional.

Por último, se analizó si la administración adaptativa (con sus variantes en los criterios de parada) afecta las evidencias de validez concurrente de la BHS con las escalas del SCL-90-R y las facetas del Neuroticismo del BIN. Para estos análisis se llevaron adelante pruebas de diferencia de r de Pearson para muestras relacionadas a fin de comparar si los coeficientes de correlación con estas escalas difieren significativamente al estimar las puntuaciones de Desesperanza con el TAI o con el test en formato de administración tradicional.

Resultados

Fase de depuración de la BHS

Se efectuó una aplicación preliminar del ML2p a los 20 ítems que componen la versión original de la BHS. Este análisis mostró la necesidad de eliminar dos ítems que presentaron parámetros con valores fuera de lo esperable. El ítem 13 adoptó valores negativos de discriminación ($a = -0.27$) y el ítem 20 presentó un valor elevado ($a = 4.27$) con un error de estimación alto ($s.e. = 1.26$). La estimación de parámetros que resulta con valores atípicos puede evidenciar problemas de ajuste de manera indirecta (Gray-Little, Williams & Hancock, 1997). Ante la sospecha de funcionamiento inadecuado se ha decidido descartar estos ítems de la calibración definitiva.

Fase de calibración de los ítems

Supuestos del ML2p. Todos los pares de ítems mostraron resultados favorables con respecto a la independencia local ($\chi^2_{LD} < 10$), siendo el máximo valor de χ^2_{LD} registrado para el par de ítems 7 y 18 ($\chi^2_{LD} = 6.3$). Estos resultados permiten concluir que los ítems son localmente independientes más allá de la similitud observada en los contenidos de algunos pares. Con respecto a la unidimensional, tanto el índice de ajuste absoluto $RMSEA = .039$ (IC 90% .031 - .046) como los índices de ajuste comparativos ($TLI = .92$ y $CFI = .93$) evidenciaron valores adecuados para sostener una única dimensión con cargas factoriales estadísticamente significativas ($p < .05$).

Aplicación del ML2p. Se requirieron 56 ciclos del proceso de estimación para alcanzar un criterio de convergencia 0.00001. Los parámetros a y b estimados para los 18 ítems retenidos luego de la depuración presentaron valores dentro de un rango esperable y con errores de estimación relativamente bajos. Como se observa en los $S-\chi^2$ de la tabla 1 todos los ítems mostraron un ajuste adecuado ($p > .01$). La tabla 1 también reporta los estadísticos χ^2 del test de Wald empleado para el estudio del DIF de los ítems en función del género. Dado que todos los p -valor fueron mayores a .05 se concluye que los ítems de la BHS no funcionan diferencialmente si son respondidos por mujeres o varones.

Los ítems tendieron a localizarse mayoritariamente por encima de la media del rasgo, lo cual es acorde con la finalidad de la escala para ser aplicada como instrumento de tamizaje. La media de los parámetros de localización fue de 1.68 ($DE = 0.78$) siendo -0.16 (ítem 4) y 2.72 (ítem 2) los respectivos valores mínimo y máximo. La discriminación de los ítems resultó media-alta con un valor promedio de $a = 1.64$ ($DE = 0.71$). Un 39% de los ítems mostró parámetros a superiores a 1.70, considerados como muy elevados según el criterio definido por Baker y Kim (2017).

Tabla 1. Calibración de los ítems y estudio del funcionamiento diferencial

Ítem	Calibración				Ajuste			Test de Wald	
	a	se	b	se	S- χ^2	gl	p	χ^2 (gl=2)	p
1	1.95	0.26	1.19	0.11	19.3	8	.02	0.3	.86
2	1.55	0.32	2.72	0.36	17.2	9	.05	1.6	.45
3	1.27	0.22	2.22	0.28	14.3	11	.22	1.3	.52
4	0.74	0.13	-0.06	0.13	11.7	9	.23	0.2	.88
5	0.64	0.12	1.22	0.25	13.7	11	.25	1.6	.44
6	1.77	0.33	2.38	0.27	13.1	11	.29	1.8	.41
7	2.15	0.35	1.71	0.15	17.6	9	.04	0.1	.96
8	0.82	0.15	1.71	0.27	16.9	10	.08	1.2	.54
9	2.57	0.53	2.18	0.21	11.5	8	.18	4.4	.11
10	1.41	0.21	1.6	0.17	8.9	11	.64	0.7	.70
11	3.54	0.77	1.8	0.14	3.9	7	.79	1.7	.43
12	1.43	0.23	2.06	0.23	24.4	11	.02	0.5	.79
14	1.17	0.17	1.33	0.16	12.9	9	.16	0.7	.71
15	2.4	0.33	0.93	0.08	14.2	8	.08	0.1	.96
16	1.86	0.4	2.7	0.34	6.5	7	.48	1.9	.38
17	1.43	0.29	2.71	0.37	10.5	10	.40	2.3	.31
18	1.5	0.2	0.31	0.08	13.3	7	.06	3.7	.16
19	1.24	0.19	1.61	0.19	10.7	10	.38	0.2	.92

Autoría Propia

Nota. a = Parámetro de discriminación del Modelo Logístico de dos parámetros. b = Parámetro de localización del Modelo Logístico de dos parámetros. se = Error de estimación del parámetro. S- χ^2 = estadístico de ajuste de Kang y Chen (2011). gl = grados de libertad. χ^2 = estadístico del test de Wald.

Fase de simulación de la administración adaptativa

Versión convencional de 18 ítems. Los estudios de simulación se llevaron adelante con la muestra de 235 participantes a partir de sus respuestas a los 18 ítems calibrados (tabla 2). A través de la administración de la BHS se obtuvo una mediana de estimación de θ de **0.20**, oscilando entre **-0.92 y 2.73**. En esta muestra el error de estimación de los θ mostró una mediana de 0.42 por lo que es posible afirmar que, globalmente, la evaluación con los 18 ítems presentó una confiabilidad

clásica de aproximadamente .82. El error de estimación de θ mínimo obtenido al responder la BHS completo fue de 0.19. Este dato indica el error más bajo que se podría alcanzar al medir la Desesperanza de un sujeto, ya que la mejor estimación de un nivel de rasgo (en términos de precisión) es la que se logra al administrar todos los ítems que componen la escala. Por otro lado, el error máximo presentó un valor relativamente elevado de 0.68 debido a que los ítems se localizaron mayoritariamente en los niveles altos del rasgo. Por ende, los valores bajos de Desesperanza se estimaron con errores más elevados. Esta tendencia es apreciable a partir de la intensa correlación negativa ($r = -.97, p < .001$) registrada entre los θ estimados con la BHS y sus respectivos errores.

Tabla 2. Resultados de la administración de la BHS de 18 ítems y sus versiones adaptativas

	BHS18		TAI9		TAI.35	
	θ	EE(θ)	θ	EE(θ)	θ	EE(θ)
Mediana	0.20	0.42	0.21	0.46	0.22	0.46
Mínimo	-0.92	0.19	-0.91	0.21	-0.91	0.25
Máximo	2.73	0.68	2.75	0.75	2.60	0.74

Autoría Propia

Nota. BHS18 = Versión de la Beck Hopelessness Scale con ítems calibrados; TAI9 = Versión adaptativa de formato fijo.; TAI.35 = Versión adaptativa con criterio mixto; θ = Nivel de rasgo estimado para el evaluado. EE(θ) = Error de estimación del rasgo.

Comparación con versiones adaptativas. Si se consideran los resultados obtenidos al administrar adaptativamente la mitad de los ítems calibrados de la BHS (versión de longitud fija), se obtuvo una mediana en el nivel estimado de rasgo de **0.21**, oscilando entre **-0.91 y**

2.75. La mediana del error de estimación fue **0.46 (mínimo = 0.21 y máximo = 0.75)**, lo que resulta equivalente a una confiabilidad clásica aproximada de 0.79. Esto indica que al administrar una menor cantidad de reactivos (50% menos) se verifica un error levemente superior

al obtenido con la versión tradicional de la BHS (0.42 versus 0.46). En esta línea, es posible inferir que para la mitad de los sujetos evaluados la confiabilidad del TAI, en términos clásicos, fue similar o superior a la que se obtuvo con el test en su formato convencional, pero requirió administrar un 50% menos de ítems. Además, las correlaciones fueron altas y positivas entre los valores θ estimados con 9 y 18 ítems ($r = .95, p < .001$) y, también, entre sus errores de estimación ($r = .93, p < .001$).

Similares resultados se encontraron al analizar el TAI con criterio de parada mixto (i.e. la administración se interrumpe al alcanzar un error de estimación menor o igual a 0.35 o, en su defecto, al administrar hasta 9 ítems). La mediana del nivel estimado de rasgo fue de 0.22 (mínimo = -0.91 y máximo = 2.60). Los errores de estimación variaron entre 0.25 y 0.74 con una mediana 0.46. Como es posible observar, al igual que para el TAI de longitud fija, el error de estimación también resultó levemente superior al obtenido con los 18 ítems de la BHS (0.42 contra 0.46). Los θ estimados con la BHS completa se asociaron de manera intensa y positiva ($r = .95, p < .001$) con las estimaciones obtenidas del TAI con criterio de parada mixto. También se registró una correlación elevada entre sus errores de estimación ($r = .92, p < .001$). En esta variante adaptativa se necesitó que los sujetos de la muestra respondan un promedio de 8.35 (DE= 0.82) ítems, con un mínimo de 4 y un máximo de 9 ítems (que estaba fijado como límite). Sólo 12.8% de los participantes alcanzaron un error de 0.35 antes de la presentación

del ítem número 9. Esto significa que la mayoría de los evaluados no alcanzó el error prefijado y la administración se interrumpió por alcanzar el tope de cantidad de ítems.

Considerando que los errores máximos de los TAIs fueron elevados se decidió corroborar si los niveles altos de Desesperanza, de particular interés en un tamizaje, presentaban un error de estimación óptimo. Para ello se analizaron los errores de estimación pertenecientes a los sujetos de la muestra que superaron el percentil 75 en el rasgo. Efectivamente, los errores máximos registrados por esta submuestra en las versiones adaptativas fueron de 0.34 (longitud fija) y 0.36 (criterio mixto), evidenciando que es menor que el error máximo con que fueron evaluados todos los sujetos de la muestra (errores de 0.75 y 0.74 respectivamente).

Relación con criterios externos. Finalmente, la asociación de la BHS (y sus versiones adaptativas) con las escalas del BIN y SCL-90-R arrojaron resultados que responden a lo esperable a nivel teórico y aportan evidencias de validez concurrente (tabla 3). Asimismo, los coeficientes de correlación r de Pearson obtenidos no difieren significativamente si las asociaciones con los criterios externos se calculan con estimaciones de θ surgidas del test completo o con alguna las versiones adaptativas. En consecuencia, la reducción en la cantidad de ítems de las administraciones adaptativas no muestra un impacto negativo en las evidencias de validez basadas en la relación con otras variables.

Tabla 3. Correlaciones de las estimaciones del TAI y el test completo con criterios externos.

	rBHS18	rTAI9	Diferencia de r de Pearson: rBHS18- rTAI9	rTAI.35	Diferencia de r de Pearson: rBHS18- rTAI.35
Facetas de Neuroticismo (BIN)					
Ansiedad	.409**	.390**	t(232)=1; ns	.386**	t(232)=1.21; ns
Hostilidad	.232**	.197*	t(232)=1.74; ns	.220**	t(232)=0.59; ns
Depresión	.519**	.487**	t(232)=1.8; ns	.500**	t(232)=1.07; ns
Autoconciencia	.362**	.326**	t(232)=1.86; ns	.339**	t(232)=1.19; ns
Impulsividad	.305**	.276**	t(232)=1.47; ns	.275**	t(232)=1.52; ns
Vulnerabilidad	.395**	.365**	t(232)=1.57; ns	.371**	t(232)=1.26; ns
Escalas del SCL90-R					
Som	.313*	.277**	t(232)=1.83; ns	.292**	t(232)=1.03; ns
Obs	.336**	.331**	t(232)=0.77; ns	.335**	t(232)=0.05; ns
SI	.278**	.270**	t(232)=0.4; ns	.275**	t(232)=0.15; ns
Dep	.357**	.351**	t(232)=0.31; ns	.351**	t(232)=0.30; ns
Ans	.332**	.329**	t(232)=0.15; ns	.316**	t(232)=0.79; ns
Hos	.212**	.182*	t(232)=1.48; ns	.196*	t(232)=0.77; ns
AF	.277**	.252**	t(232)=1.25; ns	.260**	t(232)=0.83; ns
IP	.231**	.240**	t(232)=-0.45; ns	.215**	t(232)=0.77; ns
Psi	.369**	.364**	t(232)=0.26; ns	.375**	t(232)=-0.30; ns
ISG	.373**	.360**	t(232)=0.67; ns	.363**	t(232)=0.50; ns
TSP	.342**	.322**	t(232)=1.02; ns	.332**	t(232)=0.50; ns
IMP	.266**	.251**	t(232)=0.75; ns	.242**	t(232)=1.17; ns

Autoría Propia

Nota. BHS18 = Versión de la Beck Hopelessness Scale con ítems calibrados; TAI9 = Versión adaptativa de formato fijo.; TAI.35 = Versión adaptativa con criterio mixto; BIN = Banco de Ítems de Neuroticismo; Som = Somatización; Obs = Obsesiones/Compulsiones; Dep = Depresión; SI = Sensitividad Interpersonal; Ans = Ansiedad; Hos = Hostilidad; AF = Ansiedad Fóbica; IP = Ideación Paranoide; Psi = Psicoticismo; ISG = Índice de Severidad Global; TSP = Total de Síntomas Positivos; IMP = Índice de Malestar Positivo. ** $p < .001$

Discusión y Conclusiones

Aun cuando la BHS es un instrumento desarrollado hace más de 40 años resulta innegable su vigencia considerando la numerosa cantidad de investigadores que adaptan la escala a sus culturas y actualizan periódicamente sus propiedades psicométricas. La mayoría de estos estudios instrumentales aportan evidencias de validez y confiabilidad a partir del análisis psicométrico con teoría clásica. Las aplicaciones de la TRI a la BHS, en cambio, son muy escasas (e.g. Aish & Wasserman, 2001; Balsamo et al., 2020; Abal et al., 2020; Toro Tobar, Avendaño Prieto & Castrillón, 2016; Young, Halper, Clark, Scheftner & Fawcett, 1992) y se circunscriben exclusivamente a la calibración de sus ítems. Pero las ventajas de la TRI más relevantes pueden apreciarse a través de sus aplicaciones, por ejemplo, el estudio del funcionamiento diferencial de los ítems o la construcción de TAIs (Abal, Lozzia, Aguerri, Galibert & Attorresi, 2010; Muñiz, 2018; Thomas, 2019).

En esta investigación se utilizó la TRI con el objetivo de encontrar una solución alternativa a las propuestas realizadas desde la teoría clásica para el problema de la extensión de la BHS. El análisis de los resultados obtenidos con ambas variantes de administración adaptativa corroboró que las correlaciones de los θ con la BHS completo fueron elevadas y alcanzaron el criterio de calidad más exigente ($r \geq .95$) definido por Thompson (2009) para estudios con TAIs.

El grado de precisión de las estimaciones de Desesperanza que tiene la BHS depende de la localización del sujeto en la dimensión del rasgo. Esta característica ya fue señalada por Young et al. (1992) y obedece a que la mayoría de los ítems muestran capacidad discriminativa en niveles medio-altos del rasgo. La concentración de los b en niveles elevados es frecuente en variables del contexto clínico para las que se registran variaciones significativas sólo en los puntajes altos mientras que los puntajes bajos se asocian a ausencia de manifestaciones del rasgo (Morris, Bass, Howard & Neapolitan, 2019; Reise & Revicki, 2015; Thomas, 2019). Claro está, las versiones adaptativas reproducen esta propiedad y, como consecuencia, los θ estimados para los niveles bajos del rasgo presentan un error más alto que los estimados para niveles elevados de Desesperanza.

La comparación de las versiones adaptativas ensayadas (longitud fija y criterio mixto) muestra que bajo ambas condiciones se alcanzaron niveles de calidad similares para predecir el puntaje que obtendría el sujeto si respondiera el test completo. Los nueve primeros reactivos presentados a cada persona fueron los más eficientes para la evaluación de la Desesperanza en tanto que los nueve subsiguientes no aportan información relevante para mejorar la precisión de la medida. La interrupción basada en el error de estimación menor o igual a .35 (criterio mixto) no brindó un ahorro considerable en los tiempos de administración como se esperaba. Sólo una escasa proporción de evaluados finalizó la aplicación con menos de nueve ítems por lo que su impacto parece irrelevante.

Otros TAI elaborados para la medición de rasgos de la personalidad han adoptado un error prefijado más laxo (e.g. error menor o igual a .50 equivalente a una confiabilidad clásica de .75) para garantizar un nivel de precisión mínimo y priorizar el ahorro de tiempo de administración (e.g. Abal et al., 2019; Nieto et al., 2017). Pero esta estrategia no parece una alternativa razonable porque propicia que la estimación de los niveles de θ en las versiones adaptativas se alejen de las que se obtendrían con la administración de la prueba completa. Además, se vería perjudicada justamente la precisión de los niveles altos del rasgo, para los que se tiene especial atención en pruebas de tamizaje. En con-

traposición, otros autores han diseñado TAIs proponiendo criterios de interrupción más exigentes, como un error de estimación de por lo menos .30 (e.g. Fonseca-Pedrero et al., 2013; Sunderland, et al., 2017). No obstante, la cantidad total de elementos que componen la versión original de la BHS es relativamente escasa si se compara con los bancos de ítems usados para alcanzar un grado de eficiencia tan elevado.

La metodología TAI aplicada aquí ha requerido la corroboración del supuesto de unidimensionalidad del constructo, en consonancia con los hallazgos reportados por otros autores que aplicaron TRI en la BHS (Aish & Wasserman, 2001; Balsamo et al., 2020). Young et al., 1992). Sin embargo, es notoria la heterogeneidad de resultados registrados en el análisis de la estructura factorial de la BHS. Citando solamente investigaciones realizadas con población no-clínica es posible encontrar estructuras con único factor (Steed, 2001; Tanaka et al., 1998) pero también con dos (Satorres et al., 2016), tres (Illiceto & Fino, 2015; Kocalevent et al., 2016) y cinco factores (Zhang et al., 2015). También se han ajustado modelos bifactoriales en el que se define un factor general de Desesperanza y se identifican dos factores específicos vinculados al sentido positivo o negativo de la redacción de los ítems (Boduszek & Dhingra, 2016; Kliem et al., 2018).

Esta falta de consenso sobre los aspectos internos que constituyen la Desesperanza complica una medición adaptativa que controle la cobertura de contenido de los ítems administrados a cada sujeto. La consolidación de una estructura interna con subfactores con mayor fundamentación empírica podría contribuir al desarrollo de un TAI multidimensional. Este tipo de TAIs ha mostrado una ganancia incremental con respecto a los TAIs unidimensionales en la precisión y eficiencia de las mediciones (e.g. Gibbons et al., 2016; Makransky et al., 2013; Paap et al., 2017). Más allá de eso, las correlaciones elevadas entre los θ estimados con los 18 ítems y sendas versiones adaptativas brindan evidencias de que las variaciones en el contenido no incidieron la medición del constructo. Tampoco variaron significativamente las asociaciones de la Desesperanza con las variables externas evaluadas por SCL-90-R y BIN al comparar los coeficientes de Pearson entre la versión completa y las adaptativas. En definitiva, las evidencias de validez de contenido y concurrente no se vieron afectadas de manera sustantiva por la medición adaptativa. A su vez, la variedad de modelos que aíslan subfactores en la Desesperanza no han tenido incidencia en la interpretación tradicional del puntaje que arroja el instrumento. En el plano aplicado siempre se ha validado e interpretado una única puntuación global de Desesperanza como la analizada en el presente estudio.

Una de las limitaciones de la presente investigación y que será motivo de futuros estudios es la validación de puntos de corte que permitan una valoración clínico-patológica de los niveles de Desesperanza. De esta manera se podrá avanzar en la obtención de índices de sensibilidad y especificidad requeridos para el uso de la BHS con fines de tamizaje (Balsamo et al., 2020). La ausencia de un criterio externo con probada evidencia para discriminar entre el grupo de riesgo y el grupo control imposibilita realizar el análisis con los datos de la presente investigación. Asimismo, se podrán ensayar versiones adaptativas con criterios de selección de ítems e interrupción de la aplicación centrados únicamente en determinar si el evaluado se localiza superando o no el punto de corte (Thompson, 2009). Aunque debe ser corroborado empíricamente, es probable que esta estrategia permita disminuir aún más la cantidad de ítems administrados porque el énfasis se coloca en la precisión para clasificar correctamente al sujeto como persona con o sin un nivel crítico de Desesperanza, en lugar de reducir el error en la estimación del nivel de θ (Barrada, 2012).

Para esta investigación se ha utilizado la adaptación local de la BHS en su formato original con ítems de respuesta dicotómica habida cuenta de la equivalencia lingüística y conceptual documentada.

No obstante, la implementación de formatos de respuesta politómica ha mostrado adecuadas propiedades psicométricas en versiones fijas de cuatro ítems analizadas en el marco de la TCT alcanzando correlaciones de .88 con las puntuaciones de la BHS de 20 ítems (Perczel Forintos et al., 2013; Yip & Cheung, 2006). En posteriores investigaciones se analizará desde la perspectiva de la TRI si la inclusión de más alternativas de respuestas puede mejorar la eficiencia de una medición adaptativa.

En conclusión, se ha demostrado que un TAI basado en los ítems de la BHS puede brindar estimaciones de los niveles de Desesperanza de los evaluados con un grado de precisión aceptable con sólo administrar nueve elementos de la escala. Esto supone un ahorro de tiempo considerable que podría ser aprovechado para la recolección de otros datos o simplemente para evitar la fatiga de los evaluados. El avance de la tecnología y el uso cada vez más frecuente de prácticas de evaluación mediadas por computadoras hace factible la implementación de una administración adaptativa para optimizar la medición de la Desesperanza.

Referencias

- Abal, F.J.P., Auné, S.E. y Attorresi, H.F. (2019). Construcción de un banco de ítems de Facetas de Neuroticismo para el desarrollo de un test adaptativo. *Psicod debate. Psicología, Cultura y Sociedad*, 19 (1), 31 - 50. doi: [10.18682/pd.v19i1.85](https://doi.org/10.18682/pd.v19i1.85).
- Abal, F. J. P., Lozzia, G. S., Aguerri, M. E., Galibert, M. S. & Attorresi, H. F. (2010). La escasa aplicación de la Teoría de Respuesta al Ítem en tests de ejecución típica. *Revista Colombiana de Psicología*, 19 (1), 111 - 122. doi: [10.35670/1667-4545.v19.n1.23877](https://doi.org/10.35670/1667-4545.v19.n1.23877)
- Abal, F. J. P., Lozzia, G. S., Auné, S. E., Attorresi, H. F. (2020). Item Banking for an adaptive measurement of Neuroticism. *Psychological Thought*, 13 (2), 459-485. doi: [10.37708/psyct.v13i2.503](https://doi.org/10.37708/psyct.v13i2.503)
- Achtyes, E. D., Halstead, S., Smart, L., Moore, T., Frank, E., et al. (2015). Validation of computerized adaptive testing in an outpatient non-academic setting: the VOCATIONS Trial. *Psychiatr. Serv.* 66, 1091-96. <https://doi.org/10.1176/appi.ps.201400390>
- Aish, A. M., & Wasserman, D. (2001). Does Beck's Hopelessness Scale really measure several components? *Psychological Medicine*, 31(02), 367-372. doi: <https://doi.org/10.1017/S0033291701003300>
- Alamo, C., Baader, T., Antúnez, Z., Bagladi, V. & Bejer, T. (2019). Escala de desesperanza de Beck como instrumento útil para detectar riesgo de suicidio en universitarios chilenos. *Revista chilena de neuro-psiquiatría*, 57(2), 167-175. doi: [10.4067/S0717-92272019000200167](https://doi.org/10.4067/S0717-92272019000200167).
- Aloba, O., Awe, O., Adelola, A., Olatunji, P., & Aloba, T. (2018). Psychometric Adaptation of the Beck Hopelessness Scale as a Self-Rated Suicide Risk Screening Instrument Among Nigerian University Students. *Journal of the American Psychiatric Nurses Association*, 24(5), 433-443. <https://doi.org/10.1177/1078390318762054>
- Aloba, O., Esan, O., & Alimi, T. (2017). Adaptation of the Beck Hopelessness Scale as a suicide risk screening tool among Nigerian patients with schizophrenia. *International Journal of Psychiatry in Clinical Practice*, 22(1), 19-24. doi: [10.1080/13651501.2017.1356928](https://doi.org/10.1080/13651501.2017.1356928)
- American Educational Research Association [AERA], American Psychological Association [APA] y National Council on Measurement in Education [NCME] (2014). *Standards for Educational and Psychological Testing*. Washington, DC: AERA
- American Psychological Association (2010). *Ethical Principles for Psychologists and Code of Conduct*. Washington, D.C.: APA
- Ato, M., López, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29 (3), 1038-1059. <https://doi.org/10.6018/analesps.29.3.178511>
- Baker, F. B. & Kim, S. H. (2017). *The Basics of Item Response Theory Using R*. Cham, Suiza: Springer International Publishing. DOI: [10.1007/978-3-319-54205-8](https://doi.org/10.1007/978-3-319-54205-8)
- Balsamo, M., Carlucci, L., Innamorati, M., Lester, D., & Pompili, M. (2020). Further Insights Into the Beck Hopelessness Scale (BHS): Unidimensionality Among Psychiatric Inpatients. *Frontiers in Psychiatry*, 11. doi: [10.3389/fpsy.2020.00727](https://doi.org/10.3389/fpsy.2020.00727).
- Barrada, J. R. (2012). Tests adaptativos informatizados: una perspectiva general. *Anales de Psicología*, 28 (1), 289-302. Obtenido de: <https://revistas.um.es/analesps/article/view/140792>.
- Beck, A. T., & Steer, R. A. (1989). Clinical predictors of eventual suicide: a 5- to 10-year prospective study of suicide attempters. *Journal of Affective Disorders*, 17(3), 203-209. DOI: [10.1016/0165-0327\(89\)90001-3](https://doi.org/10.1016/0165-0327(89)90001-3)
- Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology*, 42(6), 861-865. DOI: [10.1037/h0037562](https://doi.org/10.1037/h0037562)
- Beiser, D., Vu, M., & Gibbons, R. (2016). Test-Retest Reliability of a Computerized Adaptive Depression Screener. *Psychiatric Services in Advance*. <https://doi.org/10.1176/appi.ps.201500304>
- Boduszek, D. & Dhingra, K. (2016). Construct validity of the Beck Hopelessness Scale (BHS) among university students: A multitrait-multimethod approach. *Psychol Assess*, 28(10), 1325-1330. doi: [10.1037/pas0000245](https://doi.org/10.1037/pas0000245).
- Boffa, J. W., King, S. L., Turecki, G., & Schmidt, N. B. (2018). Investigating the role of hopelessness in the relationship between PTSD symptom change and suicidality. *Journal of Affective Disorders*, 225, 298-301. DOI: [10.1016/j.jad.2017.08.004](https://doi.org/10.1016/j.jad.2017.08.004)
- Bouvard, M., Charles, S., Guerin, J., & Aimard, G. (1992). Study of the Hopelessness Scale: Validation and factorial analysis. *Encephale*, 18(3), 237-240. Obtenido de: <https://europepmc.org/article/med/1299593>.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basics, concepts, applications, and programming*. New York: Routledge.
- Cai, L., Thissen, D., & Du Toit, S. (2011). *IRTPRO users guide*. Lincolnwood, IL: Scientific Software International.
- Chang, H. H. (2015). Psychometrics Behind Computerized Adaptive Testing. *Psychometrika*, 80 (1), 1-20. Recuperado de: <https://link.springer.com/article/10.1007/s11336-014-9401-5>
- Chen, W., & Thissen, D. (1997). Local dependence indices for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*, 22, 265-289. <https://doi.org/10.2307/1165285>
- Choi, S. W. (2018). Firestar. En W. J. van der Linden (Ed.). *Handbook of Item Response Theory, Volume 3: Applications* (pp. 547-556). Boca Raton: Chapman y Hall/CRC.
- De Beurs, D.P., de Vries, A.L., de Groot, M. H, et al. (2014). Applying computer adaptive testing to optimize online assessment of suicidal behavior: a simulation study. *Journal of Medical Internet Research*, 16:e207. doi: [10.2196/jmir.3511](https://doi.org/10.2196/jmir.3511)
- Derogatis, L. (1994). *SCL-90-R. Symptom Checklist-90-R. Administration, Scoring and Procedures Manual*. Minneapolis: National Computer System.
- Devine, J., Fliege, H., Kocalevent, R., Mierke, A., Klapp, B. F., & Rose, M. (2016). Evaluation of Computerized Adaptive Tests (CATs) for longitudinal monitoring of depression, anxiety, and stress reactions. *Journal of Affective Disorders*, 190, 846-853. <https://doi.org/10.1016/j.jad.2014.10.063>
- Dunn S. L., Olamijulo G. B., Fuglseth H. L., Holden T. P., Swieringa L. L., Sit M. J., Rieth N. P., & Tintle N. L. (2014). The State-Trait Hopelessness Scale: development and testing. *West J Nurs Res*. 36(4), 552-70. <https://doi.org/10.1177/0193945913507634>
- Flores, F., & Almanza, M. (2013). Un estudio exploratorio de las consecuencias psicosociales del estigma en varones que viven con VIH en México. *Revista Iberoamericana De Psicología*, 6(1), 51-60. <https://doi.org/10.33881/2027->

- Fonseca-Pedrero, E., Menéndez, L. F., Paino, M., Lemos-Giráldez, S., & Muñiz, J. (2013). Development of a Computerized Adaptive Test for Schizotypy Assessment. *PLoS ONE*, 8(9), 73-201. <https://doi.org/10.1371/journal.pone.0073201>
- Gibbons, R. D., Weiss, D. J., Frank, E., & Kupfer, D. (2016). Computerized Adaptive Diagnosis and Testing of Mental Health Disorders. *Annual Review of Clinical Psychology*, 12(1), 83-104. doi:10.1146/annurev-clinpsy-021815-093634. DOI: [10.1146/annurev-clinpsy-021815-093634](https://doi.org/10.1146/annurev-clinpsy-021815-093634)
- Gray Little, B., Williams, V. S. L. & Hancock, T. D. (1997). An Item Response Theory Analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, 23, 443-451. <https://doi.org/10.1177/0146167297235001>
- Gutiérrez Rincón, K., & Escobar Melo, H. (2013). Avance del dominio del inglés como lengua extranjera por medio de una estrategia didáctica efectiva. *Revista Iberoamericana De Psicología*, 6(1), 61-72. <https://doi.org/10.33881/2027->
- Hermosillo-de la Torre A. E., Méndez-Sánchez C., & Gonzalez-Betanzos F. (2020). Evidencias de validez factorial de la Escala de desesperanza de Beck en español con muestras clínicas y no clínicas. *Acta Colombiana de Psicología*, 23(2). <http://www.doi.org/10.14718/ACP.2020.23.2.7>
- Illiceto, E., & Fino, E. (2015). Beck Hopelessness Scale (BHS). A Second-Order Confirmatory Factor Analysis. *European Journal of Psychological Assessment*, 31(1), 31-37. <https://doi.org/10.1027/1015-5759/a000201>
- Innamorati, M., Lester, D., Balsamo, M., Erbutto, D., Ricci, F., Amore, M., & Pompili, M. (2013). Factor Validity of the Beck Hopelessness Scale in Italian Medical Patients. *Journal of Psychopathology and Behavioral Assessment*, 36(2), 300-307. <https://doi.org/10.1007/s10862-013-9380-3>
- International Test Commission (2014, april 10). ITC Statement on the Use of Tests and Other Assessment Instruments for Research Purposes. Obtenido de: <http://www.intestcom.org/page/22>
- Kang, T., & Chen, T. T. (2011). Performance of the generalized S-X2 item fit index for the graded response model. *Asia Pacific Education Review*, 12(1), 89-96. <https://doi.org/10.1007/s12564-010-9082-4>
- Kao, Y.C., Liu, Y.P., & Lu, C.W. (2012) Beck hopelessness scale: exploring its dimensionality in patients with schizophrenia. *Psychiatr Q.*, 83, 241-255. doi: <https://doi.org/10.1007/s11126-011-9196-9>
- Kim, S., Lee, E. H., Hwang, S. T., Hong, S. H., Lee, K., Kim, J. H. (2015). Reliability and Validity of the Korean Version of the Beck Hopelessness Scale. *J Korean Neuropsychiatr Assoc.*, 54 (1), 84-90. doi: <https://doi.org/10.4306/jknpa.2015.54.1.84>
- Kliem, S., Lohmann, A., Mößle, T., & Brähler, E. (2018). Psychometric properties and measurement invariance of the Beck hopelessness scale (BHS): results from a German representative population sample. *BMC Psychiatry*, 18(1). Doi: <https://doi.org/10.1186/s12888-018-1646-6>
- Kocalevent, R. D., Finck, C., Pérez Trujillo, M., Sautier, L., Zill, J., & Hinz, A. (2016). Standardization of the Beck Hopelessness Scale in the general population. *Journal of Mental Health*, 26(6), 516-522. DOI: [10.1080/09638237.2016.1244717](https://doi.org/10.1080/09638237.2016.1244717)
- Lester D. (2013). Hopelessness in undergraduate students around the world: A review. *J Affect Disord*, 150, 1204-8. DOI: [10.1016/j.jad.2013.04.055](https://doi.org/10.1016/j.jad.2013.04.055)
- Magis D., Yan D., & von Davier A. A. (2017). Computerized Adaptive and Multistage Testing with R. Use R! Cham: Springer.
- Magnée, T., de Beurs, D. P., Terluin, B., & Verhaak, P. F. (2017). Applying Computerized Adaptive Testing to the Four-Dimensional Symptom Questionnaire (4DSQ): A Simulation Study. *JMIR Ment Health*, 4(1), 7. doi:[10.2196/mental.6545](https://doi.org/10.2196/mental.6545).
- Makransky, G., Mortensen, E. L., & Glas, C. A. (2013). Improving personality facet scores with multidimensional computer adaptive testing: An illustration with the NEO PI-R. *Assessment*, 20, 3-13. doi: <https://doi.org/10.1177/1073191112437756>
- McCrae, R. R., & Costa P. T., Jr. (2010). *NEO Inventories professional manual*. Odessa, FL: Psychological Assessment Resources.
- McMillan, D., Gilbody, S., Beresford, E., & Neilly, L. (2007). Can we predict suicide and non-fatal self-harm with the Beck Hopelessness Scale? A meta-analysis. *Psychological Medicine*, 37(6), 769-78. doi:[10.1017/S0033291706009664](https://doi.org/10.1017/S0033291706009664)
- Mikulic, M. I., Cassullo, G.L., Crespi, M.C., & Marconi A. (2009). Escala de Desesperanza BHS (A. Beck, 1974): estudio de las propiedades psicométricas y baremización de la Adaptación Argentina. *Anu Investig* (16), 365-73. Obtenido de: <https://www.redalyc.org/articulo.oa?id=3691/369139945034>.
- Morris, S. B., Bass, M., Howard, E., & Neapolitan, R. E. (2019) Stopping Rules for Computer Adaptive Testing When Item Banks Have Nonuniform Information, *International Journal of Testing*. doi:[10.1080/15305058.2019.1635604](https://doi.org/10.1080/15305058.2019.1635604).
- Muñiz, J. (2018). *Introducción a la psicometría*. Madrid: Pirámide.
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide* (6th ed.). Los Angeles: Muthén & Muthén.
- Nieto, M. D., Abad, F. J., Hernández-Camacho, A., Garrido, L. E., Barrada, J. R., Aguado, D. & Olea, J. (2017). Calibrating a new item pool to adaptively assess the Big Five. *Psicothema*, 29(3), 390-395. Recuperado de: <http://www.psicothema.com/english/psicothema.asp?id=4411>
- Nissim, R., Flora, D. B., Cribbie, R. A., Zimmermann, C., Gagliese, L., & Rodin, G. (2010). Factor structure of the Beck Hopelessness Scale in individuals with advanced cancer. *Psycho-Oncology*, 19(3), 255-263. Recuperado de: <https://pubmed.ncbi.nlm.nih.gov/19274620/>
- Olea, J. & Ponsoda, V. (2013). *Tests adaptativos informatizados*. Madrid: Ediciones UNED.
- Paap, M.C.S., Kroeze, K.A., Terwee, C.B. et al. (2017). Item usage in a multidimensional computerized adaptive test (MCAT) measuring health-related quality of life. *Qual Life Res* 26, 2909-2918. doi: <https://doi.org/10.1007/s11136-017-1624-3>
- Parpa, E., Tsilika, E., Galanos, A., Nikoloudi, M., & Mystakidou, K. (2019). Depression as mediator and or moderator on the relationship between hopelessness and patients' desire for hastened death. *Supportive Care in Cancer*, 27(11), 4353-4358. doi: [10.1007/s00520-019-04715-2](https://doi.org/10.1007/s00520-019-04715-2)
- Pedrosa García, I. (2015). *Evaluación de la personalidad emprendedora mediante un test adaptativo informatizado*. Tesis doctoral, Universidad de Oviedo, España. Obtenido de: <http://hdl.handle.net/10651/33327>.
- Perczel Forintos, D., Rózsa, S., Pilling, J., & Kopp, M. (2013). Proposal for a Short Version of the Beck Hopelessness Scale Based on a National Representative Survey in Hungary. *Community Mental Health Journal*, 49(6), 822-830. DOI:[10.1007/s10597-013-9619-1](https://doi.org/10.1007/s10597-013-9619-1)
- Perczel Forintos, D., Sallai, J., & Rózsa, S. (2001). Psychometric analysis of the Beck Hopelessness Scale: A pilot study. *Psychiatria Hungarica*, 6, 632-643. doi: <https://doi.org/10.1007/s10597-013-9619-1>
- Pompili, M., Innamorati, M., Gonda, X., Serafini, G., Sarno, S., Erbutto, D., & Girardi, P. (2013). Affective temperaments and hopelessness as predictors of health and social functioning in mood disorder patients: A prospective follow-up study. *Journal of Affective Disorders*, 150(2), 216-222. <https://doi.org/10.1016/j.jad.2013.03.026>
- Quiñonez Tapia, F., Méndez Luevano, T., & Castañeda Camey, N. (2019). Análisis factorial confirmatorio y propiedades psicométricas de la Escala de Desesperanza de Beck en estudiantes en contextos de pobreza en México. *Revista de Psicopatología y Psicología Clínica*, 24(2), 59-70. <https://doi.org/10.5944/rppc.24104>

- Reise, S. P. & Revicki, D. A. (2015). Handbook of Item Response Theory Modeling Applications to Typical Performance Assessment. Nueva York: Routledge.
- Reise, S. P., & Rodriguez, A. (2016). Item response theory and the measurement of psychiatric constructs: some empirical and conceptual issues and challenges. *Psychological Medicine*, 46(10), 2025–2039. doi:<https://doi.org/10.1017/S0033291716000520>
- Sanchez Gonzalez, J. F., Attorresi, H. F. & Abal, F. J. P. (2020). Escala de Desesperanza de Beck analizada con teoría de respuesta al ítem. *Revista De Psicología*, 19 (1), 138-156. doi: [10.24215/2422572Xe054](https://doi.org/10.24215/2422572Xe054)
- Sánchez, R. O. & Ledesma, R. D. (2009). Análisis psicométrico del Inventario de Síntomas Revisado (SCL-90-r) en población clínica. *Revista Argentina de Clínica Psicológica*, XVIII, 265-274. Obtenido de: <http://www.clinicapsicologica.org.ar/>.
- Satorres, E., Ros, L., Meléndez, J. C., Serrano, J. P., Latorre, J. M., & Sales, A. (2016). Measuring elderly people's quality of life through the Beck Hopelessness Scale: a study with a Spanish sample. *Aging & Mental Health*, 22(2), 239–244. doi: <https://doi.org/10.1080/13607863.2016.1247427>
- Schalet, B. D., Cook, K. F., Choi, S. W., & Cella, D. (2014). Establishing a common metric for self-reported anxiety: Linking the MASQ, PANAS, and GAD-7 to PROMIS Anxiety. *Journal of Anxiety Disorders*, 28(1), 88–96. doi: <https://doi.org/10.1016/j.janxdis.2013.11.006>
- Spangenberg, L., Zenger, M., Garcia-Torres, F., Mueller, V., Reck, M., Mehnert, A., & Vehling, S. (2016). Dimensionality, Stability, and Validity of the Beck Hopelessness Scale in Cancer Patients Receiving Curative and Palliative Treatment. *Journal of Pain and Symptom Management*, 51(3), 615–622. Doi: <https://doi.org/10.1016/j.jpainsymman.2015.11.008>
- Steed, L. (2001). Further Validity and Reliability Evidence for Beck Hopelessness Scale Scores in a Nonclinical Sample. *Educational and Psychological Measurement*, 61(2), 303–316. doi: <https://psycnet.apa.org/doi/10.1177/00131640121971121>
- Steeg, S., Haigh, M., Webb, R. T., Kapur, N., Awenat, Y., Gooding, P., ... Cooper, J. (2016). The exacerbating influence of hopelessness on other known risk factors for repeat self-harm and suicide. *Journal of Affective Disorders*, 190, 522–528. <https://doi.org/10.1016/j.jad.2015.09.050>
- Stochl, J., Boehnke, J., Pickett, K. E., & Croudace, T. J. (2016). Computerized adaptive testing of population psychological distress: simulation-based evaluation of GHQ-30. *Social Psychiatry and Psychiatric Epidemiology*, 51(6), 895-906. doi: <https://doi.org/10.1007/s00127-015-1157-4>
- Sunderland, M., Slade, T., Krueger, R. F., Markon, K. E., Patrick, C. J., & Kramer, M. D. (2017). Efficiently measuring dimensions of the externalizing spectrum model: Development of the Externalizing Spectrum Inventory-Computerized Adaptive Test (ESI-CAT). *Psychological Assessment*, 29(7), 868–880. doi: <https://psycnet.apa.org/doi/10.1037/pas0000384>
- Tanaka, E., Sakamoto, S., Ono, Y., Fujihara, S. & Kitamura, T. (1998). Hopelessness in a Community Population: Factorial Structure and Psychosocial Correlates. *The Journal of Social Psychology*, 138 (5), 581–590. doi: <https://psycnet.apa.org/doi/10.1080/00224549809600413>
- Thomas, M. L. (2019). Advances in applications of item response theory to clinical assessment. *Psychological Assessment*, 31(12), 1442–1455. doi:[10.1037/pas0000597](https://doi.org/10.1037/pas0000597).
- Thompson, N. (2009). Ability estimation with item response theory. Minnesota: Assessment Systems Corporation.
- Thompson, N. A. & Weiss, D. A. (2011). A Framework for the Development of Computerized Adaptive Tests. *Practical Assessment, Research, and Evaluation*, 16(1). doi: <https://doi.org/10.7275/wqzt-9427>
- Toland, M. (2013). Practical guide to conducting an item response theory analysis. *The Journal of Early Adolescence*, 34(1), 120-151. doi: <https://doi.org/10.1177/0272431613511332>
- Toro Tobar, R. A., Avendaño Prieto, B. L., & Castrillon, D. A. (2016). Design and psychometric analysis of the Hopelessness and Suicide Ideation Inventory "IDIS". *International Journal of Psychological Research*, 9(1), 52-63. Obtenido de: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S2011-20842016000100006&lng=en&tlng=en.
- Walter, O. B. (2010). Adaptive Tests for measuring Anxiety and Depression. In W. J. van der Linden and A. W. G. Cees (Ed), *Elements of Adaptive Testing Statistics for Social and Behavioral Sciences* (pp.123-136). New York: Springer.
- Wolfe, K. L., Nakonezny, P. A., Owen, V. J., Rial, K. V., Moorehead, A. P., Kennard, B. D., & Emslie, G. J. (2019). Hopelessness as a Predictor of Suicide Ideation in Depressed Male and Female Adolescent Youth. *Suicide and Life-Threatening Behavior*, 49(1), 253-263. doi: <https://doi.org/10.1111/sltb.12428>
- Young, M. A., Halper, I. S., Clark, D. C., Scheftner, W., & Fawcett, J. (1992). An item-response theory evaluation of the Beck Hopelessness Scale. *Cognitive Therapy and Research*, 16(5), 579–587. doi: <https://doi.org/10.1007/BF01175143>
- Zhang, W.-C., Jia, C.-X., Hu, X., Qiu, H.-M., & Liu, X.-C. (2015). Beck Hopelessness Scale: Psychometric Properties Among Rural Chinese Suicide Attempters and Non-Attempters. *Death Studies*, 39(7), 442–446. doi: <https://doi.org/10.1080/07481187.2014.970300>