

Revisión sistemática de las propiedades psicométricas de la Escala de Ansiedad y Depresión Hospitalaria (HADS) en muestras mexicanas: ¿Debemos regresar al principio?

Violeta R. Castro-López¹ y Karina Franco-Paredes²

¹ Universidad de Guadalajara, Guadalajara, México

² Centro de Investigación en Riesgos y Calidad de Vida. Universidad de Guadalajara, Guadalajara, México

Systematic review of the psychometric properties of the Hospital Anxiety and Depression Scale (HADS) in Mexican samples: Should we go back to the beginning

Abstract: A systematic review was run in order to find validation studies on the Hospital Anxiety and Depression Scale (HADS) in Mexican samples. Specialized databases and referencia sections of some articles were searched to locate this evidence. We found 10 articles that analyzed the psychometric properties of the HADS. The analysis showed that in all of the studies reliability was adequate and generally a two-factor structure could be supported. Nevertheless, we identified five different versions of the HADS and did not find any information about the translation process. In conclusion, the adequate internal consistency of the HADS could be confirmed but regarding the two-factor structure and cut-off points evidence is not conclusive.

Key words: HADS; reliability; validity; psychometric properties; Mexico.

Resumen: El propósito de esta investigación fue realizar una revisión sistemática de los estudios de validación de las propiedades psicométricas de la Escala de Ansiedad y Depresión Hospitalaria (HADS) en muestras mexicanas. Se realizó una búsqueda en bases de datos especializadas y en las referencias de los artículos. Se encontraron 10 estudios en los que se analizaron las propiedades psicométricas de la HADS. El análisis mostró que en todas las investigaciones se evaluó la fiabilidad y ésta fue adecuada; en general, los estudios apoyan una estructura de dos factores. Sin embargo, se identificaron cinco versiones diferentes de la HADS y no se encontró información sobre cómo se realizó la traducción. En conclusión, se confirma que la consistencia interna de la HADS es adecuada, si bien los estudios apoyan la estructura de dos factores la evidencia no es concluyente, como tampoco lo es respecto a los puntos de corte.

Palabras clave: HADS; fiabilidad; validez; propiedades psicométricas; México.

Introducción

La Escala de Ansiedad y Depresión Hospitalaria (HADS, por sus siglas en inglés) fue creada por Zig-

mond y Snaith (1983) en Inglaterra, con el propósito de brindar a los médicos un instrumento breve que permitiera la detección de síntomas clínicamente significativos de ansiedad y depresión en pacientes que acuden a

Recibido: 03 de diciembre de 2019; aceptado: 15 de enero de 2021.

Correspondencia: Dra. Karina Franco Paredes, Centro Universitario del Sur, Universidad de Guadalajara. Av. Enrique Arreola Silva # 883, Col. Centro, Ciudad Guzmán, C.P. 49066, Jalisco, México. Correo-e: karina.franco@academicos.udg.mx

Agradecimientos: La autora principal de este trabajo agradece al Consejo Nacional de Ciencia y Tecnología (CONACYT) el financiamiento para estudios de posgrado, CVU: 971475.

ambientes hospitalarios, no psiquiátricos, aunque su uso se ha extendido a otras situaciones; de manera que su aplicación fuese rápida y de ser necesario canalizar a los pacientes para una evaluación completa. La HADS está conformada por 14 ítems, siete para ansiedad y siete para depresión, es una escala de autoinforme con formato de respuesta tipo *Likert* con cuatro opciones, con valores de cero (ausencia de sintomatología) a tres (mayor presencia de sintomatología), y con puntos de corte para cada factor.

La HADS es reconocida como un instrumento con adecuada calidad métrica, por ello a través de los años fue adaptada en diferentes poblaciones, idiomas y países. Sin embargo, se ha señalado que las puntuaciones podrían variar entre las culturas por lo que es necesario realizar estudios sobre su validez para todas las versiones traducidas. Se tiene conocimiento de que hasta el momento se han realizado más de 60 estudios sobre las propiedades psicométricas de las puntuaciones de la HADS y nueve estudios de revisión en los que se analizó la evidencia disponible sobre este instrumento, consideraron tanto sus propiedades psicométricas como los resultados clínicos generados a partir de su uso (Bjelland et al., 2002; Brennan et al., 2010; Carey et al., 2012; Cosco et al., 2012; Herrmann, 1997; Maters et al., 2013; Mitchell et al., 2010; Morse et al., 2005; Terol-Cantero et al., 2015). En general, los resultados de las revisiones sobre la HADS apoyan su uso como un instrumento útil tanto para identificar la presencia de los síntomas de ansiedad y depresión, como la gravedad de los mismos en personas con una amplia variedad de enfermedades físicas o mentales (e.g. enfermedad coronaria, parkinson, cáncer, epilepsia, accidente cerebrovascular, enfermedad renal crónica, enfermedad cardiovascular, trastornos de ansiedad, trastorno depresivo, pacientes psiquiátricos y cuidadores primarios, etc.), así como en población sana.

Sin embargo, tanto para la estructura factorial como para los puntos de corte de la HADS los resultados de las investigaciones son inconsistentes. Hasta el momento no existe evidencia concluyente respecto la estructura factorial de la HADS. En dos revisiones sistemáticas se apoya la estructura de dos factores (Bjelland et al., 2002; Herrmann, 1997); no obstante, también existe evidencia de una estructura unidimensional, de tres factores (Cosco et al., 2012; Terol-Cantero et al., 2015) y hasta cuatro factores (Cosco et al., 2012); y también se ha propuesto un factor general denominado *distrés psicológico* (Norton et al., 2013). Asimismo, se han encontrado inconsistencias respecto a los ítems de la HADS, por ejemplo, en diferentes investigaciones los ítems 3, 7, 8, 9 y 10 presentan cargas factoriales bajas (menores a .40 ó .30),

presentan cargas anómalas (mayores a .40) en el factor al que no corresponde, o bien, cargan en los dos factores que se generan. Una explicación sobre los hallazgos inconsistentes o variados respecto a la estructura factorial ha sido el tipo de análisis utilizado. Por ejemplo, para identificar la estructura factorial se ha realizado análisis factorial exploratorio (AFE), análisis factorial confirmatorio (AFC), o bien, análisis Rasch. Se ha identificado que la estructura factorial obtenida depende tanto de la técnica estadística utilizada en la investigaciones como del tipo de muestra con la que se ha trabajado (Cosco et al., 2012; Terol-Cantero et al., 2015). Se sabe que la fuerza y la solidez de la evidencia proporcionada por cada una de las técnicas estadísticas es diferente y que cuando se utilizan modelos robustos (AFC o análisis Rasch) se tiene mayor precisión en los resultados (Quiñonez-Tapia et al., 2019), por lo que los hallazgos entre investigaciones no son directamente comparables. En general, cuando se ha utilizado AFE se obtiene una estructura de dos factores y cuando se realizan AFC se generan modelos de dos o tres factores, o bien modelos bifactoriales; en tanto que utilizando análisis Rasch la evidencia apoya la unidimensionalidad de la escala. Por tanto, se sugiere que las investigaciones futuras deben concentrarse en los análisis estadísticos más sólidos y evaluar si el funcionamiento de la HADS es similar en las diferentes muestras.

Respecto a los puntos de corte de la HADS también existen variaciones importantes en los resultados de los estudios, para el caso del factor de Ansiedad el rango es de ≥ 3 a ≥ 11 y de ≥ 4 a ≥ 11 para el factor Depresión, en tanto que para la puntuación total el rango es de ≥ 10 a ≥ 17 ; obteniendo valores de sensibilidad y especificidad superiores a .70 en la mayoría de los estudios; y se ha identificado que en algunos casos no se presentan los datos respecto a la sensibilidad y especificidad (Bjelland et al., 2002; Morse et al., 2005; Terol-Cantero et al., 2015; Vordermaier y Millman, 2011). Las inconsistencias respecto a los puntos de corte propuestos por Zigmond y Snaith (1983) pueden explicarse por las diferencias en los instrumentos utilizados como criterio o las características de la muestra, por ejemplo, tipo de enfermedad, edad, género, etc. (Hinz y Brähler, 2011; Terol-Cantero et al., 2015).

También se ha advertido que las inconsistencias sobre las propiedades psicométricas de la HADS, pueden deberse no sólo a los análisis estadísticos realizados, pueden tener su origen en el proceso de adaptación de la escala a los diferentes países. Maters et al. (2013) realizaron una revisión sistemática de la literatura con el propósito de analizar si los problemas transculturales y de traducción pueden ser subyacentes a los problemas reportados en la literatura sobre las propiedades psicomé-

tricas de la HADS. Revisaron 417 artículos realizados en países donde el inglés no es la lengua materna y en los que se utilizó la HADS. Encontraron que solo en 45% de los artículos revisados se declaró que se utilizó una versión traducida de la HADS y se indicó el idioma. Asimismo, 46% no citó el estudio de validación de la escala en el idioma de su país, incluso 13% ni siquiera citaron a Zigmond y Snaith, autores de la HADS. También encontraron que en las revisiones sobre la HADS no se consideró que se incluyen estudios con traducciones a diferentes idiomas y es poco común que se mencionen o analicen las implicaciones del proceso de traducción de la escala. Por tanto, para analizar las evidencias generadas sobre las propiedades psicométricas de la escala es necesario considerar la versión traducida con la que se trabajó e incluso el proceso de traducción.

Este hecho llamó la atención de las autoras de la presente investigación debido a que al buscar información sobre cómo fue realizado el proceso de traducción de la HADS en diferentes estudios realizados con muestras clínicas mexicanas no fue posible encontrar información clara al respecto; asimismo se identificaron imprecisiones en la información sobre las propiedades psicométricas de la escala. Por ello, el propósito de este estudio fue realizar una revisión sistemática de los estudios de validación de las puntuaciones de la HADS en muestras mexicanas.

Método

Estrategia de búsqueda

Se empleó la declaración PRISMA (Urrútia y Bonfill, 2010) para realizar revisiones sistemáticas. Las dos autoras llevaron a cabo de manera independiente la búsqueda en las bases de datos PUBMED, EBSCO, Science Direct, Redalyc, Scielo, REMERI e Imbiomed, así como en el buscador Google Académico. Adicionalmente, se realizó una búsqueda manual, la cual consistió en revisar el listado de referencias de los estudios para identificar otras investigaciones sobre las propiedades psicométricas de la HADS. Para la búsqueda se utilizaron las siguientes palabras clave en español: validación, adaptación, ansiedad, depresión, HADS, HAD, Escala de Depresión y Ansiedad Hospitalaria, México; y las palabras clave en inglés fueron: *validation, adaptation, anxiety, depression, hospital anxiety and depression scale and Mexico*. También se emplearon los operadores booleanos (*AND, OR*). La búsqueda se realizó en el mes de mayo de 2019 y en el mes de octubre del mismo año se realizó una segunda búsqueda manual sólo por una de las investigadoras. La búsqueda incluyó el periodo de 1983 (año de publicación del artículo original de la HADS) al 2019.

Criterios de inclusión y exclusión

Se establecieron los siguientes criterios de elegibilidad para los artículos: estudios cuantitativos, participantes mexicanos saludables y/o en condición clínica, estudios de adaptación y/o validación de la HADS en muestras mexicanas, estudios publicados en idioma inglés o español y artículos con acceso al texto completo. El criterio de exclusión fue: artículos que no correspondían a investigaciones empíricas.

Extracción de la información

Para cada artículo se registró la siguiente información en una base de datos de Excel: autores, año, tamaño de la muestra, género y edad de los participantes, resultados de fiabilidad (consistencia interna y test-retest) y evidencias de validez (análisis factorial, número de factores, varianza explicada, punto de corte, sensibilidad y especificidad, y otras evidencia de validez). La información fue registrada de manera independiente por las dos autoras y posteriormente fue revisada. Asimismo, se realizó una revisión de los ítems que aparecían en los artículos.

Resultados

Como resultado de la búsqueda se encontraron un total de 2710 estudios en las bases de datos. Se identificaron dos investigaciones más mediante la primera búsqueda manual y dos más en la segunda búsqueda manual. Fueron eliminados 2616 estudios duplicados, se excluyeron 88 por no cumplir con los criterios de inclusión; siendo evaluados a texto completo e incluidos a síntesis cuantitativa en total 10 estudios (véase Figura 1).

El análisis de los estudios mostró que fueron publicados en el periodo 2002 a 2018. En primer lugar se analizó la información disponible respecto al proceso de traducción y adaptación de la escala y se encontró que en ningún caso fue descrito claramente; solo en tres de los 10 estudios (30%) se encontraron detalles generales sobre el proceso de adaptación de la escala (Galindo, Meneses et al., 2015; Herrera et al., 2018; López-Alvarenga et al., 2002). Por ejemplo, López-Alvarenga et al. (2002) indicaron que debido a que no hubo concordancia entre la HADS y el diagnóstico se realizaron cambios en la estructura idiomática de las preguntas (no se especificaron los cambios); Galindo, Meneses et al. (2015) aplicaron la escala a 20 cuidadores primarios para evaluar la comprensión de cada ítem, instrucciones y opciones de respuesta; y para identificar palabras confusas u ofensivas. Indicaron que se realizaron cambios menores, sin especificar en qué consistieron. Herrera et al. (2018) señalaron

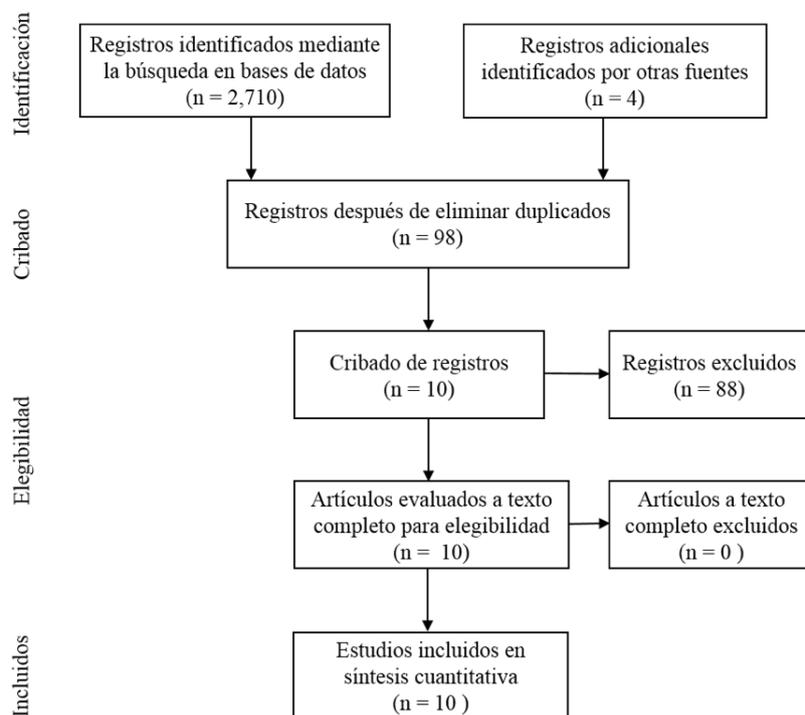


Figura 1. Diagrama PRISMA de búsqueda y selección de artículos.

que la adaptación de la escala se llevó a cabo con una muestra de 20 pacientes. Considerando los lineamientos internacionales para la traducción y estandarización de instrumentos de Koller et al. (2007), analizaron cómo estaban formuladas las preguntas o si resultaban ofensivas; así como la comprensión de las instrucciones y las opciones de respuesta.

La revisión de los ítems incluidos en los artículos mostró que existen cinco versiones diferentes de la HADS, en algunos casos se realizaron cambios importantes en su contenido, tanto de los ítems como de las opciones de respuesta al momento de la traducción.

A continuación se presentan los cinco ítems con variaciones de redacción en las versiones de la HADS revisadas en este estudio: Para el ítem 2 (*I still enjoy the things I used to enjoy*) se identificaron cuatro traducciones: todavía disfruto con lo que antes me gustaba, sigo disfrutando de las cosas de siempre, todavía disfruto de las cosas que suelen gustarme y sigo disfrutando con las mismas cosas de siempre. Para el ítem 8 (*I feel as if I am slowed down*) las dos traducciones identificadas fueron: me siento como si cada vez estuviera más lento, me siento lento y torpe. Para el ítem 9 (*I get a sort of frightened feeling like 'butterflies' in the stomach*), se identificaron tres traducciones: experimento una desagradable sensación de nervios y vacío en el estómago, experimento una desagradable sensación de «nervios y hormi-

guelos» en el estómago y tengo una sensación extraña, como de «aleteo» en el estómago. Para el ítem 12 (*I look forward with enjoyment to things*), se encontraron cuatro traducciones: me siento optimista respecto al futuro, espero las cosas con ilusión, veo las cosas con esperanza. Para el ítem 13 (*I get sudden feelings of panic*) las cuatro traducciones identificadas fueron: me asaltan sentimientos repentinos de pánico, experimento de repente sensaciones de gran angustia o temor, de repente siento pánico, experimento de repente una sensación de gran angustia o temor.

En la Tabla 1 se presentan las propiedades psicométricas evaluadas en cada estudio, así como las características de la muestra. En 90% de las investigaciones se trabajó con una muestra de personas con enfermedades físicas o mentales (e.g. cáncer, virus de inmunodeficiencia humana, enfermedad cardiovascular, asma, enfermedad renal crónica; obesidad; trastornos alimentarios) y sólo un estudio se realizó con una muestra de cuidadores primarios informales de pacientes oncológicos (Galindo, Meneses et al., 2015). El tamaño de la muestra fluctuó entre 75 y 400 participantes, en todos los estudios se incluyeron tanto mujeres ($n = 15 - 305$) como varones ($n = 17 - 174$) y en 70% de los estudios la muestra fue predominantemente de mujeres. La edad promedio mínima fue de 22 años y la máxima de 59 años, y en dos estudios no se informó la edad de los participantes.

Tabla 1. Características de los estudios analizados.

Autores, año	Población	Total <i>n</i>	Mujeres <i>n</i>	Varones <i>n</i>	Edad <i>M (DT)</i>	Análisis factorial	Otra evidencia de validez
López-Alvarenga et al. (2002)	Adultos con obesidad	75	58	17	39.70 (11.50)	No	Efectos de la intervención.
Noguera et al. (2013)	Pacientes mayores de 18 años que iniciaron tratamiento con antiretroviral	150	15	135	No se especifica	Sí	No
Galindo, Benjet et al. (2015)	Pacientes oncológicos	400	226	174	No se especifica	Sí	Correlación con otras pruebas
Galindo, Meneses et al. (2015)	Cuidadores primarios informales de pacientes oncológicos	200	131	69	47.40	Sí	Correlación con otras pruebas
Anguiano et al. (2017)	Pacientes hospitalizados en los servicios de medicina interna y hematología	111	63	48	45.40 (15.20)	Sí	No
Barriguete et al. (2017)	Personas con trastorno alimentario	325	305	20	22.00(9.70)	Sí	No
Yamamoto et al. (2018)	Pacientes con enfermedad inflamatoria intestinal	112	66	46	44.4 (15.00)	Sí	No
Herrera et al. (2018)	Pacientes con enfermedades cardiovasculares	200	82	118	59.70	Sí	Correlación con otras pruebas
Guillén-Díaz-Barriga et al. (2018)	Pacientes con asma	75	58	17	43.95 (14.68)	Sí	Correlación con otras pruebas
Lugo et al. (2018)	Pacientes con enfermedad renal crónica atendidos en un centro de hemodiálisis	253	122	131	50.59 (15.38)	Sí	Discriminación entre grupos

Respecto a las propiedades psicométricas de la HADS se encontró que en todos los estudios se evaluó la consistencia interna; sólo en un estudio se identificaron puntos de corte (López-Alvarenga et al., 2002) y en 90% de los estudios se analizó la estructura factorial de la HADS (Anguiano et al., 2017; Barriguete et al., 2017; Galindo, Benjet et al., 2015; Galindo, Meneses et al., 2015; Guillén-Díaz-Barriga et al., 2018; Herrera et al., 2018; Lugo et al., 2018; Noguera et al., 2013; Yamamoto-Furusho et al., 2018). Otras evidencias de validez generadas en los estudios se obtuvieron al correlacionar las puntuaciones de la HADS con otras pruebas. Así por ejemplo, se utilizó: el Inventario de Ansiedad de Beck (Beck et al., 1961) e Inventario de Depresión de Beck (Beck et al., 1988) en el estudio de Galindo, Benjet et al. (2015), la Escala de Evaluación de Desgaste de Zarit para Cuidadores Primarios (Zarit et al., 1980) en la investigación realizada por Galindo, Meneses et al. (2015), el Cuestionario de Calidad de Vida para personas con

asma (Juniper et al., 1999) y el Cuestionario de Control de Asma (Nathan et al., 2004) en el estudio de Guillén-Díaz-Barriga et al. (2018), y la Escala de Retiro de Patrón de Conducta tipo A (Rodríguez et al., 1996) en la validación de Herrera et al. (2018).

Los resultados específicos sobre la fiabilidad se presentan en la Tabla 2. En todos los estudios se evaluó la consistencia interna de la HADS con el coeficiente Alfa de Cronbach, para la puntuación total el rango fue de .80 a .94, para el factor Ansiedad fue de .78 a .87 y para Depresión fue de .70 a .87. Es importante señalar que en cuatro estudios se analizó la consistencia interna para 11, 12 ó 13 ítems (Barriguete et al., 2017; Galindo, Benjet et al., 2015; Lugo et al., 2018; Noguera et al., 2013), no para los 14 ítems que originalmente conforman el instrumento, debido a que los autores reportaron haber eliminado algunos. Las razones por las cuales eliminaron ítems fueron diversas, tales como que el ítem obtuvo cargas factoriales similares en ambos factores (Noguera

et al., 2013), los ítems se agruparon en un tercer factor sin corresponder conceptualmente a las subescalas originales, y se eliminaron para mantener la estructura de dos factores (Barriguete et al., 2017; Galindo, Benjet et al., 2015), también se eliminaron ítems que no cumplieron

con el criterio de carga factorial mayor a .40; además, de tener cargas factoriales similares en ambos factores de la HADS (Lugo et al., 2018). En ningún estudio se analizó la consistencia interna con el coeficiente Omega y tampoco se analizó la fiabilidad test-retest.

Tabla 2. Fiabilidad de la HADS (coeficiente alfa y número de ítems)

Autores, año	HADS-A	HADS-D	HADS-Total	Número ítems
López-Alvarenga et al. (2002)	.84	.70	.84	14
Noguera et al. (2013)	.78	.85	–	13
Galindo, Benjet et al. (2015)	.79	.80	.86	12
Galindo, Meneses et al. (2015)	.82	.78	.83	14
Anguiano et al. (2017)	–	–	.90	14
Barriguete et al. (2017)	.81	.87	.88	12
Yamamoto et al. (2018)	.81	.82	.88	14
Herrera et al. (2018)	–	–	.94	14
Guillén-Díaz-Barriga et al. (2018)	.87	.81	.91	14
Lugo et al. (2018)	.79	.71	.80	11

La estructura factorial de la escala (véase Tabla 3) fue evaluada en nueve estudios, de los cuáles en siete se informó que se realizó mediante un análisis de componentes principales (ACP), en dos casos se especifica que se realizó con el método de máxima verosimilitud; en seis estudios se empleó la rotación varimax y en otro la rotación oblimin. Solo un estudio reportó haber realizado análisis factorial exploratorio (Lugo et al., 2018). En cinco estudios que realizaron ACP y en el estudio de AFE se identificó una estructura de dos factores, como en la versión original de la HADS (varianza explicada = 42.58 a 56.3%), y en tres estudios realizados con ACP se encontró una estructura unidimensional (varianza explicada = 42.96 a 59.6%). Sin embargo, debe considerarse que 50% de los estudios se trabajó con muestras menores a 200 participantes (Anguiano et al., 2017; Guillén-Díaz-Barriga et al., 2018; López-Alvarenga et al., 2002; Noguera et al., 2013; Yamamoto-Furusho et al., 2018).

Sólo en un estudio de los diez revisados se analizaron los puntos de corte. López-Alvarenga et al. (2002) realizaron un estudio para determinar la reproducibilidad y exactitud de la HADS en una muestra de pacientes mexicanos con obesidad. Se aplicó una entrevista estructurada basada en los criterios diagnósticos del Episodio de Depresión Mayor y del Trastorno de Ansiedad Generalizada del DSM-IV, aleatorizando la aplicación en la secuencia de los procedimientos, entrevista y aplicación

de la HADS; y los síntomas de ansiedad y depresión fueron considerados como independientes. Derivaron puntos de corte para cada dimensión (Ansiedad = 8 y Depresión = 7) y proporcionan datos de sensibilidad (84% para ambos factores) y especificidad (84% para Ansiedad y 87% para Depresión).

Discusión

El propósito de este estudio fue realizar una revisión sistemática de los estudios de validación de las puntuaciones de la HADS en muestras mexicanas. El análisis de los diez artículos encontrados mostró que las propiedades psicométricas de la HADS se evaluaron principalmente en muestras de personas con enfermedades físicas o mentales (cáncer, virus de inmunodeficiencia humana, enfermedad renal crónica, trastornos alimentarios, obesidad, enfermedad cardiovascular, etc). En general, los estudios han proporcionado evidencia favorable respecto a la fiabilidad de las puntuaciones de la HADS. Por ejemplo, la consistencia interna, evaluada con el coeficiente Alfa de Cronbach, fue superior a .80 en la mayoría de los casos, tanto para la puntuación total como para los factores. Al respecto se ha sugerido que valores mayores a .70 son aceptables (Nunally y Bernstein, 1995).

La estructura factorial de la HADS se evaluó en nueve investigaciones, sin embargo, cabe señalar que en

Tabla 3. Análisis factorial de la HADS.

Autores, año, muestra	Método de análisis para factores	Número de factores	Factor con mayor varianza explicada (%)	% Varianza total
Noguera et al. (2013) <i>n</i> = 150 pacientes con tratamiento retroviral	Análisis de componentes principales con rotación varimax	Uno	Sólo se encontró un factor	54.65
Galindo, Benjet et al. (2015) <i>n</i> = 400 pacientes oncológicos	Análisis de componentes principales con rotación varimax	Dos	36.55 (Ansiedad)	48.04
Galindo, Meneses et al. (2015) <i>n</i> = 200 cuidadores primarios	Análisis de componentes principales con rotación varimax	Dos	32.51 (Ansiedad)	42.58
Anguiano et al. (2017) <i>n</i> = 111 pacientes hospitalizados	Análisis de máxima verosimilitud	Dos	No se indica	47.14
Barriguete et al. (2017) <i>n</i> = 325 pacientes con trastornos alimentario	Análisis de componentes principales con rotación oblimin	Dos	44.2 (Depresión)	56.30
Yamamoto et al. (2018) <i>n</i> = 112 pacientes con enfermedad infamatoria intestinal	Análisis de componentes principales con rotación varimax	Dos	40.22 (Depresión)	50.10
Herrera et al. (2018) <i>n</i> = 400 pacientes con enfermedad cardiaca	Análisis de componentes principales con rotación varimax	Uno	Sólo se encontró un factor	59.60
Guillén-Díaz-Barriga et al. (2018) <i>n</i> = 75 pacientes con asma	Análisis de componentes principales	Uno	Sólo se encontró un factor	42.96
Lugo et al. (2018) <i>n</i> = 253 pacientes con enfermedad renal	Análisis factorial exploratorio con método de máxima verosimilitud y rotación de varimax	Dos	35.6 (Ansiedad)	49.99

50% de los estudios el análisis se trabajó con muestras de tamaño menor al deseable ($n = 75$ a 150). Los expertos en psicometría han sugerido que para realizar análisis factoriales es recomendable al menos una muestra de 200 personas (Lloret-Segura et al., 2014; Mazor et al., 1992); asimismo, se señala que cuando se tienen muestras pequeñas (e.g. entre 50 y 100 participantes), es posible trabajar con otros análisis tales como el «*conditional item p value comparisons y delta plots*» (Muñiz, et al., 2001).

Considerando los estudios en los que se trabajó con una muestra mayor a 200 se encontró que en todos los casos se obtuvieron dos factores (Barriguete et al., 2017; Galindo, Benjet et al., 2015; Galindo, Meneses et al., 2015; Lugo et al., 2018), excepto en el estudio de Herrera et al. (2018) quienes encontraron un factor. Estos datos coinciden con la evidencia generada en las investigaciones internacionales debido a que se han identificado diferentes estructuras factoriales, siendo mayor la evidencia respecto a la estructura de dos factores (Bjelland et al., 2002; Herrmann, 1997), tal como se propuso en el estudio original (Zigmond, y Snaith, 1983).

En la mayoría de los estudios revisados, los resultados son producto de haber aplicado un ACP para analizar la estructura factorial de la HADS, aún cuando este análisis ha sido identificado desde hace años como obsoleto e incluso inadecuado si el objetivo es identificar el número de factores comunes y su composición, necesarios para explicar la varianza común de los ítems analizados. Aplicar el ACP como método para estimar factores conlleva ignorar el error en la medida, pudiendo contribuir a que aumenten las cargas factoriales de manera espuria, así como el porcentaje de varianza explicado por los factores, y la dimensionalidad del conjunto de ítems puede ser sobreestimada. Es inapropiado aplicar el ACP para la reducción de dimensiones y número de ítems (Lloret-Segura et al., 2014).

Es necesario considerar que solo en uno de los estudios revisados fue reportada la realización de AFE. Además, se identificó que en dos estudios se reportó la aplicación de un AFE mediante ACP (Galindo, Benjet et al., 2015; Yamamoto et al., 2018) evidenciando lo discutido por Lloret-Segura et al. (2014) que a pesar de la diferencia conceptual y formal entre los dos análisis, durante

décadas ha imperado el confundirlos, de manera que en muchas investigaciones el ACP se ha empleado con los propósitos que corresponden al AFE.

La HADS es una escala que evalúa dos constructos antiguos y comunes en psicología, con base en las más de dos décadas en que la escala ha sido empleada, y la amplia cantidad de estudios empíricos donde ha sido aplicada a nivel mundial, así como las revisiones para estudiarla, hay evidencia suficiente de las variables abordadas por la escala como para probar su estructura a través de confirmar la teoría que subyace a dichas variables, por lo que el AFC y el análisis Rasch se posicionan como opciones adecuadas y vigentes a emplear para analizar la estructura factorial de la escala.

Llamó la atención que sólo en un estudio fueron analizados los puntos de corte y se obtuvieron adecuados niveles de sensibilidad y especificidad (López-Alvarenga et al., 2002). Considerando que la HADS se utiliza ampliamente en el trabajo con muestras clínicas para discriminar casos, es importante generar evidencia respecto a los puntos de corte en cada muestra. Sin embargo, el análisis de los puntos de corte es el aspecto más cuestionado debido a la variedad de resultados obtenidos en las investigaciones, lo cual se ha documentado en diversas revisiones (Bjelland et al., 2002; Brennan et al., 2010; Hinz, y Brähler, 2011; Morse et al., 2005) por lo que queda pendiente la respuesta a la pregunta planteada por Terol-Cantero et al. (2015) ¿debemos usar puntos de corte diferentes en función de la muestra de estudio?

Un hecho que no permite mayor análisis y comparación de los resultados de los estudios sobre las propiedades psicométricas de las puntuaciones de la HADS en muestras mexicanas es que se identificó que se trabajó con diferentes traducciones de la escala. La revisión de los ítems presentados en los artículos mostró que en México se utilizan cinco diferentes versiones de la HADS y en algunos estudios no es posible saber con cuál versión se trabajó. Este hallazgo coincide con el encontrado por Maters et al. (2013) quienes identificaron cuatro versiones holandesas diferentes de la HADS con cambios importantes tanto en el contenido de los ítems como en las opciones de respuesta.

En la presente investigación no se encontró información sobre el proceso seguido para la traducción del instrumento. Esta situación no permite saber si en las traducciones se siguieron los lineamientos establecidos por organismos internacionales para garantizar que el test adaptado cumpla con los criterios legales, éticos y psicométricos (Elosua, 2017); por lo que es posible que existan debilidades importantes en el proceso de traducción, lo cual afecta los resultados obtenidos y limita su interpretación (Maters et al., 2013). En las investigaciones

analizadas tampoco se prestó atención a los ítems que pueden resultar críticos si se realiza una traducción literal. Por ejemplo, existen tres diferentes traducciones para el ítem 9 de la escala «*I get a sort of frightened feeling like 'butterflies' in the stomach*»; este es uno de los ítems con mayor cambio en las traducciones y a su vez de los que más inconsistencias ha presentado en los análisis estadísticos. Es importante considerar que una traducción literal no garantiza la equivalencia de los ítems con la versión original, es necesario realizar una adaptación de los significados (equivalencia semántica), considerando las diferencias culturales y lingüísticas para lograr que la versión adaptada sea lingüísticamente correcta, culturalmente adaptada y con propiedades psicométricas adecuadas (Elosua et al., 2014).

En las últimas décadas el nivel de exigencia científica, profesional y ética respecto al uso de los tests ha evolucionado de manera importante. Los avances en el campo de la psicología en general y de la investigación psicométrica en particular han resultado en la disponibilidad de modelos rigurosos para la construcción, traducción y administración de los tests (Elosua, 2017). Simultáneamente, se han impulsado proyectos para la elaboración de normas y herramientas que favorezcan prácticas correctas en el uso de los test. Este trabajo colaborativo ha derivado en la constitución de organismos nacionales e internacionales. Los organismos internacionales con mayor impacto en el establecimiento de directrices, que tienen como propósito regular la construcción, adaptación y el empleo de los tests en diferentes poblaciones, son la American Psychological Association, el National Council on Measurement in Education, y la International Test Commission.

El objetivo principal de las directrices para la traducción y adaptación de los tests es que el producto final del proceso de adaptación tenga el máximo nivel posible de equivalencia lingüística, cultural, conceptual y métrica respecto a la prueba original. Por ello, las directrices deben considerarse como una guía que los investigadores y profesionales deben seguir cuando se enfrentan al hecho de realizar la traducción y adaptación del instrumento de su interés (Muñiz et al., 2013). Sin embargo, las adaptaciones culturales adecuadas de los instrumentos no se logran fácilmente debido a que los creadores de los instrumentos no suelen anticiparse a los problemas que puede presentarse al momento de la traducción y adaptación, por tanto en ocasiones es difícil garantizar que el producto final de la traducción sea conceptualmente equivalente a la versión original (Breugelmans, 2009).

En el caso de la HADS se ha señalado que los autores optaron por utilizar un lenguaje coloquial para la construcción de los ítems, hecho que ha dificultado la traduc-

ción de algunos ítems porque en algunas culturas su contenido no tiene sentido; por ejemplo, cuando se preguntó sobre el proceso de traducción de la HADS a una muestra de investigadores, respondieron que habían tenido muchas complicaciones para la traducción porque el contenido de algunos ítems era extraño o resultaba divertido para las personas que hablaban árabe (Maters et al., 2013). Por tanto, es necesario hacer una traducción y adaptación muy cuidadosa considerando las pautas y procedimientos existentes, para garantizar que los instrumentos están midiendo lo mismo en los diferentes países o culturas. El cambio constante en las situaciones sociales y la evolución de los fenómenos psicológicos, son algunas de las razones para llevar a cabo estudios de validación y de análisis de propiedades psicométricas de los tests; no obstante, la realización de estos análisis o la publicación de ellos en un artículo no son garantía del funcionamiento adecuado del instrumento.

Al analizar los estudios incluidos en esta revisión se documentó una práctica que pareciera común en algunos estudios de validación de instrumentos en general; la eliminación de ítems sin otra explicación que la mejora de la fiabilidad, las cargas factoriales e índices de ajuste de los modelos. En algunos de los estudios revisados la eliminación de ítems en la HADS se basó en los resultados del ACP, con ello se obtuvo evidencia de estructura interna a expensas de perder amplitud de constructo, y como se mencionó anteriormente, esta no es la estrategia adecuada ni vigente para analizar la estructura factorial de acuerdo con los objetivos de los estudios revisados; y tampoco es un fundamento suficiente para discernir la eliminación o permanencia de uno o varios ítems analizados en una sola muestra.

El hecho de eliminar ítems de un instrumento debe estar respaldado por expertos en la variable de estudio y en teoría de la medida, ya que podría estar afectándose la validez de contenido (Sireci y Faulkner-Bond, 2014), lo cual tendría implicaciones mayores tratándose de instrumentos como la HADS en los que la eliminación de un ítem representa la omisión de un síntoma y el posible sesgo en las puntuaciones totales de las dimensiones, lo cual puede repercutir en el cribado para tomar la decisión de canalizar al paciente a una evaluación a profundidad.

Otra implicación de la eliminación de ítems en un instrumento es la existencia de versiones no homogéneas, por lo que debería valorarse si esa versión puede ser considerada como la versión fiel y adaptada del instrumento original. Por otra parte, es importante resaltar el respeto a los derechos de autor, ya que la eliminación de ítems significa la modificación del propio instrumento, por lo que este tipo de decisiones deberían estar consultadas con quienes crearon el instrumento. En esta re-

visión se encontraron estudios en los que se decidió eliminar ítems, sin embargo, las decisiones sólo se basaron en los análisis estadísticos.

Eliminar ítems en los instrumentos puede tener implicaciones más allá de los beneficios obtenidos al mejorar el resultado estadístico, el más común es la mejora de la consistencia interna del instrumento. La evidencia de fiabilidad en los estudios revisados se obtuvo mediante el cálculo del Alfa de Cronbach, que hasta hace algunos años era considerado como una medida suficiente, sin embargo se ha documentado que en algunos casos dependiendo de las características de los datos este coeficiente puede subestimar o sobreestimar la consistencia interna de los tests. Hoy en día también se propone su cálculo a través del coeficiente omega de McDonald con el que el coeficiente de fiabilidad de consistencia interna es considerado un subproducto del modelo de medición, que permite ajustar un modelo de medición para que a partir de los parámetros estimados sean obtenidos los coeficientes de fiabilidad (Viladrich et al., 2017).

El análisis realizado en esta revisión sistemática permite concluir lo siguiente. En México, las propiedades psicométricas de las puntuaciones de la HADS se han evaluado en diferentes muestras de personas con enfermedades físicas o mentales; sin embargo, se trabaja con cinco versiones diferentes de la HADS y la ausencia o limitada información respecto al proceso de traducción y adaptación del mismo pone en duda la calidad del instrumento. Este hecho no permite valorar adecuadamente los resultados sobre la fiabilidad y validez de las puntuaciones de la HADS. El amplio uso de la escala por parte de los clínicos e investigadores, así como la aceptación por parte de los pacientes hacen recomendable seguir utilizándola; no obstante, es necesario realizar una traducción y adaptación de la escala considerando las directrices y avances en el campo de la psicometría.

En la actualidad, el abordaje de la psicología en los diversos fenómenos multifactoriales demanda la incorporación de un mayor número de variables para su estudio. Desde la práctica y la investigación psicológica es posible medir, evaluar, reportar y analizar los constructos a través de los tests, sin embargo, incluir un mayor número de variables conlleva estructurar una amplia batería de pruebas, situación que puede afectar la factibilidad del estudio al entorpecer la calidad del proceso de respuesta, prolongando el tiempo de la evaluación, contribuyendo al cansancio o la poca disposición para permanecer tiempos prolongados contestando los ítems, entre otras cuestiones que pueden afectar la calidad de los datos. En este sentido el empleo de instrumentos en versiones cortas beneficia la recolección de datos y favorece la inclusión de un mayor número de variables; como es el caso de la HADS, siendo

una escala corta de dos dimensiones que permite el tamizaje de dos de las condiciones clínicas de mayor prevalencia, la ansiedad y la depresión.

Lo que se documentó en la presente investigación probablemente puede estar pasando con otros instrumentos, por lo tanto es recomendable que quien vaya a utilizar un instrumento realice una búsqueda exhaustiva sobre las propiedades psicométricas del mismo, analizando detalladamente no sólo los resultados sobre estas propiedades, también el proceso de traducción y adaptación; y considerando siempre la revisión del trabajo original y el propósito por el que fue creada la prueba. No obstante, la revisión exhaustiva en lo relativo a las propiedades psicométricas del instrumento no será suficiente para detectar anomalías si los profesionales e investigadores en psicología no se mantienen actualizados en cuanto a los avances en psicometría, no bastará con solo considerar si las estrategias para análisis de los datos psicométricos han sido repetidas en un gran número de estudios para considerarlas pertinentes (Lloret-Segura et al., 2014), sobre todo cuando estos fueron producidos años atrás. Es necesario estar informado para poder cuestionar los procedimientos y con ello tomar decisiones apropiadas, incluso mejores a las ya realizadas en la validación de un instrumento para poder emplearlo de manera adecuada.

Conflictos de intereses

Los autores declaran no tener conflictos de intereses.

Referencias

- *Anguiano, S. A., Mora, M. A., Reynoso, L., y Vega, C. Z. (2017). Prevalencia de ansiedad y depresión en pacientes hospitalizados. *Eureka*, 14(1), 24–38.
- *Barriguet, J. A., Pérez, A. R., de la Vega, R. I., Barriguet, P., y Rojo, L. (2017). Validation of the Hospital Anxiety and Depression Scale in Mexican population with eating disorders. *Revista Mexicana de Trastornos Alimentarios*, 8(2), 123-130. <https://doi.org/10.1016/j.rmta.2017.05.001>
- Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of Consulting Clinical Psychology*, 56, 893-897. <https://doi.org/10.1037/0022-006X.56.6.893>
- Beck, A., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 53-63. <https://doi.org/10.1001/archpsyc.1961.01710120031004>
- Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002a). The validity of the Hospital Anxiety and Depression Scale: An updated literature review. *Journal of Psychosomatic Research*, 52(2), 69–77. [https://doi.org/10.1016/S0022-3999\(01\)00296-3](https://doi.org/10.1016/S0022-3999(01)00296-3)
- Brennan, C., Worrall-Davies, A., McMillan, D., Gilbody, S., & House, A. (2010). The Hospital Anxiety and Depression Scale: A diagnostic meta-analysis of case-finding ability. *Journal of Psychosomatic Research*, 69(4), 371–378. <https://doi.org/10.1016/j.jpsychores.2010.04.006>
- Breugelmans, R. (2009). Dangers in using translated medical questionnaires: the importance of conceptual equivalence across languages and cultures in patient-reported outcome measures. *Chest*, 136(4), 1175–1177. <https://doi.org/10.1378/chest.09-1684>
- Cosco, T. D., Doyle, F., Ward, M., & McGee, H. (2012). Latent structure of the Hospital Anxiety and Depression Scale: A 10-year systematic review. *Journal of Psychosomatic Research*, 72(3), 180–184. <https://doi.org/10.1016/j.jpsychores.2011.06.008>
- Elosua, P. (2017). Avances, proyectos y retos internacionales ligados al uso de tests en Psicología. *Estudios de Psicología*, 34(2), 201-210. <https://doi.org/10.1590/1982-02752017000200002>
- Elosua, P., Mujika, J., Almeida, L. S., y Hermsilla, D. (2014). Procedimientos analítico-rationales en la adaptación de tests. Adaptación al español de la batería de pruebas de razonamiento. *Revista Latinoamericana de Psicología*, 46(2), 117-126. [https://doi.org/10.1016/s0120-0534\(14\)70015-9](https://doi.org/10.1016/s0120-0534(14)70015-9)
- *Galindo, O., Benjet, C., Juárez, F., Rojas, E., Riveros, A., Aguilar, J. L., Álvarez, M. A. y Alvarado, S. (2015). Propiedades psicométricas de la Escala Hospitalaria de Ansiedad y Depresión (HADS) en una población de pacientes oncológicos mexicanos. *Salud Mental*, 38(4), 253-258. <https://doi.org/10.17711/sm.0185-3325.2015.035>
- *Galindo, O., Meneses, A., Herrera, A., Caballero, M. R., y Aguilar, J. L. (2015). Escala Hospitalaria de Ansiedad y Depresión (HADS) en cuidadores primarios informales de pacientes con cáncer: propiedades psicométricas. *Psicooncología*, 12(2-3), 383-392. https://doi.org/10.5209/rev_PSIC.2015.v12.n2-3.51016
- *Guillén-Díaz-Barriga, C., González-Celis, A. L., Sánchez, C. T., y Fernández, M. (2018). Escala de Ansiedad y Depresión Hospitalaria (HADS) en mexicanos con asma: confiabilidad y validez. *Psicología Iberoamericana*, 26(2), 75-87.
- *Herrera, B., Galindo, O., y Lerma, A. (2018). Inventario de Ansiedad y Depresión Hospitalaria y Escala de Retiro de Patrón de Conducta tipo A en pacientes con enfermedades cardiovasculares: propiedades psicométricas. *Psicología y Salud*, 28(1), 121-130. <https://doi.org/10.25009/pys.v28i1.2544>
- Herrmann, C. (1997). International experiences with the hospital anxiety and depression scale - A review of validation data and clinical results. *Journal of Psychosomatic Research*, 42(1), 17-41. [https://doi.org/10.1016/S0022-3999\(96\)00216-4](https://doi.org/10.1016/S0022-3999(96)00216-4)
- Hinz, A., & Brähler, E. (2011). Normative values for the hospital anxiety and depression scale (HADS) in the general German population. *Journal of Psychosomatic Research*, 71(2), 74-78. <https://doi.org/10.1016/j.jpsychores.2011.01.005>
- Juniper, E. F., Guyatt, G., Cox, F., Ferrie, P., & King, D. (1999). Development and validation of the Mini Asthma Quality of Life Questionnaire. *European Respiratory Journal*, 14(1), 32-38. <https://doi.org/10.1034/j.1399-3003.1999.14a08.x>
- Koller, M., Aaronson, N., Blazeby, J., Bottomley, A., Dewolf, L., Fayers, P., & EORTC Quality of Life Group (2007). Translation procedures for standardized quality of life questionnaires: The European Organisation for Research and Treatment of Cancer (EORTC) approach. *European Journal of Cancer*, 43(12), 1810-1820. <https://doi.org/10.1016/j.ejca.2007.05.029>
- Lloret-Segura, S., Ferreres-Traver, A., y Tomás-Marco, A. H. I. (2014). El análisis factorial exploratorio de los ítems: una guía

- práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169. <https://doi.org/10.6018/analesps.30.3.199361>.
- *López-Alvarenga, J. C., Vázquez-Velázquez, V., Arcila-Martínez, D., Sierra-Ovando, Á. E., González-Barranco, J., y Salín-Pascual, R. J. (2002). Exactitud y utilidad diagnóstica del Hospital Anxiety and Depression Scale (HAD) en una muestra de sujetos obesos mexicanos. *Revista de Investigación Clínica*, 54(5), 403-409.
- *Lugo, V. I., Pérez, Y. Y., y Sánchez, C. T. (2018). Validación de la Escala de Ansiedad y Depresión Hospitalaria en pacientes con ERC en hemodiálisis. *Aportaciones a la Psicología Social*, IV, 1209-1225.
- Maters, G. A., Sanderman, R., Kim, A. Y., & Coyne, J. C. (2013). Problems in Cross-Cultural Use of the Hospital Anxiety and Depression Scale: «No Butterflies in the Desert.» *PLoS ONE*, 8(8). <https://doi.org/10.1371/journal.pone.0070975>
- Mazor, K., Clauser, B., & Hambleton, R. (1992). The effect of sample size on the functioning on the Mantel-Haenszel Statistic. *Educational and Psychological Measurement*, 52, 443-451. <https://doi.org/10.1177/0013164492052002020>
- Morse, R., Kendell, K., & Barton, S. (2005). Screening for depression in people with cancer: the accuracy of the hospital anxiety and depression scale. *Clinical Effectiveness in Nursing*, 9(3-4), 188-196. <https://doi.org/10.1016/j.cein.2006.08.012>
- 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>
- Muñiz, J., Hambleton, R. K., & Xing, D. (2001). Small sample studies to detect flaws in item translations. *International Journal of Testing*, 1(2), 115-135. https://doi.org/10.1207/S15327574IJT0102_2
- Nathan, R. A., Sorkness, C. A., Kosinski M, Schatz, M., Li, T., Marcus, P., Murray, J., & Pendergraft, T. (2004). Development of the asthma control test: A survey for assessing asthma control. *Journal of Allergy and Clinical Immunology*, 113(1), 59-65. <https://doi.org/10.1016/j.jaci.2003.09.008>
- *Noguera, M. J., Pérez, B., Barrientos, V., Robles, R., y Sierra, J. G. (2013). Escala de Ansiedad y Depresión Hospitalaria (HADS): Validación en pacientes mexicanos con infección por VIH. *Psicología Iberoamericana*, 21(2), 29-37.
- Norton, S., Cosco, T., Doyle, F., Done, J., & Sacker, A. (2013). The Hospital Anxiety and Depression Scale: A meta confirmatory factor analysis. *Journal of Psychosomatic Research*, 74, 74-87. <https://doi.org/10.1016/j.jpsychores.2012.10.010>
- Nunnally, J. C., & Bernstein, I. J. (1995). Teoría psicométrica. Madrid: McGraw-Hill.
- Quiñonez-Tapia, F., Méndez, T. E., y 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>
- Rodríguez, S. C., Gil, C. P., y Martínez, R. (1996). Presentación de la Escala Retiro de Patrón de Conducta Tipo A (ERCTA). *Psicothema*, 8(1), 207-213.
- Sireci, S., & Faulkner-Bond, M. (2014). Evidencia de validez basada en el contenido del test. *Psicothema*, 26(1), 100-107. <https://doi.org/10.7334/psicothema2013.256>
- Terol-Cantero, M., Cabrera-Perona, V., & Martín-Aragón, M. (2015). Hospital Anxiety and Depression Scale (HADS) review in Spanish samples. *Anales de Psicología*, 31(2), 494-503. <https://doi.org/10.6018/analesps.31.2.172701>
- Urrutia, G., & Bonfill, X. (2010). PRISMA declaration: A proposal to improve the publication of systematic reviews and meta-analyses. *Medicina Clínica*, 135(11), 507-511. <https://doi.org/10.1016/j.medcli.2010.01.015>
- Viladrich, C., Ángulo-Brunet, A., y Doval, E. (2017). Un viaje alrededor de alfa y omega para estimar la fiabilidad de consistencia interna. *Anales de Psicología*, 33(5), 755-782. <https://doi.org/10.6018/analesps.33.3.268401>
- Vodermaier, A., & Millman, R. D. (2011) Accuracy of the Hospital Anxiety and Depression Scale as a screening tool in cancer patients: a systematic review and meta-analysis. *Support Care Cancer*, 19(12): 1899–1908. <https://doi.org/10.1007/s00520-011-1251-4>
- *Yamamoto-Furusho, J. K., Sarmiento-Aguilar, A., García-Alanis, M., Gómez-García, L. E., Toledo-Mauriño, J., Olivares-Guzmán, L., & Fresán-Orellana, A. (2018). Hospital Anxiety and Depression Scale (HADS): Validation in Mexican patients with inflammatory bowel disease. *Gastroenterología y Hepatología (English Edition)*, 41(8), 477-482. <https://doi.org/10.1016/j.gastre.2018.08.009>
- Zarit, H. S., Reever, E. K., & Bach-Peterson, J. (1980). Relatives of the impaired elderly: correlates of feelings burden. *Gerontologist*, 20, 649-55. <https://doi.org/10.1093/geront/20.6.649>
- Zigmond, A., & Snaithe, R. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, 67(6), 361-370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>