

**Evaluación dimensional de la depresión: Una revisión sistemática de estudios  
instrumentales**

**Brigith Natyé Sarmiento Rada**



**UNIVERSIDAD  
DE LA COSTA**  
1970

**Corporación Universidad de la Costa “CUC”**

**Facultad de Ciencias Humanas y Sociales**

**Programa de Psicología**

**Barranquilla**

**2020**

**Evaluación dimensional de la depresión: Una revisión sistemática de estudios  
instrumentales**

**Brigith Natyé Sarmiento Rada**

**Tutor(a): Mg. Paola Andrea García Roncallo**

**Cotutor(a): Mg. Eileen Andrea García Montaña**

**Notas de autor:**

**Corporación Universidad de la Costa “CUC”**

**Facultad de Ciencias Humanas y Sociales**

**Programa de Psicología**

**Barranquilla**

**2020**

**NOTA DE ACEPTACIÓN**

---

---

---

**Firma de la Decana del Departamento de Ciencias Sociales**

---

**Firma del Decano de Facultad de Ciencias Humanas y Sociales**

---

**Firma de la directora del Programa de Psicología**

---

**Firma Líder de Grupo de Investigación de Evolutiva y Psicopatología**

---

**Firma de Tutora**

---

**Firma de Co-tutora**

---

**Firma Juez Interno**

---

**Firma Juez Externo**

**Barranquilla,**

**PhD. Ketty Herrera**

Líder del grupo de investigación Evolutiva y Psicopatología

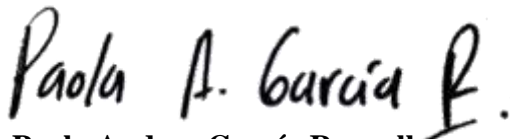
E.S.M.

Atentamente me permito hacerle la entrega formal del Trabajo de grado titulado “Evaluación dimensional de la depresión: una revisión sistemática de estudios instrumentales” presentado por la estudiante Brighth Natyé Sarmiento Rada.

Como tutora considero que el trabajo de grado cumple satisfactoriamente con los criterios de calidad y rigor científico. El trabajo consolidado en este documento evidencia la suficiencia de las competencias investigativas de la estudiante, por lo tanto, comedidamente solicito su colaboración para someterlo a evaluación de jurados.

Sea esta la oportunidad también, para expresar mis felicitaciones por su trabajo, compromiso y dedicación a Brighth, quien desarrolló su estudio cumpliendo con los estándares de calidad exigidos para los estudios de revisión sistemática, materializando una contribución científica significativa a la consolidación de la evidencia probada de los avances psicométricos al servicio de la evaluación clínica, psicológica de los trastornos depresivos.

Gracias por su atención,



**Paola Andrea García Roncallo**

**CC. 1.140.841.407**

**Profesora Tiempo Completo - Investigadora**

**Departamento de Ciencias Sociales**

### **Agradecimientos**

Infinitas gracias a Dios, por la vida, quien ha forjado mi camino y me ha dirigido por el sendero correcto.

Le dedico este triunfo a mi madre Beatriz y mi abuela Rita, quienes son mi mayor fuente de motivación, por su amor incondicional, por su paciencia, esfuerzos y sacrificios, gracias por inspirarme y convertirme en la mujer que soy ahora.

Gracias a mis profesores, personas de gran sabiduría que me han formado en esta hermosa profesión.

Especial agradecimiento a mis tutoras Paola García, Eileen García y Daniella Abello, por su apoyo constante, dedicación y preocupación por mi formación académica. Reitero mi respeto, admiración y gran cariño a ustedes porque sin su ayuda y conocimientos no hubiese sido posible realizar este proyecto. Gracias familia C2 por cada momento compartido con ustedes, este triunfo es de todos.

Gracias amiga Jou, por tu amistad incondicional, por las alegrías, las tristezas y ser mi soporte en muchas ocasiones.

Y gracias a aquellos que de alguna manera aportaron a la realización de este proyecto.

Con amor,

**Brigith**

### Resumen

La presente investigación tuvo como objetivo analizar la evidencia científica de los estudios de medición de sintomatología depresiva empleados en la práctica clínica e investigativa en adultos hispanohablantes entre 18 a 65 años a través de una revisión sistemática de estudios instrumentales. Se realizaron búsquedas electrónicas en las bases de datos Scopus, ScienceDirect y Web of Science (WOS) de estudios originales sobre el desarrollo, uso, validación o adaptación de instrumentos que evalúen sintomatología depresiva. La calidad metodológica de los estudios seleccionados y los resultados de la adecuación conceptual, los modos de aplicabilidad y las propiedades psicométricas se evaluaron críticamente utilizando una lista de verificación diseñada para evaluar específicamente estudios de propiedades psicométricas y atiendo a las normas para el desarrollo y revisión de estudios instrumentales. La estrategia de búsqueda dio como resultado un total de 1.793 estudios, de los cuales 53 artículos cumplieron con los criterios de elegibilidad y se incluyeron en la revisión. Los resultados indican un aumento del número de publicaciones en los últimos años de estudios instrumentales en depresión. El Inventario de Depresión de Beck y la Escala de Depresión del Centro de Estudios Epidemiológicos fueron los instrumentos de medición más evaluados. En términos generales, la mayoría de los estudios no dieron un reporte completo de aspectos relevantes del proceso de construcción/adaptación de los instrumentos de medida.

***Palabras clave:*** Revisión sistemática, depresión, estudios instrumentales, propiedades psicométricas

### **Abstract**

The aim of this research was to analyze the scientific evidence of the depressive symptomatology measurement studies used in clinical and research practice in spanish-speaking adults between 18 and 65 years old through a systematic review of instrumental studies. Electronic searches were conducted in Scopus, ScienceDirect and Web of Science (WOS) databases of original studies on the development, use, validation or adaptation of instruments that evaluate depressive symptomatology. The methodological quality of the selected studies and the results of the conceptual adequacy, the modes of applicability and the psychometric properties were critically evaluated using a checklist designed to specifically evaluate studies of psychometric properties and following the norms for the development and revision of instrumental studies. The search strategy resulted in a total of 1,793 studies, of which 53 articles met the eligibility criteria and were included in the review. The results indicate an increase in the number of publications in recent years of instrumental studies in depression. The Beck Depression Inventory and the Center for Epidemiologic Studies Scale (CES-D) were the most evaluated measuring instruments. In general terms, most of the studies did not give a complete report of relevant aspects of the construction / adaptation process of the measuring instruments.

**Keywords:** Systematic review, depression, instrumental studies, psychometric properties.

**Contenido**

Lista de tablas y figuras .....	10
Introducción .....	11
Definición del problema .....	13
Planteamiento del problema .....	13
Formulación del problema .....	17
Justificación .....	19
Objetivos .....	22
Objetivo general .....	22
Objetivos específicos .....	22
Marco teórico .....	23
Antecedentes del Trastorno depresivo .....	23
Manifestaciones clínicas de los trastornos depresivos .....	28
Teorías explicativas de la depresión .....	42
Perspectiva biológica de los trastornos depresivos .....	42
Teorías conductuales de los trastornos depresivos .....	43
Teorías cognitivas de los trastornos depresivos .....	48
Determinantes psicosociales de la depresión en la edad adulta .....	59
Evaluación de las manifestaciones clínicas de la depresión en la edad adulta .....	64
Evaluación de las variables psicológicas .....	65
Tipos de evaluación .....	66
Calidad de los instrumentos de medida: contribuciones de la psicometría .....	68
Los principios del modelo psicométrico .....	70
Teorías de los Test .....	79
Practica psicométrica basada en la evidencia .....	81



Metodología .....	87
Población y muestra .....	87
Variables y/o indicadores del estudio.....	87
Instrumentos y técnicas .....	89
Procedimiento.....	89
Parte A. Búsqueda bibliográfica.....	90
Parte B. Evaluación de las propiedades psicométricas.....	92
Parte C. Selección de los instrumentos.....	94
Resultados .....	96
Población y muestra .....	96
Características generales de los estudios e instrumentos .....	97
Potencial impacto de los estudios instrumentales sobre el análisis de los indicadores bibliométricos.....	105
Análisis de la adecuación conceptual de las medidas .....	105
Análisis de las propiedades psicométricas de los instrumentos .....	108
Discusión.....	130
Conclusiones y recomendaciones .....	145
Limitaciones de la revisión .....	145
Recomendaciones finales .....	146
Referencias.....	149
Referencias de los estudios incluidos .....	178

**Lista de tablas y figuras****Tablas**

Tabla 1 Criterios diagnósticos para los Trastornos Depresivos.....	31
Tabla 2 Etapas de la adultez.....	61
Tabla 3 Matriz de variables e indicadores del estudio.....	88
Tabla 4 Caracterización de los estudios.....	98
Tabla 5 Caracterización de los instrumentos .....	102
Tabla 6 Propiedades psicométricas de los instrumentos: diferenciación entre sujetos.....	109
Tabla 7 Propiedades psicométricas de los instrumentos: Reportes de análisis factorial y confiabilidad .....	114
Tabla 8 Propiedades psicométricas de los instrumentos: validez externa .....	123

**Figuras**

Figura 1. Diagrama de flujo.....	97
----------------------------------	----

## Introducción

La depresión se ha constituido como un trastorno mental frecuente alrededor del mundo que afecta significativamente a la persona que lo padece, consolidándose como la principal causa de discapacidad a nivel mundial y contribuye a la carga general de morbilidad (Organización Mundial de la Salud [OMS], 2020). Frente a esta situación, han surgido iniciativas para dar respuesta a esta problemática como la generación y estandarización de instrumentos psicométricos que evalúen de manera objetiva la presencia y frecuencia de síntomas depresivos, que permitan el establecimiento de objetivos, selección de técnicas y estrategias para un proceso adecuado de intervención, por lo tanto, los instrumentos deben dar reporte de elementos científicos que permita enjuiciar su calidad. En este sentido, la presente investigación tiene como objetivo evaluar los instrumentos de medición de sintomatología depresiva en adultos hispanohablantes de 18 a 65 a través de la metodología de revisión sistemática.

El marco teórico, de manera general, se centra en la fundamentación teórica en la que se enmarca el diseño de la presente investigación tomando aspectos como la evolución de la medición de la depresión, los diferentes postulados de las teorías explicativas, principales síntomas, determinantes psicosociales de la depresión en la edad adulta, importancia de la evaluación psicométrica y fundamentos, y finalmente, la relevancia de las revisiones sistemáticas de instrumentos de medición.

La presente investigación corresponde a un estudio documental (Páramo, 2020) con un diseño metodológico de revisión sistemática instrumental de evaluación de sintomatología depresiva en adultos hispanohablantes de 18 a 65 años, en la medida que sintetiza la mejor evidencia científica disponible con el objetivo de contribuir al mejoramiento del desarrollo y uso de instrumentos implementados comúnmente por profesionales de la salud mental.

La estrategia de búsqueda se realizó mediante la elaboración de un “*Banco de códigos de búsqueda*” conformado principalmente por términos metodológicos y clínicos del constructo, bases de datos y operador booleano. La extracción de datos y la evaluación de la calidad metodológica atendió las directrices de los Estándares basados en el censo para la Selección de Instrumentos de medición de la Salud [COSMIN por sus siglas en inglés] (Prinsen, Mokkink, Bouter, Alonso, Patrick, de Vet & Terwee, 2018) y por las Normas para el Desarrollo y Revisión de estudios instrumentales (Carretero-Dios y Pérez, 2005).

## **Definición del problema**

### **Planteamiento del problema**

La salud mental se caracteriza por un estado completo de bienestar físico, mental y social, y no solo por la ausencia de enfermedades o trastornos mentales, en el que el individuo tiene la capacidad de hacer frente al estrés cotidiano de la vida, de trabajar productivamente y contribuir a su comunidad (Organización Mundial de la Salud [OMS], 2018). Aproximadamente 450 millones de personas padecen algún trastorno mental (OMS, 2004). Los trastornos afectivos son responsables de buena parte de esta prevalencia, puesto que 300 millones de personas sufren de trastorno depresivo y se constituye como la principal causa de discapacidad y contribuye a la carga mundial general de morbilidad (OMS, 2020), y en los casos de peor pronóstico, puede llevar al suicidio (World Federation for Mental Health [WFMH], 2012); en la región de las Américas aproximadamente 65.000 personas mueren cada año por esta causa (Organización Panamericana de la Salud [OPS], 2014). La depresión abarca una serie de síntomas tanto psíquicos como físicos, sin causa orgánica, que suelen ser recurrentes o de larga duración, afectando significativamente la capacidad de realizar actividades laborales, académicas y afrontar la cotidianidad (OMS, 2020).

En Colombia, la Encuesta Nacional de Salud Mental, reportó que los adultos entre los 18 y 44 años manifestaron más de 7 síntomas depresivos, siendo más prevalentes en mujeres (4,9%) que en hombres (3,2%) (Ministerio de Salud y Protección Social y Departamento Administrativo de Ciencia, Tecnología e Innovación [Colciencias], 2015).

La OMS (2020) aclara que, pese a que existen tratamientos eficaces para abordar la depresión, más de la mitad de los afectados a nivel mundial (y más del 90% en muchos países)

no acceden a estos tratamientos, siendo la falta de recursos y personal sanitario capacitado, la estigmatización de los trastornos mentales, la evaluación clínica inexacta o errónea (en algunos casos son tratados con medicamentos inadecuados que aceleran las fases críticas de sus enfermedades), los obstáculos para una atención eficaz.

La OPS (2014) refiere que la disponibilidad de datos de servicios de salud mental ha proporcionado una mejor comprensión acerca de la magnitud de la brecha de tratamiento en las américas, para los trastornos afectivos, de ansiedad y por consumo de sustancias en adultos, graves y moderados, la mediana de la brecha de tratamiento es de 73,5% en la región, 47,2% en América del Norte y 77,9% en América Latina y el Caribe (ALC).

Kohn, Saxena, Levav y Saraceno (2004) aseguran que el aumento de la brecha terapéutica tiene un impacto en años de vida ajustados por discapacidad (AVAD) y en años vividos con discapacidad, pero que además, genera costos económicos indirectos como la pérdida de productividad en el trabajo, muerte prematura, el alcoholismo, mayor riesgo de vivir en la pobreza, nivel socioeconómico más bajo, menor nivel educativo, deterioro de la función familiar, incrementa el riesgo de parto adolescente, riesgo de violencia doméstica, impacto en las condiciones médicas crónicas e individuos asintomáticos.

En Colombia, la brecha terapéutica responde a la percepción de la sociedad frente a la enfermedad mental y la no búsqueda de ayuda (barreras actitudinales: “descuido” o “no considerarlo importante”), lo que limita el acceso oportuno y pertinente para todos los solicitantes que ascienden ya a un 38,5% de adultos (entre 18 a 44 años) con problemas mentales, con un promedio de solicitud de 23,5 veces. Esta situación amerita la implementación de protocolos o guías adecuadas para trastornos mentales, mantener la educación continua a los

profesionales de la salud y favorecer la investigación que permita un mayor reconocimiento y atención adecuada (Ministerio de Salud y Protección Social y Colciencias, 2015).

Frente al problema de la depresión, la OMS (2020) aclara que, existen tratamientos psicológicos para solventar el impacto de la depresión en los individuos que la padecen, siendo la activación conductual y la terapia cognitiva conductual (TCC) los métodos más eficaces para abordarla. Sin embargo, para llevar a cabo un buen proceso de intervención, la evaluación es un requerimiento indispensable, dado que no es posible planear una intervención si no se tiene información acerca del problema que aqueja a la persona, y que sólo la puede proporcionar una correcta evaluación (Aragón, 2004). En este sentido, Lara, Navarro, Mondragón, Rubí y Lara (2002) señalan que la evaluación de la depresión puede efectuarse mediante la determinación del cumplimiento o no de criterios diagnósticos (categorial), o mediante la valoración de parámetros de intensidad, frecuencia o cronicidad en continuos objetivos (dimensional). La evaluación categorial permite el establecimiento de un diagnóstico mediante el uso de la entrevista estructurada, sin embargo, las entrevistas diagnósticas como la entrevista clínica estructurada o las entrevistas para la evaluación clínica en neuropsiquiatría no es posible emplearlas rutinariamente debido a la falta de personal entrenado y el tiempo de su aplicación (Lara et al., 2002).

Por su parte, Carretero-Dios y Pérez (2007) resaltan la importancia del uso de los instrumentos en la investigación y en la práctica clínica, debido a que estos posibilitan la identificación de la presencia de síntomas en su intensidad y frecuencia. No obstante, este tipo de herramientas puede incurrir en la sobrediagnóstico de cuadros clínicos como la depresión en comparación con la evaluación categórica (Lara et al., 2002), que podría derivar un diagnóstico

erróneo e impactando en la calidad del diseño terapéutico (Abello y Cortés, 2012), por lo que la verificación de su calidad psicométrica es esencial para brindar garantías en su uso.

Ahora bien, Carretero y Pérez (2007) señalan que la selección de los test, en el uso de la investigación, tiene una influencia directa en los resultados finales, por lo tanto, una *selección razonada* es un requerimiento ineludible; sin embargo, tal selección suele estar orientada por criterios no científicos y que en mucho de los casos no se informan las razones que llevaron a emplearlos; muchos de los test publicados carecen de información necesaria sobre la elaboración en cuanto a la definición de la variable, objeto de evaluación, los valores de los coeficientes de fiabilidad y validez, proceso de edición de los ítems, justificación del número de ítems suficientes que representen el constructo a evaluar, representación de las dimensiones a través de los ítems y otros aspectos relevantes que enjuicien la calidad de los test.

En este sentido, la selección de los instrumentos en la práctica e investigación en el área de la salud mental pueden llevarse a cabo a través de las directrices de la Práctica Basada en la Evidencia, cuya finalidad se centra en la selección de las mejores evidencias científicas para la toma de decisiones en el ámbito profesional mediante un proceso de búsqueda, selección y análisis de estudios; sin embargo, tal proceso se entorpece en la medida que influyen una serie de factores que impiden llevar a cabo un proceso preciso y coherente, como el exceso de estudios empíricos, el corto tiempo que cuenta el profesional y la alta carga laboral, de ahí, radica la implementación de las revisiones sistemáticas (RSs) y metaanálisis (MAs) (Sánchez-Meca y Botella, 2010), en la medida que sintetizan sistemáticamente toda la evidencia empírica de un mismo tema de estudio, y su empleo se ha extendido como herramientas eficaces para la selección en la investigación y la práctica clínica de instrumentos de medición (Trujols, Ballesteros, Solá y Portella, 2015).



Las revisiones sistemáticas instrumentales pueden ser de diferentes tipos, algunas enfocadas en un instrumento en particular y otras en la evaluación de diversos instrumentos de una misma variable. En el primer caso existen investigaciones basadas en la evaluación de un instrumento de medición, por ejemplo, la Escala de Depresión Geriátrica-versión breve (Pocklington, Gilbody, Manea y McMillan, 2016), el Inventario Rápido de Sintomatología Depresiva de 16 Ítems (Reilly, MacGillivray, Reid y Cameron, 2014), el Inventario de Depresión de Beck (Wang y Gorenstein, 2013), la Escala de Ansiedad Depresión Hospitalaria (Cosco, Doyle, Ward y McGee, 2012), la Escala de Depresión Postnatal de Edimburgo (Gibson, McKenzie-McHarg, Shakespeare, Price y Gray, 2009) y la Escala de Depresión Geriátrica (Wancata, Alexandrowicz, Marquart, Weiss y Friedrich, 2006). Por otra parte, en el caso de revisiones de distintos instrumentos, se encontraron investigaciones instrumentales de medición de síntomas depresivos en pacientes con esquizofrenia (Lako, Bruggeman, Knegeter, Wiersman, Schoevers, Slooff y Taxis, 2012) y en personas con discapacidades intelectuales (Hermans y Evenhuis, 2010; Pérez-Achiaga, Nelson y Hassiotis, 2009).

Con base a lo anterior, la revisión de la literatura sugiere un vacío investigativo entorno a revisiones sistemáticas que evalúen diversos instrumentos de medición de sintomatología depresiva en población hispanohablantes adultos entre 18 a 65 años con respecto a la adecuación conceptual, aplicabilidad, y propiedades psicométricas implementados en la práctica e investigación clínica.

### **Formulación del problema**

Atendiendo este panorama científico, la presente investigación busca consolidar la información relevante de los instrumentos psicométricos a través de la elaboración de una revisión sistemática, que responda la siguiente pregunta problema: *¿Cuáles son los soportes*

*teóricos, las propiedades psicométricas y la aplicabilidad de los instrumentos de medición de sintomatología depresiva empleados en la práctica e investigación clínica en adultos de 18 a 65 años?*

### **Justificación**

Los organismos internacionales y nacionales pretenden dar respuesta efectiva a la problemática expuesta mediante la implementación de planes de acción para el mejoramiento y acceso a la atención primaria, mejorando los procesos de diagnóstico, tratamiento y rehabilitación, así como fortalecer la promoción y prevención de los trastornos mentales a nivel mundial y en la región de las Américas.

En la 65ª Asamblea Mundial de la Salud de 2012, se adoptó la resolución WHA65.4 sobre la carga mundial de trastornos mentales y la necesidad de una respuesta integral acorde a las especificidades de cada nación, mediante la elaboración y refuerzo de políticas y estrategias en el marco de la salud mental (OMS, 2012). Posteriormente, la 66ª Asamblea Mundial de la Salud de 2013 (OMS, 2013), adoptó la resolución WHA66.8 sobre el Plan de Acción Integral sobre Salud Mental 2013-2020, el cual se basa en fomentar el bienestar mental, prevención de los trastornos mentales, su mejoramiento atencional y recuperación, promoción de los derechos, reducción de la mortalidad, morbilidad y discapacidad de las personas con trastornos mentales a través del fortalecimiento de los sistemas de información, datos científicos e investigaciones en torno a la salud mental.

Así mismo, el Programa de Acción para Superar las Brechas en Salud Mental (OMS, 2008), tiene como objetivo reforzar el compromiso de todos los agentes interesados para el aumento de asignación en cuanto a recursos humanos y financieros para la atención de los trastornos (mentales, neurológicos y por abuso de sustancias) y a la vez lograr una mayor cobertura, especialmente en países de bajos y medios ingresos.

Por su parte, el Plan de Acción sobre Salud Mental 2015-2020 (OPS, 2014) valora la importancia de las investigaciones científicas para el suministro de la mejor evidencia para las intervenciones de promoción, prevención y tratamiento de los trastornos mentales. En este sentido, el Plan Estratégico de la OPS 2014-2019 (OPS, 2014) se centra en proporcionar a la población de las Américas una salud óptima, contribuir al bienestar de las familias y comunidades, promover la equidad en salud, combatir la enfermedad, mejorar la calidad y prolongar la duración de la vida.

En Colombia, el Plan Estratégico Sectorial 2014-2018, entre sus objetivos se encuentra el mejoramiento de las condiciones de salud de la población y la disminución de las brechas en resultados de salud, la cual requiere cerrar las brechas entre regiones, comunidades y grupos económicos, así como, promover la equidad en el acceso a los servicios sanitarios y de calidad (Ministerio de Salud y Protección Social, 2015). Así mismo, a nivel nacional se aprobó la Ley 1616 del 2013, la cual, pretende garantizar a la población colombiana la promoción de la salud mental y prevención de trastornos mentales, atención integral que incluya diagnóstico, tratamiento y rehabilitación (Ley N° 1616, 2013). Por su parte, la Encuesta Nacional de Salud Mental (Ministerio de Salud y Protección Social y Colciencias, 2015) resalta la necesidad de crear, validar, difundir e implementar guías de prácticas clínicas y protocolos de evaluación, por lo que propone favorecimiento de investigación asociadas.

En este sentido, la presente investigación busca contribuir al desarrollo científico, a través de la elaboración de una revisión sistemática, herramienta de gran utilidad en el proceso de toma de decisiones en la práctica clínica, que integra de forma objetiva y sistemática los resultados de estudios empíricos (Argimón y Jiménez, 2013; Sánchez-Meca, 2010). En este estudio, se propenderá por la caracterización psicométrica de los instrumentos de medición de

sintomatología depresiva aportando con ello un insumo valioso para el personal de salud mental, que facilite la elección eficiente de instrumentos de exploración clínica e investigación de depresión.

Así mismo, contribuirá a la producción científica del Semillero Cognitivo Conductual, suscrito a la línea de investigación Neurociencia cognitiva y Salud mental y a la sublínea de neurociencia cognitiva aplicada de la Universidad de la Costa (Facultad de Ciencias Humanas y Sociales- Programa de Psicología).

## Objetivos

### **Objetivo general.**

Analizar la evidencia científica sobre la adecuación conceptual, las propiedades psicométricas y aplicabilidad de los instrumentos de medición de sintomatología depresiva empleados en la práctica clínica e investigativa en adultos hispanohablantes de 18 a 65 años.

### **Objetivos específicos.**

1. Identificar los instrumentos psicométricos disponibles, en la literatura científica, para la medición de sintomatología depresiva en hispanohablantes.
2. Caracterizar el impacto potencial de los estudios instrumentales identificados, sobre la medición de la depresión en hispanohablantes, mediante sus indicadores bibliométricos
3. Caracterizar la adecuación conceptual, los modos de aplicabilidad y las propiedades psicométricas de los instrumentos psicométricos identificados.
4. Comparar la evidencia empírica disponible sobre los instrumentos psicométricos de medición de la depresión en hispanohablantes.

### **Marco teórico**

El presente apartado se centra en la fundamentación teórica en la que se enmarca el diseño de la investigación, teniendo en cuenta: aspectos históricos más relevantes de la evolución de la medición de los Trastornos Afectivos enfocándose en la medición de la depresión, descripción de la sintomatología, visión general de las diferentes teorías explicativas, implicancia de la depresión en el desarrollo evolutivo, importancia de la evaluación psicométrica y sus fundamentos, y finalmente, la importancia de las revisiones sistemáticas de instrumentos de medición.

### **Antecedentes del Trastorno depresivo**

Se considera que los sentimientos depresivos han estado presentes a lo largo de la historia de la humanidad; en la antigüedad se denominaba “melancolía” (término de origen médico) como producto de la doctrina antigua “el humorista” planteada por Hipócrates, quien explicaba la salud y la enfermedad como un equilibrio-desequilibrio de cuatro humores: sangre, flema, bilis y atrabilis, y su expulsión: sudor, vómitos, expectoración, orina y deyecciones, conocidas como las crisis (Zarragoitía, 2011). Los griegos, reconocían un estado de locura delirante, con ánimo exaltado; Soranus comprobó una posible relación entre el estado de locura delirante con melancolía y describió los episodios mixtos; Areteo de Capadocia reconoció la melancolía como una alteración del ánimo que está siempre frío y adherido a un pensamiento de tristeza (causada por la bilis negra), estableció una relación entre episodios depresivos y maníacos y consideraba la melancolía como forma fundamental de la locura; para los griegos, cuatro eran los temperamentos: sanguíneo, colérico, melancólico y flemático (Zarragoitía, 2011).

Galeno de Pérgamo (131-201) describió la melancolía como una alteración crónica (sin fiebre) con diversos síntomas: temor, miedo o desesperación, suspicacia, misantropía y cansancio de la vida, etc.; Isaq ibn Imran, en el siglo X, refería la melancolía como un sentimiento de aflicción y aislamiento que se crea en el alma debido a algo que la persona cree real, pero es irreal, con síntomas somáticos (pérdida de peso y sueño) y de causa puramente psíquica (miedo, aburrimiento o ira,); Constantino, en el siglo XI, planteó que la melancolía perturba el espíritu más que otras enfermedades del cuerpo, cuya ubicación se encuentra en lo íntimo del cerebro (Zarragoitía, 2011).

En la Edad Media, los religiosos consideraban que su causa era por la posesión de los espíritus malos. Santo Tomás (1225-1274) afirmaba que la melancolía era provocada por demonios e influencias astrales; Martín Lutero sostenía que toda la pesadez de la mente y la melancolía provenían del diablo; su tratamiento se realizaba a través de azotes, sangrías, exorcismo y baños; San Isidoro de Sevilla señaló cuatro derivadas de la tristeza: el rencor, la pusilanimidad, la amargura y la desesperación (Zarragoitía, 2011).

En el siglo XIV, se fortaleció la especificación de las características clínicas de la depresión y se efectuaron algunas taxonomías asociadas; Paracelso subclasificó a la depresión en endógena y exógena; Robert Burton (1621) escribió el libro *La anatomía de la melancolía*; Thomas Willis (1621-1675) atribuyó la causa de la depresión a alteraciones químicas del cerebro y el corazón, y estableció cuatro tipos de melancolías: por desorden inicial del cerebro, derivada por los hipocondrios (bazo), la que comprende todo el cuerpo y originada en el útero (Zarragoitía, 2011). Algunos autores reconocieron la presencia de síntomas mixtos y algunas comorbilidades relevantes; Bonetus (1679) trató la relación y significación patogénica entre melancolía y manía; Robert Whytt (1764) estableció relación entre la depresión mental con el



espíritu bajo, la hipocondría y la melancolía; Chiarugi (1784) propuso que durante la evolución de la depresión pueden suceder diversos estados psíquicos, y planteó tres formas de locura: la melancolía, la manía y la demencia; y Wilhem Griesinger introdujo el término *estados de depresión mental* como sinónimo de melancolía (Zarragoitía, 2011).

En el siglo XVIII, se avanzó en la generación de teorías etiológicas de este cuadro clínico. Pitcairn explicó que el desorden de la hidrodinámica de la sangre afectaba el flujo de los espíritus animales de los nervios provocando pensamientos desordenados y delirios de melancolía; William Cullen asoció la melancolía con un estado menor de energía cerebral (hipoergia) (Zarragoitía, 2011).

A principios del siglo XIX, Philippe Pinel señaló las causas posibles de la melancolía (psicológicas y físicas), y poco tiempo después fue aceptada como una enfermedad mental; Robert Burton la definió como un delirio (sin fiebre) acompañado de miedo y tristeza sin causas aparentes; Guislain, sostenía que la melancolía era el fenómeno inicial, seguido por la manía, las ideas delirantes, y, por último, la demencia; Zeller (1849) consideró la melancolía como forma fundamental de la mayor parte de las enfermedades mentales, siendo la manía, la paranoia y la demencia como agravaciones sucesivas a la melancolía; Griesinger (1861) planteó la idea de psicosis única, se hablaba de formas de trastornos psíquicos, susceptibles a ser agrupados en perturbaciones afectivas (melancolía y manía) y perturbaciones del pensamiento (paranoia y demencia), las primeras procedían a las segundas (Zarragoitía, 2011).

A mediados del siglo XIX, Hoffmann (1861) y Snell (1865) plantearon el concepto de psicosis única; consideraban que la paranoia no precedía de la melancolía, sino que constituía la posibilidad de una paranoia primaria; Khalbaum (1863) comenzó a utilizar el término distimia; Kraepelin, empleó el término de locura depresiva (sin desprenderse del término melancolía), y

utilizó la depresión para denominar un estado de ánimo, atribuyendo las melancolías como formas de depresión mental, utilizó como categoría diagnóstica la locura maníaco-depresiva y diferenció la enfermedad maníaco-depresiva de la demencia *praecox*, señaló que el estrés psicológico puede desencadenar los episodios iniciales de manía y los episodios posteriores ocurren espontáneamente (Zarragoitía, 2011).

A medida que transcurría el tiempo, los conceptos de melancolía y depresión eran más acertado y específicos; Mayer, propuso eliminar totalmente el término melancolía y reemplazarlo por depresión; Kahlbaum (1863) implementó el término ciclotimia a las variaciones de las fases depresiva y maníaca; finalmente, la melancolía fue aceptada como concepto de depresión y dio paso a otros términos que especifican las clasificaciones: ansiedad, histeria, hipocondría, obsesión, fobia y trastorno psicossomático, etc. (Zarragoitía, 2011).

En 1908, Gaston Dany, introduce el concepto de *ciclotimia* en un trabajo publicado, y en la tesis de su alumno Pierre Kahn un año después; Kraepelin, en la octava edición de su *Tratado* agrupa prácticamente todos los trastornos depresivos y maníacos en una única enfermedad bajo la noción de maniaco depresiva, recibió una fuerte oposición por psiquiatras europeos, quienes preferían mantener la melancolía como categoría independiente, siendo el caso de Philippe Chaslin (1912) y Carl Wernicke (1906), el primero distinguió cuatro tipos de melancolías (simple o sin delirio, delirante, ansiosa, y con estupor), mientras que el segundo estableció cinco tipos (afectiva, depresiva, parálitica, crónica y agitada); Kleist (1925) denominó “*phasophrenias*” a un grupo conformado por el trastorno afectivo y por el trastorno bipolar (Luque y Berrios, 2011).

En 1924, Bleuler planteó la relación entre maníaco-depresiva y la demencia precoz como un continuo; Mayer brindó la noción de que la psicopatología surgía de interacciones

biopsicosociales; en 1952, el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM) en su primera edición, incluyó el término reacción maníaco-depresiva (Zarragoitía, 2011). En 1957, Karl Leonhard (alumno de Kleist) observó que algunos pacientes presentaban historia de depresión como historia de manía, mientras que otros solo con historia de depresión, notó que los pacientes con historias de manía tenían incidencia de manía en sus familias cuando los comparaba con aquellos que solo tenían historia de depresión, a los primeros los denominó bipolares y los segundos monopolares; gracias a las clasificaciones propuestas por Kleist y Leonhard se dio paso a la aceptación de la separación unipolar/bipolar (Zarragoitía, 2011; Luque y Berrios, 2011).

En 1966, Jules Angst en Suiza y Carlos Perris en Suecia, defendían la distinción nosológica entre los trastornos unipolar y bipolar, cuyos resultados de sus investigaciones apoyaban la existencia de tres entidades independientes dentro del marco de la psicosis: dos formas unipolares (depresiva y maniaca) y una bipolar (Luque y Berrios, 2011). Frente a la aceptación internacional de las categorizaciones propuestas por Angst y Perris, el DSM III, en 1980, incorporó la distinción bipolar-unipolar, que posteriormente fue continuada por el DSM IV (1994) y la Clasificación Internacional de Enfermedades (CIE-10) (1992) (Zarragoitía, 2011; Luque y Berrios, 2011).

La reciente publicación del DSM en su quinta versión, ha generado controversias y polémicas debido a los cambios significativos que en él se presentan, particularmente, el manual divide los “trastornos del estado de ánimo” en “trastornos depresivos” y “trastornos bipolares” manteniendo la característica primordial de un período de dos semanas consecutivas y la presencia de al menos cinco síntomas en el que al menos uno debe ser estado de ánimo depresivo

y/o disminución del interés/placer por todos o casi todas las actividades (García, León, Gálvez y Hernández, 2015).

Las principales modificaciones realizadas en el DSM-5 son (García et al., 2015; de Dios, Goikolea, Colom, Moreno y Vieta, 2014):

- El duelo (reacción o proceso) no excluye la presencia de un episodio depresivo mayor.
- El especificador “con características mixtas” alude a cuadros de coexistencia de síntomas depresivos con al menos tres síntomas maníacos/hipomaníacos de un grupo de 7 síntomas, pero si reúnen los criterios de manía o hipomanía se trasladará a trastorno bipolar I o II, respectivamente.
- El especificador “con ansiedad” alude a cuadros de depresión mayor con al menos dos síntomas de ansiedad de un grupo de cinco síntomas.
- El especificador “con el inicio en el periparto” alude a cuadros de depresión mayor durante la etapa de embarazo (además de las primeras cuatro semanas del postparto).
- El especificador de “crónico” del trastorno depresivo mayor se traslada a los trastornos depresivos persistentes (distimia), contemplando la posibilidad de un trastorno depresivo mayor crónico o con episodios intermitentes.

### **Manifestaciones clínicas de los trastornos depresivos**

Los cambios en el estado de ánimo en los seres humanos son muy frecuentes, sentimientos como tristeza, euforia o irritabilidad se manifiestan como respuesta psicológica normal ante determinadas situaciones, por lo que, sentirse “triste” o “deprimido” no es una característica suficiente para determinar un diagnóstico, de ahí, la importancia de diferenciar la

depresión como síntoma y la depresión como patología (González, Martín, Sánchez y Mirapeix, 2015; Sanz y Vásquez, 2009).

La depresión patológica se caracteriza por una tristeza sin razón aparente que la justifique de forma grave y persistente (Zarragoitía, 2011), acompañado de un conjunto de síntomas como melancolía, nerviosismo, sensación de vacío, deterioro en las relaciones sociales (debido al desinterés por mantenerlas), enlentecimiento físico (o en alguno casos se presenta inquietud), trastornos del sueño (dificultad para conciliar, sueño inquieto o despertar precoz) alteración en la alimentación (reducción o incremento de peso), disminución de la libido y por ende, incapacidad para experimentar placer, aumento de molestias físicas (dolor de cabeza, cansancio, malas digestiones, etc.), apatía, desesperanza (en casos más graves puede llegar al suicidio), los cuales, alteran significativamente áreas de la vida de la persona, provocando cambios en el estado de ánimo, en el pensamiento y la conducta (González et al., 2015; Currin, Haschke, Kluge, Niebuhr & Stüttgen, 2001).

Para el establecimiento de un diagnóstico clínico, los dos principales sistemas de clasificación de los trastornos mentales son: Manual Diagnóstico y Estadístico de los Trastornos Mentales y la Clasificación Internacional de Enfermedades (Hernández-Guzmán, del Palacio, Freyre y Alcázar, 2011)

El Manual Diagnóstico y estadístico de los trastornos mentales en su quinta versión (DSM-V) (American Psychiatric Association, [APA], 2014), establece que los trastornos depresivos incluyen: el trastorno de depresión mayor (incluye el episodio depresivo mayor), el trastorno depresivo persistente (distimia), el trastorno disfórico premenstrual, el trastorno depresivo inducido por una sustancia/medicamento, el trastorno depresivo debido a otra afección médica, otro trastorno depresivo especificado y otro trastorno depresivo no especificado. La

característica en común de todos estos trastornos es la presencia de un ánimo triste, vacío o irritable, acompañado de cambios somáticos y cognitivos que afectan significativamente a la capacidad funcional del individuo, lo que los diferencia es la duración, la presentación temporal o la supuesta etiología.

La Clasificación Internacional de Enfermedades, décima versión (CIE-10; Organización Mundial de la Salud [OMS], 1994) ubica la depresión dentro de la categoría de Trastornos del humor (afectivos) entendidos como un grupo de trastornos cuya alteración fundamental es un cambio en el humor o en la efectividad hacia la depresión (con o sin ansiedad asociada) o la euforia.

La CIE-10 y el DSM-V clasifican y establecen los criterios los trastornos depresivos de la siguiente manera:

Tabla 1

*Criterios diagnósticos para los Trastornos Depresivos*

Manual Diagnóstico y estadístico de los trastornos mentales (DSM-V)		Clasificación Internacional de Enfermedades (CIE-10)	
Trastorno depresivo	Criterios	Trastornos depresivos	Criterios
Trastorno depresivo mayor	<p>A. Cinco o más de los síntomas siguientes han estado presente en un periodo de dos semanas; el estado de ánimo deprimido o la pérdida de interés o de placer deben estar incluidos dentro de los síntomas.</p> <ol style="list-style-type: none"> <li>1. Estado de ánimo deprimido la mayor parte del día, todos los días, por autopercepción de sentimiento de tristeza, vacío o desesperanza, u observada por terceros.</li> <li>2. Disminución importante del interés o placer por todas o casi todas las actividades la mayor parte del día, casi todos los días.</li> <li>3. Pérdida importante de peso o aumento, o bien disminución o aumento del apetito casi todos los días.</li> <li>4. Insomnio o hipersomnia casi todos los días.</li> <li>5. Agitación o retraso psicomotor casi todos los días que puede ser identificado por terceros.</li> <li>6. Fatiga o pérdida de la energía casi todos los días.</li> </ol>	Episodio depresivo	<p><b>Síndrome somático:</b></p> <ol style="list-style-type: none"> <li>1. Pérdida de interés o capacidad para disfrutar actividades que normalmente eran placenteras.</li> <li>2. Ausencia de reacciones emocionales ante acontecimientos que normalmente producen una respuesta emocional.</li> <li>3. Despertarse por la mañana dos o más horas antes de la hora habitual.</li> <li>4. Empeoramiento matutino del humor depresivo.</li> <li>5. Presencia objetiva de enlentecimiento psicomotor o agitación (observada por terceros)</li> <li>6. Pérdida marcada de apetito.</li> <li>7. Pérdida de peso.</li> <li>8. Notable disminución de la libido.</li> </ol> <p><b>Episodio depresivo leve:</b></p> <ol style="list-style-type: none"> <li>A. Deben cumplirse los criterios de episodio depresivo.</li> <li>B. Se encuentran presentes los siguientes síntomas:</li> </ol>

*Criterios diagnósticos para los Trastornos Depresivos*

Manual Diagnóstico y estadístico de los trastornos mentales (DSM-V)		Clasificación Internacional de Enfermedades (CIE-10)	
Trastorno depresivo	Criterios	Trastornos depresivos	Criterios
	<p>7. Sentimientos de inutilidad o de culpabilidad excesiva o inapropiada casi todos los días.</p> <p>8. Disminución de la capacidad para pensar, concentrarse o tomar decisiones casi todos los días, puede ser reportado por el propio individuo o por otros sujetos.</p> <p>9. Pensamientos de muerte e ideas suicidas de forma recurrente, así como intento de suicidio o un plan fraguado para llevarlo a cabo.</p> <p>B. Los síntomas producen malestar clínicamente significativo o deterioro en el área social, laboral u otras importantes para el normal funcionamiento.</p> <p>C. No se puede atribuir a los efectos de una sustancia o de otra afección médica.</p> <p>D. El episodio de depresión mayor no se explica mejor por un trastorno esquizoafectivo, esquizofrenia, esquizofreniforme, delirante u otro.</p> <p>E. Nunca ha habido un episodio maníaco o hipomaníaco.</p>		<p>1. Humor depresivo en grado anormal, la mayor parte del día, cada día y persiste durante 2 semanas.</p> <p>2. Pérdida de interés o de la capacidad para disfrutar actividades que solían ser placenteras.</p> <p>3. Disminución de la energía o aumento de la fatiga.</p> <p>C. Además, deben aparecer uno o más de los siguientes síntomas:</p> <p>1. Pérdida de confianza o disminución de la auto estima.</p> <p>2. Sentimientos no razonables de auto reproche o de culpa.</p> <p>3. Pensamientos recurrentes de muerte o suicidio, o cualquier conducta suicida.</p> <p>4. Indecisión o vacilación.</p> <p>5. Agitación o enlentecimiento de la actividad psicomotriz.</p> <p>6. Alteraciones del sueño.</p> <p>7. Cambio en el apetito y el peso.</p> <p><b>Episodio depresivo moderado:</b> Están presentes cuatro o más de los síntomas arriba descritos y es</p>



*Criterios diagnósticos para los Trastornos Depresivos*

Manual Diagnóstico y estadístico de los trastornos mentales (DSM-V)		Clasificación Internacional de Enfermedades (CIE-10)	
Trastorno depresivo	Criterios	Trastornos depresivos	Criterios
			<p>probable que el paciente tenga dificultades para realizar sus actividades diarias.</p> <p><b>Episodio depresivo grave sin síntomas psicóticos:</b> En el que varios de los síntomas mencionados anteriormente son marcados y afectan al paciente, se presenta pérdida de la autoestima, ideas de autodesprecio o culpa, ideas o actos suicidas y presencia de unos cuantos síntomas somáticos.</p> <p><b>Episodio depresivo grave con síntomas psicóticos:</b> Episodio de depresión tal como se describe en el criterio previamente, pero con la presencia de alucinaciones, ideas delirantes, enlentecimiento psicomotriz o estupor suficientemente graves que imposibilitan las actividades sociales diarias. Alto riesgo de suicidio, deshidratación o inanición.</p> <p><b>Otros episodios depresivos:</b> Depresión atípica y episodios aislados de depresión “enmascarada” sin especificación”</p> <p><b>Episodio depresivo sin especificación:</b> Depresión sin especificación y trastorno depresivo sin especificación.</p>
Trastorno depresivo persistente (distimia)	A. Estado de ánimo deprimido la mayor parte del día, con más presencia que	Trastorno depresivo recurrente	<b>Trastorno depresivo recurrente, episodio actual leve:</b> Caracterizado por episodios

*Criterios diagnósticos para los Trastornos Depresivos*

Manual Diagnóstico y estadístico de los trastornos mentales (DSM-V)		Clasificación Internacional de Enfermedades (CIE-10)	
Trastorno depresivo	Criterios	Trastornos depresivos	Criterios
	ausencia a lo largo de los días por autopercepción o de terceros.		repetidos de depresión, siendo el episodio actual leve, sin antecedentes de manía.
	B. Presencia durante la depresión, de dos o más de los siguientes síntomas:		<b>Trastorno depresivo recurrente, episodio actual moderado:</b> Caracterizado por episodios repetidos de depresión, siendo el episodio actual moderado, sin antecedentes de manía.
	1. Poco apetito o sobre alimentación.		
	2. Insomnio o Hipersomnía.		
	3. Poca energía o fatiga.		
	4. Baja autoestima.		<b>Trastorno depresivo recurrente, episodio actual grave sin síntomas psicóticos:</b>
	5. Falta de concentración o dificultad para tomar decisiones.		Caracterizado por episodios repetidos de depresión, siendo el episodio actual grave sin síntomas psicóticos, sin antecedentes de manía.
	6. Sentimientos de desesperanza.		
	C. Durante el lapso de dos años de padecimiento, el individuo nunca ha estado sin los síntomas de los criterios A y B durante dos meses seguidos.		<b>Trastorno depresivo recurrente, episodio actual grave con síntomas psicóticos:</b>
	D. Los criterios para un trastorno de depresión mayor pueden estar presentes durante dos años.		Caracterizado por episodios repetidos de depresión, siendo el episodio actual grave con síntomas psicóticos, sin antecedentes de manía.
	E. Nunca se ha presentado un episodio maniaco o hipomaniaco y no se han cumplido los criterios para el trastorno ciclotímico.		<b>Trastorno depresivo recurrente actualmente en remisión:</b> El paciente ha tenido dos o más episodios depresivos en el

*Criterios diagnósticos para los Trastornos Depresivos*

Manual Diagnóstico y estadístico de los trastornos mentales (DSM-V)		Clasificación Internacional de Enfermedades (CIE-10)	
Trastorno depresivo	Criterios	Trastornos depresivos	Criterios
	<p>F. La alteración no se explica por un trastorno del espectro de la esquizofrenia u otro trastorno psicótico.</p> <p>G. Los síntomas no se pueden atribuir a efectos fisiológicos de una sustancia u otra afección médica.</p> <p>H. Los síntomas producen malestar clínicamente significativo o deterioro en el área social, laboral u otras significativas para el funcionamiento.</p>		<p>pasado, sin síntomas depresivos durante varios meses.</p>
Trastorno disfórico premenstrual	<p>A. Al menos cinco síntomas han estado presentes en la última semana antes del comienzo de la menstruación, empezar a mejorar unos días después del inicio y hacerse mínimos o desaparecer en la semana después:</p> <p>B. Uno o más de los siguientes síntomas deben estar presentes:</p> <ol style="list-style-type: none"> <li>1. Labilidad afectiva intensa</li> <li>2. Irritabilidad intensa, enfado o aumento de conflictos interpersonales.</li> </ol>	Trastornos del humor (afectivos) persistentes	<p><b>Ciclotimia:</b></p> <p>A. Presencia de un período de al menos dos años en la inestabilidad del estado de ánimo (episodios de depresión e hipomanía), son o sin períodos intermedios de estado de ánimo normal.</p> <p>B. Ninguna manifestación de depresión o hipomanía durante este período es lo suficientemente severa o prolongada para el cumplimiento de episodio maniaco o depresivo, sin</p>

*Criterios diagnósticos para los Trastornos Depresivos*

Manual Diagnóstico y estadístico de los trastornos mentales (DSM-V)		Clasificación Internacional de Enfermedades (CIE-10)	
Trastorno depresivo	Criterios	Trastornos depresivos	Criterios
	3. Estado de ánimo deprimido, desesperanza o ideas de autodesprecio. 4. Ansiedad, tensión y/o sensación intensa de estar excitada. C. Uno o más de los siguientes síntomas deben estar presentes, hasta llegar a un total de cinco síntomas al combinarse con los síntomas del Criterio B. 1. Disminución del interés por las actividades habituales. 2. Dificultad subjetiva de concentración. 3. Letargo, fatigabilidad fácil o intensa falta de energía. 4. Cambio importante del apetito (sobrealimentación o anhelo específico). 5. Hipersomnias o insomnio. 6. Sensación de estar agobiada o sin control.		embargo, pudieron haberse producido antes o después de un período de inestabilidad persistente del estado de ánimo. C. Por lo menos en algunos de los períodos de depresión deben haberse presentado tres o más de los siguientes: 1. Disminución de la energía o de la actividad. 2. Insomnio. 3. Pérdida de la confianza en sí mismo o sentimientos de inferioridad. 4. Dificultad para concentrarse. 5. Aislamiento social. 6. Pérdida de interés o satisfacción por el sexo y otras actividades placenteras. 7. Disminución de la locuacidad. 8. Pesimismo sobre el futuro o cavilaciones sobre el pasado. D. Por lo menos en algunos de los períodos de elevación del estado de

*Criterios diagnósticos para los Trastornos Depresivos*

Manual Diagnóstico y estadístico de los trastornos mentales (DSM-V)		Clasificación Internacional de Enfermedades (CIE-10)	
Trastorno depresivo	Criterios	Trastornos depresivos	Criterios
	<p>7. Síntomas físicos (dolor o tumefacción mamaria, dolor articular o muscular, sensación de hinchazón o aumento de peso).</p> <p>D. Asociación clínicamente significativa de los síntomas o interferencias en diferentes áreas (laboral, académico, sociales. Etc.)</p> <p>E. No es exacerbación de síntomas de otros trastornos (trastorno de depresión mayor, de pánico, depresivo persistente o de personalidad).</p> <p>F. El criterio A se debe confirmar mediante evaluaciones diarias prospectivas por al menos dos ciclos sintomáticos.</p> <p>G. No se atribuye a efectos fisiológicos de una sustancia o afección médica.</p>		<p>ánimo deben haberse presentado tres o más de los siguientes:</p> <ol style="list-style-type: none"> <li>1. Aumento de la energía o de la actividad.</li> <li>2. Disminución de las necesidades de sueño.</li> <li>3. Autoestima exagerada.</li> <li>4. Pensamiento agudizado o excepcionalmente creativo.</li> <li>5. Mayor sociabilidad de los normal.</li> <li>6. Aumento de la locuacidad o del ingenio.</li> <li>7. Optimismo desmesurado o exageración de los logros del pasado.</li> </ol> <p><b>Distimia:</b></p> <ol style="list-style-type: none"> <li>A. Presencia de un período de al menos dos años constantemente recurrente. Los períodos intermedios de estado de ánimo normal raramente duran más que pocas semanas, sin episodios de hipomanía.</li> <li>B. Ninguno o muy pocos de los episodios de depresión debe ser lo</li> </ol>

*Criterios diagnósticos para los Trastornos Depresivos*

Manual Diagnóstico y estadístico de los trastornos mentales (DSM-V)		Clasificación Internacional de Enfermedades (CIE-10)	
Trastorno depresivo	Criterios	Trastornos depresivos	Criterios
			suficientemente severo o prolongado para cumplir criterios de trastorno depresivo recurrente leve.
			C. Presencia de por lo menos tres de los siguientes durante, al menos, alguno de los episodios de depresión:
			1. Disminución de la energía o de la actividad.
			2. Insomnio.
			3. Pérdida de la confianza en sí mismo o sentimientos de inferioridad.
			4. Dificultad para concentrarse.
			5. Llanto fácil.
			6. Pérdida de interés o satisfacción por el sexo y otras actividades placenteras.
			7. Sentimientos de desesperación o desesperanza.
			8. Percepción de incapacidad para afrontar las responsabilidades rutinarias de la vida diaria.
			9. Pesimismo sobre el futuro o cavilaciones sobre el pasado.

*Criterios diagnósticos para los Trastornos Depresivos*

Manual Diagnóstico y estadístico de los trastornos mentales (DSM-V)		Clasificación Internacional de Enfermedades (CIE-10)	
Trastorno depresivo	Criterios	Trastornos depresivos	Criterios
			10. Aislamiento social. 11. Disminución de la locuacidad.
			<b>Otros trastornos del humor (afectivos) persistentes:</b> Es una categoría residual para trastornos del humor (afectivos) persistentes que no son lo suficientemente severos o prolongados para ajustarse a los criterios de ciclotimia o distimia pero que son clínicamente significativos.
Trastorno depresivo inducido por una sustancia/medicamento	<p>A. Alteración importante y persistente del estado de ánimo caracterizada por un estado de ánimo deprimido, disminución del interés o placer por todas o casi todas las actividades.</p> <p>B. Existen evidencias clínicas de:</p> <ol style="list-style-type: none"> <li>Síntomas del criterio A desarrollados durante o después de la intoxicación o abstinencia de una sustancia o después de la exposición a un medicamento.</li> <li>La sustancia/medicamento implicado puede producir los síntomas del criterio A.</li> </ol>	Otros trastornos del humor (afectivos)	Cualquier otro trastorno del humor (afectivo) que no esté justificada su clasificación, debido a su poca gravedad o duración

*Criterios diagnósticos para los Trastornos Depresivos*

Manual Diagnóstico y estadístico de los trastornos mentales (DSM-V)		Clasificación Internacional de Enfermedades (CIE-10)	
Trastorno depresivo	Criterios	Trastornos depresivos	Criterios
	<p>C. El trastorno es explicado por un trastorno depresivo no inducido por sustancia/medicamento.</p> <p>D. El trastorno no se produce durante el curso de un síndrome confusional.</p> <p>E. El trastorno produce malestar clínicamente significativo o deterioro en áreas significativas del funcionamiento.</p>		
Trastorno depresivo debido a otra afectación medica	<p>A. Un periodo importante y persistente de estado de ánimo deprimido o una disminución del interés o placer en todas o casi todas las actividades.</p> <p>B. Existen pruebas clínicas de que el trastorno es la consecuencia fisiopatológica directa de otra afectación médica.</p> <p>C. La alteración no se explica por otro trastorno mental.</p> <p>D. El trastorno no se produce exclusivamente durante el curso de un síndrome confusional.</p>	Otros trastornos del humor (afectivo) sin especificación	Psicosis afectiva sin especificación.



*Criterios diagnósticos para los Trastornos Depresivos*

Manual Diagnóstico y estadístico de los trastornos mentales (DSM-V)		Clasificación Internacional de Enfermedades (CIE-10)	
Trastorno depresivo	Criterios	Trastornos depresivos	Criterios
	E. El trastorno produce malestar clínicamente significativo o deterioro en áreas significativas del funcionamiento.		
Otro trastorno depresivo especificado	Esta categoría se presenta predominancia de síntomas característicos de un trastorno depresivo que causan malestar clínicamente significativo pero que no cumple todos los criterios de ninguna de las categorías de los trastornos depresivos. El clínico comunica el motivo por el cual no cumple los criterios.		
Otro trastorno depresivo no especificado	Esta categoría se presenta predominancia de síntomas característicos de un trastorno depresivo que causan malestar clínicamente significativo pero que no cumple todos los criterios de ninguna de las categorías de los trastornos depresivos. El clínico no comunica el motivo por el cual no cumple los criterios.		

*Nota.* Recuperado de Guía de consulta de los criterios diagnósticos del DSM-5. Copyright © 2014 Asociación Americana de Psiquiatría y Guía de consulta de los criterios diagnósticos de la Guía de bolsillo de la CIE-10 © 2000 EDITORIAL MÉDICA PANAMERICANA, S. A.

### **Teorías explicativas de la depresión**

La depresión se caracteriza por un estado de tristeza patológico que se da a partir del entrecruzamiento de factores predisponentes y precipitantes, los primeros, abarcan la vulnerabilidad psicológica y la genética, que podrían influir en el desencadenamiento de un cuadro clínico depresivo, mientras que, los factores precipitantes hacen referencia a las situaciones del entorno que podrían activar la vulnerabilidad del individuo (Beato, 1993).

En otras palabras, en el desarrollo de la depresión, intervienen una serie de variables de diversos orígenes, de tipo psicológicos, psicosociales, ambientales y biológicos (bioquímico, genéticos, cerebral) que influyen en su aparición (Álvarez, 2016; Dávila, González, Liangxiao y Xin, 2016; González et al., 2015; Campagne, 2012; Guadarrama, Escobar y Zhang, 2006) y la postura de los autores varía de acuerdo a la relevancia dada a cada una de ellas, de ahí, el surgimiento de las diferentes teorías explicativas (Beato, 1993).

**Perspectiva biológica de los trastornos depresivos.** El trastorno depresivo es concebido como una respuesta al estrés psicosocial con la incapacidad de afrontamiento y como un trastorno biológico, de esta manera, la depresión se considera de dos formas básicas exógenas y endógenas (Guadarrama et al., 2006).

La depresión exógena (o reactiva) parte de una causa externa generalmente bien definida (pérdida de un familiar/ser amado/monetario/posición social, enfermedad invalidante, etc.), mientras que la depresión endógena no es generada por una situación externa, lo que lleva a considerar una alteración biológica en el individuo (González et al., 2015; Campagne, 2012; Guadarrama et al., 2006).

En la cotidianidad, existen diversos estímulos que pueden generar depresión, los cuales, construyen una carga estresora de valor afectivo en la persona afectada (que varía en cada individuo), en consecuencia, el estrés es un elemento importante para la generación de la depresión exógena y que no pueden separarse de los cambios biológicos (fisiológicos y hormonales) que normalmente se asocian al estrés, y todos ellos se relaciona con el eje hipotálamo-hipófisis-suprarrenal (González et al., 2015; Guadarrama et al., 2006).

Diversos estudios han demostrado que, neurotransmisores como la serotonina, la noradrenalina, la dopamina, la colina, la histamina, el ácido gammaaminobutírico (GABA) o el glutamato, están relacionados con la depresión, así como, la implicación de factores endocrinos como la alteración de la hipersecreción de cortisol, disminución de la hormona tiroidea, alteración en la secreción de la hormona de crecimiento, hiperplasia, hipofisaria y suprarrenal (Álvarez, 2016; González et al., 2015; Campagne, 2012).

A partir de esto, se puede inferir que, la depresión exógena depende de los estímulos ambientales estresores (en calidad y cantidad) y de los substratos biológicos (genéticos, bioquímicos y moleculares) que determinan las alteraciones en la homeostasis y, por ende, en la función cerebral.

**Teorías conductuales de los trastornos depresivos.** El enfoque conductual ha planteado una serie de teorías explicativas enmarcadas siempre en la conducta observable de la depresión, evitando tratar cualquier proceso interno como elemento explicativo de la misma (Beato, 1993). Estos postulados toman como base las teorías del aprendizaje, especialmente el condicionamiento operante, enfatizándose en la idea original de Skinner (1953) que la principal característica de la depresión es una *reducción generalizada en la frecuencia de las conductas*.

Charles Ferster (1965, 1973) fue uno de los primeros psicólogos en estudiar la depresión dentro de una perspectiva clínico/experimental, estableció que la principal característica de la depresión es la reducida frecuencia de conductas reforzadas positivamente que sirven para controlar el medio, mientras que, por el contrario, se produce un exceso de conductas de evitación ante estímulos aversivos, plantea los siguientes procesos (ya sea aislados o en combinación) para explicar la patología depresiva:

- Cambios inesperados y rápidos en el medio que supongan pérdidas de fuentes de refuerzos o de estímulos discriminativos importantes para el individuo. Muchas conductas están concadenadas a una(s) fuente de reforzamiento, y tal asociación explicaría por qué en la depresión, tras la pérdida de una fuente de reforzamiento, se produce una *generalización* de efectos a conductas no relacionadas a esta fuente (por ejemplo, pérdida de apetito, de interés en el sexo, aislamiento social, etc.).
- Programas de refuerzo de gran coste, que requieren grandes cambios de conducta para producir consecuencias en el medio.
- Imposibilidad de desarrollar repertorios conductuales debido al surgimiento de *espirales patológicas*, donde una disminución de refuerzos positivos produce una reducción de conductas (siendo a su vez una mala adaptación al medio), y, por tanto, un menor número de refuerzos positivos que producirían una nueva reducción conductual.
- Repertorios de observación limitados que llevarían a los depresivos a distorsionar la realidad y a conductas que son reforzadas en pocas ocasiones.

Por su parte, Costello (1972) señala que la principal característica de la depresión es la pérdida de interés general (apetito, deseo sexual, etc.), estos rasgos se explican por una *pérdida general de la efectividad de los reforzadores* que serían la causa suficiente para la depresión,

siendo los siguientes factores (aislados o combinados) los responsables de esta pérdida en general:

- cambios endógenos bioquímicos y neurofisiológicos.
- pérdida de uno o varios de los estímulos condicionados, teniendo en cuenta que la efectividad de los reforzadores de una cadena conductual está supeditada a la realización completa de dicha cadena conductual.

***La disminución del refuerzo positivo.*** Lewinsohn y sus colaboradores (1974, 1976) desarrollaron el modelo conductual más conocido denominado disminución del refuerzo positivo, el cual establece una teoría coherente para su aplicación en la práctica clínica.

Desde esta mirada, la depresión es entendida como una respuesta a la *pérdida o falta de refuerzos positivos contingente a la conducta*. Un reforzamiento insuficiente en los principales dominios vitales de una persona produciría un estado depresivo y una disminución en el número de conductas (fenómenos primarios de la depresión); la autoestima y la desesperanza serían la consecuencia lógica de la disminución del nivel de actividad. La falta de refuerzo positivo en la conducta, de acuerdo con Lewinsohn, sería la causa de la depresión que, podrían deberse a ciertos factores (aislados o en combinación), tales como:

- un ambiente que no proporciona el reforzamiento suficiente.
- falta de habilidades sociales para obtener reforzamiento de un ambiente dado.
- incapacidad de disfrutar de los reforzadores disponibles debido a los altos niveles de ansiedad social que, al interferir con las habilidades sociales, le impiden al individuo acceder a los refuerzos sociales.

Para Lewinsohn y colaboradores, la reformulación de la teoría de disminución de refuerzo positivo incluye el fenómeno de la secuenciación de conexiones causales, que inicia cuando un acontecimiento potencialmente evocador de depresión interrumpe los patrones de conductas automáticos, provocando así, una disminución de la tasa de refuerzo positivo, y como consecuencia, da lugar a un aumento de la conciencia de uno mismo (la autocrítica y las expectativas negativas), generando un aumento de la disforia con consecuencia de tipo conductual, cognitivas, interpersonales, etc., además, propusieron una serie de factores que podrían incrementar la probabilidad de depresión (factores de vulnerabilidad) tales como: ser mujer, tener una edad entre 20 y 40 años, historia previa de depresión, susceptibilidad frente a acontecimientos aversivos, bajo estatus socioeconómico, baja autoestima y tener hijos con edades inferiores a 7 años. También plantearon factores de protección frente a la depresión: capacidad de iniciativa, competencia social autopercebida, exposición a una frecuencia elevada de acontecimientos positivos (producidos en la mente del individuo o por el medio ambiente) y un grado elevado de apoyo social (Lewinsohn, Hoberman, Teri, y Hautzinger, 1985).

*El modelo de Nezu.* Nezu y sus colaboradores (1987, 1989) plantearon que, la ineficiencia de las habilidades de solución de problemas representa el comienzo y el mantenimiento de la depresión. Este argumento se sustenta en trabajos que demostraban la asociación entre déficits en la solución de problemas y sintomatología depresiva (en niños y adultos) (Nezu, Nezu, y Perri, 1989; Nezu, 1987). De esta manera, Nezu y Ronan (1985) sugirieron que las habilidades de solución de problemas podían moderar la relación entre el estrés y la depresión, solucionar problemas eficazmente podría ser positivo a los efectos de las experiencias negativas, por ello, los pacientes que tengan déficit en este aspecto tendrían un mayor riesgo para el desarrollo de un episodio depresivo, la depresión podría ser el resultado de

la carencia de los siguientes componentes de estrategias de solución de problemas: orientación hacia el problema, definición y formulación del problema, generación de soluciones alternativas, toma de decisiones, y puesta en práctica de la solución y verificación.

***Teoría de la indefensión aprendida.*** Seligman (1975) estudió los efectos de choques eléctricos inescapables en animales, descubrió que desarrollaban un patrón de conductas y cambios neuroquímicos similares a los deprimidos, en una explicación cognitiva, los animales desarrollan estas conductas solo cuando no tienen esperanza de poder controlar la situación aversiva en la que se encuentran. Seligman (1975) extrapoloó este modelo en el estudio de la conducta humana y postuló que la pérdida percibida de control del ambiente o expectativa de incontrolabilidad (aprendizaje del individuo de la existencia entre sus respuestas y las consecuencias ambientales) es causa suficiente de la aparición de una reacción depresiva en humanos y tal expectativa es producto de dos factores: a) una historia de fracasos en el manejo de situaciones, y b) una historia de reforzamientos sobre una base no contingente que no haya permitido que la persona aprenda las complejas aptitudes necesarias para controlar el ambiente.

***Teoría reformulada de la indefensión aprendida.*** Abramson, Seligman y Teasdale (1978) sugirieron que la teoría de la indefensión presentaba cuatro problemas: 1. No explica la baja autoestima típica de la depresión; 2. La teoría no daba dé cuenta de la autoinculpa presente en los depresivos de sus fallos; 3. No explica la cronicidad y generalidad de los déficits depresivos; y 4. No explicaba satisfactoriamente el estado de ánimo deprimido como síntoma. Para dar respuesta a estos problemas, Abramson y colaboradores (1989) postulan que las situaciones incontrolables por sí misma no desencadenan reacciones depresivas (teniendo en cuenta las teorías atribucionales de la psicología social) , sino que en efecto, al experimentarlas la persona intenta darse explicación sobre su causa, y si se atribuye la falta de control a:

- Factores internos: da lugar a un descenso de la autoestima
- Factores externos: no hay baja autoestima
- Factores estables: conduce a una expectativa de incontrolabilidad en situaciones futuras y, en consecuencia, la indefensión y los déficits depresivos se prolongan.
- Factores globales: produce expectativas de incontrolabilidad en otras situaciones, por tanto, la indefensión y los déficits depresivos se generalizan a otras situaciones.
- Factores inestables y específicos: supondrían sentimientos de indefensión y déficits depresivos de corta duración específicos a la situación en cuestión.

Este análisis atribucional (internalidad-externalidad, estabilidad-inestabilidad y globalidad-especificidad) responden a los tres primeros problemas; mientras que para resolver el último problema postularon el factor motivacional: la depresión solo aparece cuando la expectativa de incontrolabilidad hacía referencia a la pérdida de control de un suceso altamente deseable o de un hecho altamente aversivo.

La teoría de la indefensión aprendida surge después de la teoría cognitiva de Beck, y con base a sus postulados, es un modelo que funciona como puente entre el modelo explicativo conductista y el modelo cognitivo (Beato, 1993).

**Teorías cognitivas de los trastornos depresivos.** Las teorías cognitivas de la depresión se centran en la percepción errónea de uno mismo, de los otros y de los sucesos, que puede conducir a un trastorno depresivo. Ingram y Wisnicki (1991) delimitan tres clases de teorías cognitivas de la depresión: teorías del procesamiento de la información, teorías cognitivo-sociales y teorías conductuales-cognitivas.



*Teorías del procesamiento de la información.* Estas teorías se centran en el paradigma del procesamiento de la información y los constructos teóricos provenientes de la psicología experimental cognitiva. El hombre se concibe como un sistema de procesamiento de información (metafóricamente) donde se desarrolla en un ambiente continuamente cambiante que demanda capacidades de adaptación, y construye la realidad mediante la selección, codificación, almacenamiento y recuperación de la información. La información sobre sí mismo y el medio experimenta transformaciones cognitivas y afecta en las respuestas emocionales y conductuales, provocando así, una interdependencia entre cognición, emoción y conducta (Sanz & Vásquez, 2009).

*Teoría cognitiva de Beck.* La teoría cognitiva de Aaron Beck (1963, 1964, 1976, 1987) surgió en el campo de la psicología clínica, centrándose en el papel activo de la persona y los factores ambientales, por lo que, desde esta mirada, la depresión se concibe como el resultado de una interpretación errónea de las situaciones que debe afrontar el individuo. Esta teoría gira en torno a tres conceptos básicos: la cognición, el contenido cognitivo y los esquemas.

- cognición: es el contenido, proceso y estructura de la percepción, del pensamiento y de las representaciones sociales mentales.
- contenido cognitivo: es el significado que le da el sujeto a sus pensamientos e imágenes mentales.
- los esquemas: son creencias centrales que son los principios directrices que sirven para procesar la información procedente del medio.

Desde esta teoría, en los trastornos depresivos se presenta una *distorsión o sesgo sistemático* en el proceso de la información de sucesos que implican una pérdida o privación, y que conduce a los síntomas afectivos, conductuales, motivacionales y fisiológicos; estas

experiencias negativas, proporcionan estilos cognitivos alterados que hacen que la persona se vea de forma negativa a sí misma, a las experiencias y el futuro (triada cognitiva) y que lo lleva a su vez, a un estado de baja autoestima y una expectativa de desesperanza; este procesamiento distorsionado es una parte intrínseca del síndrome depresivo y factor responsable de su mantenimiento, sin importar su etiología (factores genéticos, evolutivos, hormonales, físicos y psicológicos), sin embargo, la teoría de Beck afirma que la etiología de la depresión no endógena tendría que ver con la interacción de los siguientes tres factores (Beck y Clark, 1988):

- la presencia de actitudes o creencias disfuncionales sobre el significado de ciertas clases de experiencias, actitudes que impregnan y condicionan la construcción de la realidad.
- una alta valoración subjetiva de la importancia de esas experiencias que resulta de la estructura de personalidad del individuo.
- la ocurrencia de un estresor específico a los anteriores factores, es decir, un suceso considerado importante y que incide directamente sobre las actitudes disfuncionales del individuo.

Beck se centra en los esquemas para explicar cómo las creencias disfuncionales están representadas mentalmente y afectan el procesamiento de la información; los esquemas podrían ser el resultado del aprendizaje a través de situaciones anteriores, jugando un papel importante en las experiencias tempranas y la identificación con personas significativas (Beck y Clark, 1988)

En los individuos depresivos, los esquemas responsables del procesamiento sesgado de la información son los esquemas disfuncionales depresógenos, para la activación de estos, es necesario la aparición de un acontecimiento estresante similar a sucesos que influyeron en la formación durante el proceso de socialización de los esquemas (infancia) y también va determinado por los rasgos de personalidad individuales (esquemas cognitivos), una vez dada su

activación, orientan y canalizan el procesamiento de información, actuando como filtro a través de los cuales se percibe, interpreta, sintetiza y recuerda la realidad (Beck y Clark, 1988).

Beck (1996) propone que la activación de ciertos esquemas cognitivos constituye el problema central de la depresión y que estos a su vez, podrían considerarse como los responsables de la activación de distintos síntomas cognitivos, afectivos y conductuales; así mismo, propone que el objetivo de la intervención es enfocarse en la modificación de las interpretaciones disfuncionales para disminuir el trastorno clínico.

La noción de modos es entendida como suborganizaciones específicas contenidas en la organización de la personalidad que retoman elementos de los sistemas básicos de la misma (cognitivo, afectivo, motivacional y conductual), que a su vez se componen por esquemas, es decir, el sistema afectivo contiene esquemas afectivos, el sistema cognitivo contiene esquemas cognitivos, el sistema motivacional contiene esquemas motivacionales y el sistema conductual contiene esquemas conductuales, en este sentido, el modo es la integración de la red de componentes afectivos, cognitivos, motivacionales y conductuales, el cual se activa por las demandas exógenas y proporciona el mecanismo como respuesta para el logro de los objetivos (Beck, 1996).

Existen dos tipos de modos: los primarios que constituyen patrones básicos para la supervivencia y su contenido pueden incluir miedo, ansiedad, impulso de huida, y los modos que se activan en situaciones de la cotidianidad (Beck, 1996).

A continuación, brevemente se explica los sistemas básicos:

- Sistema cognitivo: se compone por unas series de estructuras que implica el procesamiento de la información, la asignación de significados, selección de datos,

atención, interpretación, memoria y evocación de recuerdos y que se consolida por las interacciones sociales. Su estructura básica es el esquema cognitivo.

- Sistema afectivo: se concibe como una parte fundamental y estratégica de la psicobiología para la supervivencia y la procreación que produce estados emocionales y se encarga del reforzamiento de las conductas adaptativas o desadaptativas. Su estructura básica es el esquema afectivo.
- Sistema motivacional y Sistema conductual: constituyen los mecanismos automáticos de movilización que comprende estrategias de emergencias para emprender la huida. La motivación es entendida como los impulsos biológicos que involucra la sexualidad, el apetito, etc. El sistema conductual, por su parte, actúa independientemente de la intención consciente y el impulso es anticipado por una alerta consciente. Sus estructuras básicas son los esquemas motivacionales y esquemas conductuales respectivamente.
- Sistema fisiológico: su activación ocurre cuando un modo responde ante una situación amenazante, incluye el sistema nervioso central, el sistema nervioso periférico, el sistema nervioso automático, el sistema nervioso motor y el sistema nervioso sensorial. Es un sistema separado dado que cumple una única función cuando se activan los modos.

*Hipótesis de la activación diferencial de Teasdale.* John Teasdale (1983, 1988) desarrolló la teoría denominada *hipótesis de la activación diferencial* a partir de la teoría cognitiva de Beck y el modelo de redes asociativas propuesto por Bower para explicar la relación entre emoción y cognición. El modelo de Bower (1981) establece que cada emoción está representada en la memoria por una unidad o *nodo* (constructo cognitivo), el cual está conectada a un conjunto de rasgos asociados a esa emoción (por ejemplo, la experiencia subjetiva, el patrón característico de respuestas autonómicas, los términos verbales que la describen y cogniciones que contienen

descripciones de los sucesos que evocan la emoción, pensamientos, creencias y recuerdos que se han almacenado en la memoria producto de experiencias tempranas de aprendizaje que han asociado el nodo de esa emoción con dichas cogniciones). La teoría de la activación diferencial Teasdale (1983, 1988) es un modelo de vulnerabilidad-estrés en el que un tipo de situación que provoque una depresión clínica es capaz de producir un estado de ánimo deprimido transitorio o leve; asume que el inicio de los síntomas depresivos resulta de la activación del nodo como consecuencia de un suceso estresante, si los nodos reactivan el nodo depresivo a través de un mecanismo de tipo cíclico (bucles cognitivos) y se establece un círculo vicioso (relación recíproca entre el estado de ánimo deprimido y el procesamiento cognitivo negativo), entonces el estado depresivo inicial se intensificará o se mantendrá, tal ciclo se establecerá si los nodos activados *interpretan* los sucesos como altamente aversivos e incontrolables; estas interpretaciones dependen de interacciones entre diversos factores: ambientales, biológicos y psicológicos.

En la reformulación de su hipótesis, Teasdale incluye una reconceptualización de *Kindling* (*activación inducida propagada*) y *sensibilización* (en términos cognitivos) para explicar la susceptibilidad de episodios depresivos a sujetos con historia de depresión.

***Teorías cognitivo-sociales.*** Las teorías cognitivo-sociales se centran en los constructos desarrollados en las teorías de aprendizaje social y de la psicología social.

*Teoría de la desesperanza.* Abramson, Metalsky y Alloy (1989) publicaron una revisión de la reformulación para resolver sus principales deficiencias:

- no presenta una teoría explícitamente articulada de la depresión.

- no incorpora los hallazgos de la psicopatología descriptiva acerca de la heterogeneidad de la depresión.
- no incorpora de forma plena descubrimientos significativos obtenidos por la psicología social, de la personalidad y cognitiva.

Para el primer inconveniente, la teoría explicita un modelo de diátesis-estrés y especifica una cadena de causas distantes y próximas que incrementan la probabilidad de depresión y termina con la desesperanza, cuya cadena causal empieza con la ocurrencia de *sucesos vitales negativos*, los cuales, dependiendo de las atribuciones y el grado de importancia que se les otorguen, contribuye al desarrollo de la desesperanza y a los síntomas de la depresión por desesperanza, cuando los sucesos vitales negativos se atribuyen a factores estables y globales y se ven como importantes la probabilidad de síntomas es mayor, y, además, si interviene la internalidad, la desesperanza puede acompañarse (aunque no necesariamente) de una baja autoestima (Abramson et al., 1989).

Para resolver el segundo inconveniente, se postula la *depresión por desesperanza* como nueva categoría nosológica de los trastornos depresivos, cuya causa suficiente y próxima es la desesperanza: expectativa negativa sobre la ocurrencia de un suceso valorado como muy importante vinculado a sentimientos de indefensión sobre la posibilidad de cambiar la probabilidad de ocurrencia de tal suceso (Abramson et al., 1989).

Para el tercer problema, Abramson y colaboradores (1989), toman de la psicología social la importancia de la *información situacional* al momento de determinar el tipo de atribuciones que las personas hacen. La información situacional sugiere que un suceso negativo es bajo en consenso, alto en consistencia y bajo en distintividad que favorece una explicación atribucional conducente a la desesperanza.

Aspectos generales de la teoría de la desesperanza:

- No requiere que ninguno de los elementos distantes del modelo esté presente para desencadenar la cadena causal depresiva, esta puede activarse por elementos, dependiendo de las circunstancias, las personas, etc. La desesperanza es el único elemento que se requiere para la aparición de la depresión por desesperanza.
- Además de las atribuciones, la desesperanza puede ser producto de las conclusiones que una persona llega a tener sobre las consecuencias del suceso negativo. Sin embargo, existe la probabilidad que otros elementos sean antecedentes de la desesperanza.
- No incluye síntomas de depresión por desesperanza del tipo de errores cognitivos de Beck. Su exclusión se debe a que los autores han descubierto la posibilidad de las personas depresivas son más precisos en su visión de la realidad que los no deprimidos, denominado como “realismo depresivo”.

*Teoría de los estilos de respuesta.* Susan Nolen-Hoeksema (1991) propuso la *teoría de los estilos de respuestas* enfocada en los factores que explican el curso de la depresión, establece que la forma en que la persona responde a los primeros síntomas depresivos influye en su duración y gravedad, de tal manera que, las personas cuyas respuestas sean rumiativas sufrirán durante más tiempo y aquellas que sean capaces de distraerse serán más intensos los síntomas, dichas *respuestas rumiativas de la depresión* hacen referencia a las conductas y pensamientos focalizados en la atención de una persona en sus síntomas y en las implicaciones de estos, que incluye conductas como: aislarse para pensar sobre los síntomas depresivos que experimentan, hablar todo el tiempo sobre ellos y, pensar recurrentemente y contemplativa sobre las posibles causas y consecuencias.

Las respuestas ruminativas difieren de otras conductas y pensamientos de afrontamiento como las estrategias de solución de problemas (no hacen nada para aliviar los síntomas o buscar una solución); la teoría también establece que las personas tienen *estilos de respuestas ante la depresión*, que implica patrones de conductas y pensamientos manifestada de forma consistente y estable en diferentes situaciones, además, existen aquellos que manifiestan patrones de conductas y pensamientos de distracción, que implican intentos activos de ignorar los síntomas a través de actividades neutras o placenteras (Nolen-Hoeksema, 1991).

Nolen-Hoeksema (1991, 2000), propone mecanismos psicológicos para explicar el efecto del estilo ruminativo de respuesta en el mantenimiento o exacerbación de los síntomas depresivos, que son:

- las rumiaciones intensifican los efectos del estado deprimido sobre las cogniciones, haciendo que los pensamientos negativos sobre el pasado, el presente y el futuro estén más accesibles y sobresalientes, los cuales, exacerbaban el estado deprimido y los demás síntomas al establecerse los círculos viciosos (que establece Teasdale), y que hacen que el estado inicial se agrave y persista durante más tiempo.
- la actividad ruminativa reduce la generación de soluciones afectivas a los problemas.
- las respuestas ruminativas interfieren la puesta en marcha de conductas instrumentales que permiten obtener refuerzo positivo, aumentar sensación de control o elevar el estado de ánimo.
- el estilo ruminativo disminuye el efecto beneficioso del apoyo social.

La teoría sugiere que el origen de los estilos ruminativos son aprendidos, fundamentalmente durante la infancia por conductas de los padres y de otras figuras importantes



en la vida del niño(s), tal aprendizaje tendría una doble vía: 1. Mediante el modelado de los estilos de respuesta de los padres y otras figuras relevantes ante la depresión, y 2. A través de la práctica de socialización, las cuales, se enseñan repertorios de conductas más adaptativas para manejar las emociones negativas (estrategias eficaces de distracción o de solución de problemas).

### ***Teorías conductuales-cognitivas***

*Teoría de autocontrol.* Rehm (1977) desarrolló un modelo de autocontrol para la depresión en el que integra elementos cognitivos y conductuales, entendiendo la depresión como el resultado negativo de la relación entre recompensa y castigos, la persona depresiva es de tendencia perfeccionista (pretende alcanzar metas inalcanzables), se centra en aspectos negativos del entorno, se aplican pocos refuerzos y aumenta la tasa de autocastigo; esta teoría se centra en la contribución del individuo de esta relación negativa a través de respuestas conductuales y cognitivas; el entrenamiento consiste en una combinación de la consecución progresiva de la meta, autoreforzo, manejo de contingencias y rendimiento conductual.

*Teoría de autofocalización.* Lewinsohn, Hoberman, Teri y Hautzinger (1985) proponen que, los factores ambientales son los responsables primarios de la depresión, sin embargo, hay factores cognitivos que median los efectos del ambiente, tal variable mediadora sería un aumento de la *autoconciencia* (aumento transitorio y situacional en la autofocalización o la cantidad de atención que dirige una persona hacia sí mismo en lugar de hacia el ambiente) como resultado de los esfuerzos infructuosos del individuo de enfrentarse y resolver los acontecimientos estresantes que le ocurren.

De acuerdo con Lewinsohn et al. (1985), la cadena de acontecimientos que conduce a la aparición de un episodio depresivo empieza con la aparición de un suceso antecedente (cualquier suceso que incremente la probabilidad de depresión a futuro), estos acontecimientos inician el proceso depresógeno en cuanto *interrumpen importantes patrones adaptativos de conducta* del individuo (patrones conductuales relativamente automáticos necesarios para las interacciones diarias del individuo y el ambiente), siendo incapaz de desarrollar otros patrones conductuales que reemplacen los anteriores. La presencia de estresores y la interrupción de conductas generan una *reacción emocional negativa* o estado de ánimo deprimido, su intensidad depende de la importancia del suceso y/o del nivel de interrupción de la conducta cotidiana, tal interrupción y malestar emocional se relacionan con la depresión en la medida que conducen a un *desequilibrio negativo en la calidad de las interacciones* de la persona con el ambiente, dicho desequilibrio se concreta en una reducción del reforzamiento positivo y en un aumento en la tasa de experiencias aversivas, la respuesta emocional negativa (producto de los sucesos estresores) y el impacto emocional negativo (generado por el fracaso del individuo en anular el efecto del estrés) desencadenan un estado elevado de *autoconciencia*, siendo el factor que media los factores de la reducción del reforzamiento positivo en la depresión (atribuciones internas, autodespreciación, autoculpación, expectativas negativas, etc.) que lleva a consecuencias conductuales negativas (retirada conductual, dificultades sociales) e intensificando las reacciones emocionales negativas anteriores, dicho esto, el incremento de autoconciencia y la intensificación de las emociones negativas destruyen las autopercepciones (sirven para proteger y ensalzar la autoestima) que conduce a una serie de cambios cognitivos, emocionales y conductuales, estos cambios exacerbaban la autoconciencia permitiendo un mantenimiento y agravamiento del estado depresivo.

La teoría asume la existencia de características de predisposición (vulnerabilidades e inmunidades) que incrementan o reducen el riesgo de un episodio depresivo, tales como: ser mujer, tener 20-40 años, tener una historia previa de depresión, tener pocas habilidades de afrontamiento, tener una sensibilidad elevada a los sucesos aversivos, ser pobre, mostrar una alta tendencia a la autoconciencia, tener baja autoestima, tener un bajo umbral para la activación de autoesquemas depresógenos, mostrar dependencia interpersonal y tener niños menores de 7 años; también plantean factores protectores contra la depresión: autoperibirse como poseedor de una alta competencia social, experimentar una alta frecuencia de sucesos positivos y disponer de una persona íntima y cercana a la que poder confiarse (Lewinsohn et al., 1985).

### **Determinantes psicosociales de la depresión en la edad adulta**

La adultez es la etapa comprendida entre los 20 y 60 años aproximadamente. Sin embargo, estos límites no son fijos, dado que pueden variar según las culturas (Amador, Monreal y Marco, 2001). Esta etapa se caracteriza por el desarrollo biológico y psíquico consolidándose la personalidad y el carácter, es el estadio de la asunción de responsabilidades: procreación y cuidado de la familia, educación y transmisión de las pautas normativas de la sociedad, trabajo, producción en el sistema económico, dirección y gestión de la sociedad política y búsqueda de elementos trascendentes (Gabarda, s.f.), que implica un proceso de crecimiento y desarrollo, a nivel intelectual, social y personal (Zapata, Cano y Moyá, 2002).

Durante el surgimiento de esta etapa muchas personas no están preparadas para asumir las típicas tareas de la juventud (llevar una vida independiente, obtener un empleo, fundar una familia, etc.); a los 30 años, casi todos han cumplido con esas tareas; en la adultez media probablemente haya una disminución de las capacidades físicas, se enfrentan a grandes desafíos en cada cambio de la vida (inicio de una nueva profesión o la partida de hijos adultos) o cuidar a

padres ancianos; y en la adultez tardía, se presenta la pérdida de las facultades, muerte de seres queridos y preparación para el fin de la vida (Papalia, Wendkos y Duskin, 2009).

La adultez es una etapa que se caracteriza por la madurez en la personalidad, y, por tanto, se asume como una persona responsable, independiente y con la capacidad de adaptarse a los cambios que demanda el medio, la ubicación de las etapas que transita el adulto puede variar, en general se plantean tres fases que se compendian en esta etapa: adultos jóvenes o adultos primarios, adultos o adultos intermedios y adultos mayores o adultos tardíos (Mansilla, 2000). La tabla 2 describe brevemente las características principales de cada etapa de la adultez.

Tabla 2

*Etapas de la adultez*

Etapa	Descripción
Adulthood temprana (20 a 40 años)	Caracterizado como un período de oportunidades, dado que muchas personas por primera vez son autosuficientes, asumen cargos, ponen a prueba actividades que eligieron, toman decisiones entorno a su salud, su profesión y estilo de vida, sin embargo, muchos adultos toman el período entre los 20 y 25 años como una etapa de exploración, denominada <i>adulthood emergente</i> propuesta por Jeffrey Arnett (2000), comprendida como un período transitorio entre 18 a 25 años, que solo existe en aquellas culturas que permiten que los jóvenes prolonguen la etapa de independencia y exploratoria, por lo tanto, al igual que la adolescencia, es un período de la vida que es culturalmente construido, no universal. Desde una postura subjetiva, el adulto emergente no se visualiza a sí mismo como un adolescente ni como un adulto completamente y alcanzar la adulthood requiere de la obtención de ciertas características individuales: la aceptación responsable de sí mismo, la toma de decisiones y convertirse financieramente independiente (Arnett, 2000).
Adulthood media (40 a 65 años)	En esta etapa, las personas consolidan sus roles sociales y profesionales; el trabajo y la familia son las principales preocupaciones y prioridades (Gabarda, s.f.), por ello, se enfrentan a cambios familiares y de pareja, producido en parte (o detonante) por el proceso de cambio de los estereotipos sexuales que han dado paso a las nuevas formas de interacción entre sexos y generaciones (Mansilla, 2000). A nivel biológico se produce un descenso de las habilidades físicas y sensoriales, las mujeres entran a la pre-menopausia, aparecen las canas, la pérdida del cabello y la flacidez muscular, los cuales, son considerados negativos (especialmente en el hombre) debido a que se contraponen al modelo actual de “juventud” especialmente en el campo laboral, lo que deriva problemas psico-sociales (Mansilla, 2000). A nivel psicológico, hay una mayor estabilidad y menos vulnerabilidad a las presiones externas (Gabarda, s.f.).
Adulthood tardía (65 en adelante)	Se caracteriza por el afrontamiento a los cambios biológicos, al alejamiento de los hijos (síndrome del nido vacío) ya sea por estudios o por la formación de nueva familia, en este período la mayoría enfrenta la posibilidad de dejar de trabajar (jubilarse) (Mansilla, 2000).

*Nota.* Fuente: Arnett, 2000; Gabarda, s.f Mansilla, 2000

En la adultez, los cambios psicológicos tienen una relevancia esencialmente funcional, dado que, los avances van dirigidos a optimizar los recursos para adaptarlos saludablemente a las demandas de la vida adulta, hay una consolidación del pensamiento abstracto, conceptual y lógico- matemático, se alcanza una madurez neurobiológica (producidas en la infancia y adolescencia), en el que, los cambios en la adultez son más de tipo funcional que estructural, por ende, conlleva a un proceso de optimización del funcionamiento de las estructuras ya construidas (Brioso, Contreras, Corral, Delgado, Díaz, Giménez, Gómez, Oliva y Sánchez, 2009).

Entre los avances cognitivos se presenta un aumento de los conocimientos de carácter declarativo, lo cual beneficia el razonamiento, y, por ende, mejora el conocimiento procedimental (aumento del léxico y el razonamiento verbal), se incorporan nuevas estrategias intelectuales, lo que permite que pueda resolver los problemas y las tareas de mejor manera, mejor comprensión de los fenómenos naturales y culturales mediante el perfeccionamiento de los métodos de estudios y la obtención de la información, se automatizan los procedimientos ya consolidados permitiendo la aplicación de recursos cognitivos en la resolución de problemas, el cálculo aritmético mejora con la práctica, se aplica la integración de previos procedimientos en procedimientos más generales (aquello que ha dado buen resultado en un campo se tiende a aplicarlo en otro) (Brioso et al., 2009).

Ocurren cambios en el desarrollo del razonamiento moral (construcción de los principios éticos de validez universal), el individuo aprende a diferenciar y a conciliar sistemas efectivos distintos pero complementarios (producto de la interacción social y emocional) y la sabiduría, entendida como la habilidad para enfrentarse a los aspectos cotidianos de forma adaptativa y positiva, también sufre un proceso de desarrollo. Por último, los teóricos sostienen que el ciclo vital es un proceso complejo de ganancias y pérdidas, lo que supone un proceso de la

capacidad del psicorganismo para adaptarse a las demandas biológicas, psicológicas, culturales y medioambientales (Briso et al., 2009).

En la edad adulta y la vejez pueden aparecer trastornos afectivos que pueden manifestarse a través de síntomas atípicos como la desesperanza, fatiga o síntomas de ansiedad (Briso et al., 2009).

En Colombia, el Sistema Integral de Información de la Protección Social (SISPRO) reportó que las atenciones por depresión han aumentado desde el año 2009. En 2015, en los casos atendidos, la tendencia por edad está entre los 50 a 54 años, en los casos atendidos por diagnóstico de depresión moderada o severa se presentó en mayor proporción en mujeres (70,4%) que en hombres (29,6%), mientras que, en los casos de depresión grave con síntomas psicóticos son menos frecuentes y representan el 6% y el 13%, sin embargo, el mayor número de casos se presenta entre los 20 a 40 años y de 50 a 54 años (Ministerio de Salud y Protección Social, 2017).

En este sentido, entre las personas de 18 años y más, quienes tienen más riesgo de sufrir depresión son: personas con historia previa de episodio depresivo, con historia familiar de depresión, con problemas psicosociales (desempleo, separación conyugal, eventos vitales estresantes), abuso y dependencia de sustancias (incluidos alcohol y cigarrillo), personas que consultan a menudo al sistema de salud, quienes tengan enfermedades médicas crónicas, especialmente, cuando hay compromiso de la funcionalidad (enfermedad cardiovascular, diabetes, las que involucran dolor crónico y trastornos neurológicos); personas con otros trastornos psiquiátricos (trastornos de ansiedad, trastornos de personalidad); por cambios hormonales (embarazo, postparto y menopausia); con historia de intentos de suicidio y personas sedentarias (Ministerio de Salud y Protección social, 2013).

Las causas de la depresión son muy variadas, entre las que se encuentran (Guadarrama, et al., 2006):

- a) Factores genéticos: la influencia de la genética se observa más manifiesta en pacientes con trastornos depresivos graves y síntomas de depresión endógenas. El riesgo de morbilidad se presenta en parientes de primer grado (padres, hermanos o hijos) siendo independiente de la influencia del entorno o la educación. Los marcadores genéticos potenciales se han localizado en los cromosomas X, 4, 5, 11, 18 y 21.
- b) Factores químicos: se presenta una alteración de neurotransmisores, donde se ha evidenciado niveles anormales de la serotonina, norepinefrina y dopamina que podrían ser importantes en la fisiopatología de la depresión.
- c) Factores psicosociales: relacionado a experiencias adversas en la infancia como dificultades cotidianas y crónicas eventos indeseables en la vida, red social limitada y baja autoestima.

**Evaluación de las manifestaciones clínicas de la depresión en la edad adulta.** Para la evaluación de la depresión en el adulto, es de gran importancia la implementación de la entrevista clínica, la exploración física, exploración mental, síntomas actuales, episodios previos, respuesta ante el tratamiento, nivel de deterioro, afectación en las diferentes áreas, apoyo social, relaciones interpersonales y el riesgo de suicidio (García et al., 2015).

De acuerdo con los manuales de criterios diagnósticos (CIE y el DSM) los síntomas deben tener una duración al menos de dos semanas, su gravedad se basa en el número, tipo e intensidad y el grado de deterioro funcional (García et al., 2015), sin embargo, estos criterios deben ser utilizados para orientar la entrevista y el juicio clínico del profesional, por lo que, la



evaluación debe basarse en un enfoque amplio y no basarse exclusivamente en el recuento de síntomas (Ministerio de Sanidad, servicios sociales e igualdad, 2014).

*Evaluación diferencial de las manifestaciones clínicas de la depresión.* El estado de ánimo deprimido por sí solo no es un determinante diagnóstico de ningún cuadro clínico, la tristeza es un síntoma que está presente en diferentes cuadros, por lo tanto, es vital evaluarse de modo *diferencial* (Sanz y Vásquez, 2009). Es importante constatar si la etiología está relacionada a otra patología médica, trastorno mental, consumo de drogas o fármacos, por lo que resulta fundamental realizar una historia clínica, exploración física y exámenes de laboratorio (García et al., 2015).

La primera medida diagnóstica a considerar es si la persona con síntomas de depresión tiene o no un trastorno depresivo, entre los síndromes clínicos que pueden confundirse con un diagnóstico de depresión mayor se encuentran el trastorno bipolar, los trastornos de ansiedad (fobias, pánicos, agorafobias, obsesiones), las reacciones normales de duelo, los trastornos de adaptación con estado de ánimo deprimido, los trastornos del sueño (insomnio, narcolepsia, apneas, nocturnas, etc.), los trastornos esquizofrénicos (en especial esquizoafectivos de tipo depresivo), las quejas somáticas vagas (hipocondría), la demencia y pseudodemencia y los trastornos sexuales (Sanz y Vásquez, 2009).

### **Evaluación de las variables psicológicas**

La evaluación psicológica es la disciplina de la psicología científica encargada de la exploración y análisis del comportamiento humano, a niveles de complejidad que se definan oportunos (motor, fisiológicos y/o cognitivo) y que varían de acuerdo a los objetivos aplicados y de investigación como son la detección, diagnóstico, descripción, clasificación, selección,

explicación, predicción, tratamiento (o intervención) y valoración (Aragón, 2015; González, 2007). Para lograr una adecuada comprensión de su estudio, es importante entender los aspectos cualitativos y cuantitativos que la componen, sin embargo, la medición psicológica no se rige estrictamente a la metodología generalmente implementada en las ciencias duras (Muñiz, 1998), sino a través de metodologías básicas como la observación, la entrevista, los test proyectivos, los autoinformes y los test psicométricos (González, 2007). Estos últimos, tienen su fundamento en la psicometría que abarca teorías, métodos y usos de la medición, que se distingue del uso formal y estructurado de las matemáticas (Aliaga, 2007).

Para llevar a cabo una rigurosa evaluación psicológica, es importante que los profesionales que la realicen tengan una adecuada y pertinente preparación, los instrumentos deben mostrar buenas propiedades psicométricas y su uso debe ser de forma correcta (Muñiz, Hernández y Ponsoda, 2015).

**Tipos de evaluación.** Estar “deprimido” puede entenderse desde tres niveles, y cada uno de ellos implica una aproximación diferente sobre su evaluación. En el primer nivel, la sintomatología, refiere a un estado de ánimo triste y decaído; el segundo nivel sindrómico, involucra un conjunto de síntomas concurrentes con la depresión (cambios en el apetito, en el sueño, pérdida de placer, etc.) con un grado de gravedad y duración que podría estar relacionado a otros trastornos mentales; y por último, el nivel nosológico, alude a que los síntomas de un episodio depresivo no se deba a otro trastorno o condición, por lo que, se habla de un trastorno depresivo (Vázquez, Hernangómez, Hervás y Nieto, 2005).

De esta manera, al evaluar cuantitativamente la tristeza (gravedad y frecuencia) junto a otros síntomas el uso de cuestionarios correspondería al modo más adecuado lo que efectúa una aproximación dimensional, mientras que, para generar un diagnóstico lo más pertinente es la

comprobación de una serie de criterios diagnósticos a través de una entrevista clínica adecuando a los manuales habitualmente utilizados lo que alude a una aproximación categórica (Vásquez et al., 2005). En otras palabras, la evaluación de lo psicológico puede ser categorial o dimensional (Widakowich, 2012; Hernández, del Palacio, Freyre, 2011; Widiger & Samuel, 2005; Lara et al., 2002).

***Evaluación categorial.*** Se trata del establecimiento de categorías de propiedades definidas, para establecer la presencia o ausencia de una categoría (nivel sindrómico), que permite establecer un diagnóstico mediante la entrevista clínica estructurada (Widakowich, 2012; Lara et a, 2002).

De acuerdo con Hernández et al. (2011), en todo sistema categórico, las categorías deben cumplir tres criterios fundamentales: ser discretas, ser mutuamente exclusivas y exhaustivas: *Ser discretas* hace referencia a que los elementos de las diferentes categorías no deben compartir las características relevantes de la clasificación, es decir, se espera que las personas diagnosticadas con diferentes trastornos no compartan los mismos síntomas; *ser mutuamente exclusivas* implica que un elemento cabe en una categoría más no en otra, en otras palabras, si una persona padece de un trastorno, no puede padecer otro al mismo tiempo; por último, *ser exhaustiva* es que todos los elementos que se encuentran en una categoría comparten las mismas características, es decir, todas las personas diagnosticadas con un trastorno específico comparten los mismos síntomas.

***Evaluación dimensional.*** Requiere de la identificación y medición de las diferencias cuantitativas individuales de un fenómeno psicológico, ordenando el número de síntomas, intensidad y variedad (Widakowich, 2012; Hernández et al., 2011). La medición de las variables psicológicas como la depresión, ansiedad, calidad de vida, neuroticismo, etc., suelen llevarse a cabo mediante el uso de indicadores observables, por ejemplo, las respuestas de un sujeto frente

a los interrogantes de un cuestionario y la bondad de ésta dependerá de la calidad científica del test (Carretero-Dios y Pérez, 2005).

La psicometría adquiere un mayor significado en la medida que desempeña su papel en la evaluación de las variables psicológicas mediante teorías, métodos y técnicas vinculados al desarrollo y aplicación de test con el objetivo de describir, clasificar, diagnosticar, explicar o predecir fenómenos psicológicos que permitan la orientación en la toma de decisiones y/o acciones en el ejercicio profesional (Meneses, Barrios, Bonillo, Cosculluela, Lozano, Turbany y Valero, 2013). El test psicométrico es un procedimiento de medida objetiva y estandarizada de una muestra de comportamientos (Anastasi, 1988), miden habilidades cognitivas y rasgos de personalidad en las personas, son estructurados, y permiten la obtención de una información más amplia, que orienta a la creación de hipótesis para la evaluación y el diagnóstico (González, 2007).

Carretero-Dios y Pérez (2005) destacan la importancia de la construcción, adaptación y uso de los test en la medición de lo psicológico y establecen la relevancia de reconsiderar periódicamente el proceso y secuenciación que conllevan a la realización de los mismos para mejorar la evaluación psicológica, para ello, proponen las Normas para el Desarrollo y Revisión de Estudios Instrumentales, guía que debe considerarse en todo proceso de construcción/adaptación de un test, cuya finalidad es la de servir como referencia para la elaboración y revisión de investigaciones instrumentales.

### **Calidad de los instrumentos de medida: contribuciones de la psicometría**

La psicometría es un conjunto de métodos, técnicas y teorías formales que cimientan las bases para ejecutar adecuadamente la medición de las variables psicológicas, es decir, implica

todo aquello relacionado con la medición de lo psicológico (Muñiz, 2018). Permite el establecimiento de métodos para transformar los hechos en datos cuantificables e interpretarlos en función a la teoría, sin embargo, toda medida psicológica contiene un término de error, y que por lo tanto, los instrumentos psicométricos deben medir con precisión (Aragón, 2015).

Las variables psicológicas, por su naturaleza no fáctica, no pueden ser observadas directamente, siendo esta una dificultad que enfrenta la psicometría, por lo que resulta conveniente recurrir a indicadores indirectos: las conductas observables (Martínez, Hernández y Hernández, 2006). Por su parte, las conductas no observables directamente, hacen referencias a los *constructos teóricos* relacionado con las teorías de las conductas observables y con otros constructos, dicha relación entre los constructos latentes y los hechos observables permiten la asignación de números y el establecimiento de relaciones empíricas, que para ser medidas requieren de instrumentos estandarizados (test) que evidencie su variabilidad y algún modelo que relacione lo que se observa con la realidad (Martínez et al., 2006)

La idea central de la implementación de los test es la existencia de un atributo único: el constructo (que pueden ser muy variados), siendo el objetivo de medición del test, y que, mediante el uso de las puntuaciones numéricas, es posible situar a las personas y los ítems, lo que supone el reflejo de la cantidad de atributo que el encuestado posee, pero no se puede asumir que refleje exactamente la verdadera cantidad, en este sentido, los errores aleatorios desempeñan un papel importante en la asignación numérica y el proceso de medición se convierte en un proceso inferencial (Martínez et al., 2006).

En primer lugar, atendiendo a los postulados de Martínez et al (2006), se construye una serie de tareas (los ítems) que se realizan a partir de las especificaciones del constructo (lo más detallados posibles), la presentación y respuestas de los ítems pueden tomar diferentes

modalidades, las respuestas son la única parte observable del proceso que reciben una puntuación numérica denominada puntuación observada a través del proceso conocido como *escalamiento*.

**Los principios del modelo psicométrico.** Para determinar si un instrumento (elaborado o adaptado) es apto para la aplicación, la comunidad científica ha desarrollado algunos indicadores de gran utilidad: la validez y la confiabilidad (Cardoso, Gómez-Conesa y Hidalgo, 2010; Muñiz, 1998).

**Validez.** La validez se considera como un requisito previo y el aspecto fundamental para el uso y administración de cualquier instrumento psicométrico (Alvarado y Santisteban, 2011). Consiste en el grado en que un instrumento mide realmente la variable que pretende medir (Hernández, Fernández y Baptista, 2015). Supone examinar las teorías sobre las que se basan los datos y comprobar su fuerza y credibilidad a través de las diversas fuentes de evidencia (Martínez et al., 2006), es decir, lo que se valida no es el test en sí, sino las inferencias realizadas a partir de él, cuya forma estándar de validarlo es a partir de la derivación de predicciones y contrastarlas con los datos (Muñiz, 2018).

Los primeros trabajos sobre validez distinguían variedades de tipos de validez, tradicionalmente los especialistas las han conceptualizado en tres categorías: validez de contenido, validez relacionada con el criterio (incluye validez concurrente y predictiva) y validez de constructo (incluye validez convergente y discriminante) (Cohen y Swerdlik, 2006; Argibay, 2006).

**Validez de contenido.** Refiere al grado en que un instrumento recoge una muestra representativa de los contenidos correspondientes al campo evaluado, normalmente definido por la literatura (teoría o trabajos antecedentes) (Hernández et al., 2015; Muñiz, 1998). Para

establecer un posible banco de reactivos es importante tener una adecuada conceptualización y operacionalización del constructo, es decir, se debe especificar las dimensiones y sus indicadores (Escobar-Pérez y Cuervo-Martínez, 2008). En términos más específicos, corresponde a la relación que existe entre los ítems y lo que se pretende medir, otorgando atención a la relevancia y a la representatividad de estos (Meneses et al., 2013).

La validez de contenido generalmente se evalúa a través de un panel o un juicio de grupo de expertos, que implica una opinión informada de personas expertas y calificadas en el tema que pueden dar información, evidencia, juicios y valoraciones (Escobar-Pérez y Cuervo-Martínez, 2008), sin embargo, para determinar la confiabilidad de un juicio de expertos, se requiere conocer el grado de acuerdo entre ellos, dado que este implica elementos subjetivos (Aiken, 2003). Diferentes autores propusieron procedimientos o métodos para calcular el grado de acuerdo entre los jueces, entre los que se encuentran: Método basado en el Análisis Factorial (Tucker, 1961), índice de Validez de Contenido (IVC) conocido como Panel de Evaluación de Contenido (Lawshe, 1975), índice de congruencia ítem-objetivo (Rovinelli y Hambleton, 1977), índice de congruencia (Hambleton, 1980, 1984), V de Aiken (Aiken, 1980), Escalamiento multidimensional y análisis de clusters (Sireci y Geisienger, 1992), Método de Capacidades Mínimas (Levine, Maye, Ulm y Gordon, 1997), Rango Interpercentil Ajustado a la Simetría (Fitch, Bernstein, Aguilar, Burnand, LaCalle, Lázaro, Loo, McDonnell, Vader y Kahan, 2001), Coeficiente de Validez de Contenido (Hernández-Nieto, 2002), índice de Validez Factorial (Rubio, Berg-Weber, Tebb, Lee y Rauch, 2003) y el índice Promediado de la Desviación Media (Claeys, Neve, Tulkens y Spinewine, 2012).

La validez de contenido se establece en diferentes situaciones, siendo dos de las comunes:

(a) El diseño de una prueba y (b) la validación de un instrumento que fue construido para una

población específica, pero que fue adaptada mediante el proceso de traducción (equivalencia semántica), teniendo en cuenta que, la sola adaptación no necesariamente produce una equivalencia cultural en aspectos como el idioma, significados culturales, interpretaciones de conductas basado en normas culturales, sino que se hace necesario validarlo en términos de su contenido, y es dónde la evaluación realizada por expertos toma especial importancia, pues son quienes tienen la función de eliminar los ítems irrelevantes y modificar aquellos que así lo requieran (Escobar-Pérez y Cuervo-Martínez, 2008).

En este sentido, Carretero-Dios y Pérez (2005) resaltan que al momento de emprender el proceso de creación/adaptación de un instrumento, se debe definir cuidadosamente el constructo en términos de la concreción inicial de los componentes operativos de éste, denominada como *definición semántica de la variable* (Lord y Novick, 1968), para ello, existen diferentes propuestas de cómo presentarla, que comúnmente proponen el uso de una tabla dónde se inserte toda la información de interés (Osterlind, 1989), cuyo objetivo principal es la evaluación por parte de un grupo de expertos en la temática, la cual, una vez finalizada, debe realizarse una propuesta operativa definitiva (también presentada en una tabla) que debe contener toda la información respecto a la definición semántica del constructo, pasando a ser la tabla de especificaciones del instrumento, de ahí radica la importancia de ser reportada en las investigaciones de carácter instrumental (Carretero-Dios y Pérez, 2005). Es este sentido, al no presentar los elementos diferenciadores, las manifestaciones operativas o los componentes, va a influenciar un proceso de construcción/adaptación impreciso con tendencia a suministrar evidencia deficiente de la validez de contenido (Nunnally y Bernstein, 1995).

Por otro lado y muy en relación con la validez de contenido se encuentra la *validez aparente*, que hace referencia al grado en que un instrumento da la impresión a los evaluados de



que mide lo que pretende medir, es importante que las personas perciban que los ítems están relacionados con la finalidad del test (Abad, Garrido, Olea y Ponsoda, 2006).

*Validez relacionada con el criterio.* Este tipo de validez alude al grado en que un instrumento se emplea para la estimación de una conducta a futuro, en una situación concurrente con la administración de la prueba (Aragón, 2015), en otras palabras, la validez de criterio se basa fundamentalmente en relacionar las puntuaciones del test con otras variables (criterio), y se implementa para hacer inferencias de los puntajes obtenidas en el test respecto a otra variable de interés (Argibay, 2006).

El interés por la predicción de un criterio suele presentarse en instrumentos de medición de variables clínicas (por ejemplo, depresión, ansiedad, etc.), donde se busca que los componentes logren diferenciarse y que se vean relacionados con indicadores considerados como una señal distintiva de la variable que se está midiendo (Carretero-Dios y Pérez, 2005).

Existen dos tipos de evidencias de validez relacionada con el criterio: la validez concurrente y la validez predictiva (Cohen y Swerdlik, 2006). La diferencia entre ambos tipos de validez radica en la temporalidad del criterio, es decir, si las puntuaciones del instrumento se emplean para predecir alguna medida del criterio que se realizará a futuro, se estaría hablando de validez predictiva, pero si por el contrario, se relaciona las puntuaciones del test con alguna medida del criterio tomada en el mismo momento, entonces sería validez concurrente, sin embargo, resulta pertinente aclarar que dichas diferencias no están relacionadas con que si son predictivas o no (desde un punto de vista científico), sino más bien con el diseño que involucran con respecto a su dimensión temporal (Argibay, 2006).

*Validez de constructo.* Un constructo es un concepto referido por los teóricos de la psicología para explicar el comportamiento humano que tiene lugar dentro de una hipótesis, una teoría o modelo teórico y que precisan de indicadores observables para su correspondiente estudio, que en muchas ocasiones, se refieren a los ítems de un test (Hernández et al., 2015; Abad et al., 2006). La validez de constructo refiere a que tan bien una prueba representa medir un constructo o rasgo teórico (Aragón 2015), cuyos métodos de estudio y técnicas estadísticas para el análisis de datos son muy variados, que se sintetizan de la siguiente manera (Abad et al., 2006):

- a) Formular las hipótesis relevantes: Son hipótesis extraídas de deducciones teóricas en las que aparezca el constructo que se pretende evaluar con el test.
- b) Efectuar en la práctica mediciones de las variables inmersas en las hipótesis: refiere a que la medición del constructo o variable se realizará con la prueba diseñada a tal efecto, que es la que se pretende medir.
- c) Determinar si se verifican o no las hipótesis planteadas: En el caso que se verifique las hipótesis planteadas, queda confirmado que el test mide el constructo o variable de interés, sin embargo, en caso que las hipótesis no se confirman, no significa inicialmente que el test no es válido, sino que puede ser debido a que las hipótesis no estaban bien formuladas, por lo que requiere de una revisión de la teoría subyacente.

Entre los métodos más frecuentes para contrastar las hipótesis son (Abad et al., 2006):

- d) Obtener las relaciones entre las puntuaciones del instrumento y en otras variables que deberían relacionarse con el constructo, si el modelo teórico está bien fundamentado, debe establecer relaciones entre la variable de interés y otras diferentes, por lo que es posible establecer diseño de investigaciones para contrastar las previsiones teóricas.

- e) Evaluar a través del instrumento a grupos que se supone que deben ser distintos a la variable, para de esta manera comprobar que realmente es así, es decir, si el test es válido, debería reflejar las diferencias entre los grupos que se predicen desde la teoría.
- f) Utilizar una estrategia de tipo experimental para comprobar si el test resulta sensible para detectar los efectos previsibles a la manipulación de los niveles en una o más variables independientes.
- g) Aplicar la técnica multivariada del Análisis Factorial (exploratorio o confirmatorio) sobre la matriz de correlaciones entre los ítems, para descubrir estadísticamente las variables o dimensiones subyacentes, lo que se conoce como factores a la covariación entre los ítems.

Esta última estrategia de análisis factorial requiere de alguna precisión que resulta pertinente dado que su fundamento es basado en una técnica estadística sofisticada y porque su utilización práctica es muy extensa.

De acuerdo con Abad et al., (2006) el análisis factorial es una técnica estadística multivariante que estudia las dimensiones que subyacen a las relaciones entre varias variables, toma normalmente como punto de partida la matriz de correlaciones entre las  $n$  variables de interés para analizar, proporciona una matriz de tamaño  $n \times p$  llamada matriz factorial rotada y se lleva a cabo con dos objetivos: 1. Determinar el número de dimensiones o factores que mide un test y el significado de cada una, y 2. Obtener la puntuación de cada evaluado en cada dimensión. Normalmente el número de factores es inferior al total de ítems y para descubrir su significado es necesario fijarse en las variables que saturan elevadamente cada dimensión y se debe realizar un proceso inferencial para encontrar la relación entre las variables que reflejan correlaciones elevadas en la dimensión; adicionalmente, cada factor no tiene la misma importancia, dado que cada uno explica una determinada cantidad de la varianza total de los ítems (expresado

porcentualmente), esto quiere decir que, si un factor explica un porcentaje elevado de la varianza total, eso es señal de que dichas saturaciones de las variables en el factor son altas, por lo tanto es una dimensión importante al momento de describir las relaciones entre las variables originales (Abad et al., 2006).

Existen dos estrategias distintas de análisis factorial: exploratorio y confirmatorio. En un proceso de Análisis Factorial Exploratorio (AFE), el investigador no tiene una idea exacta de cuántos factores subyacen a las relaciones entre las variables ni qué variables tienen un peso alto en cada factor, mientras que, en el caso del Análisis Factorial Confirmatorio (AFC), el investigador plantea hipótesis definidas a priori sobre cuál es el número de factores y cómo pesan las variables en ellos; a medida que se acumulan estudios dentro de un campo de conocimiento, los investigadores optan por utilizar técnicas confirmatorias frente a exploratorias, sin embargo, la teoría que subyace a ambos tipos de análisis es la misma (Abad et al., 2006).

Por otro lado, muy cercano a la validez factorial se encuentran la *validez convergente* y *validez discriminante*, la cual se evalúa a través de las matrices multirrasgo-multimétodo (Abad et al., 2006). La validez convergente refiere a que si dos instrumentos miden un mismo rasgo, se esperaría que la correlación entre ellas sea alta, mientras que, en la validez discriminante, si dos instrumentos miden rasgos diferentes, las correlaciones entre ellos debe ser baja o al menos menor en comparación con otros test que midan el mismo rasgo (Abad et al., 2006), es decir, si dos variables son conceptualmente diferentes y sus correspondientes escalas están justificadas suficientemente bien a nivel teórico, los análisis basados en covarianzas o correlaciones no deben dar resultados inequívocos (Martínez-García y Martínez-Caro, 2008).

**Fiabilidad.** Un instrumento de medida se considera fiable si las medidas inferidas a partir de él carecen de errores de medida, es decir, son consistentes; un test es fiable en la medida que

cada vez que se aplica a las mismas personas da el mismo resultado (Muñiz, 2018). En otras palabras, la fiabilidad es la consistencia o estabilidad de las medidas, en el momento en que el proceso de medición se repite, respondiendo hasta qué punto refleja con precisión la puntuación verdadera del individuo (Martínez et al., 2006; Prieto y Delgado, 2010), comprende un conjunto de métodos y técnicas para la estimación del grado de precisión con el que se está midiendo una variable (Muñiz, 1998). La falta de confiabilidad de un test radica en la relación con la intervención del error, considerado como algún efecto de carácter irrelevante para los resultados de la medición, pero que influye sobre la falta de confiabilidad (Aliaga, 2007).

Existen dos tipos de errores: error constante (sistemático) y error causal (no sistemático), el primero se adquiere cuando las mediciones obtenidas con una escala son sistemáticamente mayores o menores que lo que realmente deben ser, mientras que, el error causal se produce cuando las mediciones son alternativamente mayores o menores de lo que deben ser realmente e interviene cuando se afecta la confiabilidad de un test (Brown, 1980).

La confiabilidad se presenta mediante el coeficiente de confiabilidad ( $p_{xx}$ ) y del error estándar de medida (EEM). El primero, es un coeficiente de correlación entre dos grupos de puntajes e indica el grado en que los sujetos mantiene sus posiciones dentro de un grupo, el cual, abarca los valores desde 0 a 1, donde el coeficiente cero indica nula confiabilidad, mientras que uno significa un máximo de confiabilidad (Hernández et al., 2015; Aliaga, 2007), el coeficiente indica la cuantía en que las medidas están libres de errores causales.

De acuerdo con Muñiz (2018), para calcular el valor empírico de una muestra determinada hay que valerse de su definición: correlación entre las puntuaciones en dos formas paralelas, en suma de: 1. Elaborar las dos formas paralelas, 2. Aplicarlas a una muestra amplia de

personas representativas de la muestra, y 3. Calcular las correlaciones entre las puntuaciones de la muestra en ambas formas; cuyo método se denomina *método de las formas paralelas*.

Sin embargo, se suelen utilizar otros dos métodos (Muñiz, 2018):

- h) Test-retest. Este método consiste en aplicar el mismo test en dos ocasiones a las mismas personas, en el que la correlación entre las puntuaciones de las dos aplicaciones será el coeficiente de confiabilidad, la estimación obtenida se denomina “coeficiente de estabilidad” en la medida que indica en qué grado las mediciones son estables en la primera aplicación del instrumento.
- i) Dos mitades. El test se aplica una sola vez, obteniendo las puntuaciones de cada persona a cada una de las mitades en las que se divide el test, en este caso, el coeficiente de confiabilidad viene dado por la corrección entre las dos mitades más una corrección para obtener la fiabilidad del test total, la estimación obtenida indica la covariación o consistencia interna de las dos mitades (es un indicador de la consistencia interna)

Una vez estimada la fiabilidad del test por algunos de los métodos, es posible hacer ciertas estimaciones acerca de las puntuaciones verdaderas de las personas en el test, es decir, hacer estimaciones sobre la cantidad de error que afecta a las puntuaciones empíricas, tales como: estimación mediante la desigualdad Chebychev, estimación basada en la distribución normal de los errores y estimación según el modelo de regresión (Muñiz, 2018).

Por su parte, otra forma de acercarse a la fiabilidad de un test es mediante el coeficiente *alfa* ( $\alpha$ ) propuesto por Cronbach (1951), el cual refleja el grado en el que covarían los ítems que constituyen el test, por lo tanto, es un indicador de la consistencia interna del test.

Por otro lado, el error estándar de medida permite estimar la diferencia entre la puntuación empírica de un sujeto ( $x$ ) y su puntuación verdadera ( $y$ ) (Meneses et al., 2013), por lo que, el objetivo de la fiabilidad es determinar en qué medida el error afecta la puntuación empírica (Muñiz, 1998).

De acuerdo con Muñiz (1998), la estimación de los errores de medida es inherente en todas las ciencias, dado que, en toda medición, ya sea de mayor o menor grado, conlleva un cierto margen de error; en el caso de las variables psicológicas para la estimación de los errores, la propuesta pionera fue realizada por Spearman (1904, 1907, 1913) denominada *Modelo Lineal Clásico*, dando un origen a un enfoque general sobre los test llamada *Teoría Clásica de los Test*, posteriormente, alrededor de los años 60-70 se originan nuevos modelos para abordar la estimación de los errores, agrupándose entre ellos lo más implementados, bajo el nombre genérico de *Teoría de Respuesta a los Ítems*.

**Teorías de los Test.** Las teorías estadísticas de los test, mediante la estimación de las propiedades psicométricas de los instrumentos, permiten garantizar que las decisiones tomadas a partir de estos sean las más adecuadas y pertinentes, así mismo, sin dichas teorías no se podría estimar la fiabilidad y la validez de los test, lo cual es de suma importancia para el uso riguroso y científico de los instrumentos (Muñiz, 2010).

*Teoría Clásica de los Test (TCT).* El enfoque clásico de los test es considerado el predominante en la construcción y análisis de los test, propuesto originalmente por Spearman, mediante el modelo lineal y el conjunto de supuestos es posible la construcción de pruebas y el análisis de sus elementos, así como, determinar su confiabilidad y validez; este modelo se expresa de la siguiente manera (Aragón, 2015; Muñiz, 2010):  $X = V + e$  ( $X$  es la puntuación empírica de un sujeto,  $V$  es la puntuación verdadera y “ $e$ ” es el término de error).

Donde la puntuación obtenida o empírica en una prueba ( $X$ ) es igual a la puntuación verdadera ( $V$ ), más el error de medición ( $e$ ); entre más confiable sea la prueba y menor el error, la puntuación empírica se acercará más a la puntuación verdadera.

Este modelo presenta tres supuestos y varias deducciones derivadas de él propuesto por Spearman (Muñiz, 2010):

- j) Primer supuesto: consiste en definir la puntuación verdadera ( $V$ ) como la esperanza matemática de la puntuación empírica:  $V = E(X)$ . Conceptualmente, se define la puntuación verdadera en un test como aquella puntuación que obtendría como media si se le aplicara infinitas veces el mismo instrumento.
- Segundo supuesto: no existe relación entre las puntuaciones verdaderas en una prueba y sus respectivos errores de medida, dando a entender que, los errores son aleatorios y no dependen del valor obtenido de la puntuación verdadera:

$$r(v,e) = 0.$$

- Tercer supuesto: los errores de medida de los individuos en una prueba no se relacionan con los errores de medida en otro instrumento, se representa de la siguiente manera:  $r(e_j, e_k) = 0$ .

El modelo lineal, los tres supuestos y los Test Paralelos constituyen el eje central de la Teoría Clásica de los Test, entendiendo por test paralelos a aquellos test que miden exactamente lo mismo, pero con distintos ítems, las puntuaciones verdaderas serían las mismas, así como, las varianzas de los errores de medida (Muñiz, 2010).

***Teoría de Respuesta a los Ítems (TRI)***. La Teoría de Respuesta al Ítem (TRI), también denominada Teoría de Rasgo Latente, es un modelo probabilístico que permite conocer la



información proporcionada por cada ítem, y de esta manera, crear test individualizados (Aliaga, 2007). Este modelo abarca unas series de investigaciones psicométricas independientes iniciadas por Rash (1960) y Birnbaum (1968) cuyo común denominador es el reconocimiento de una relación directa entre la conducta del individuo frente a un ítem y es rasgo de dicha conducta para calcular esta relación utiliza modelos matemáticos que estiman la probabilidad para dar determinada respuesta a un ítem particular para cada nivel de rasgo medio por este (Attorresi, Aguerri, Abal, Galibert y Luzzia, 2009).

Las primeras consideraciones a esta teoría se remontan alrededor de los años 20 en los aportes de Thurstone (1925); continúan en los años cuarenta con Lawley (1943, 1944) o Tucker (1946); pero fue el psicómetra Lord (1952) quien cimenta las bases de la TRI; más adelante, Birnbaum aporta nuevos avances; Rash (1960) propuso el modelo logístico de un parámetro, que posteriormente; Lord y Novick (1968) dedicaron cinco capítulos al tema de la TRI, que permitió abrir el desarrollo aplicado de la teoría e investigación.

### **Practica psicométrica basada en la evidencia**

En sus inicios, el proceso de revisión de la literatura científica sobre un tópico en las ciencias de la salud y las ciencias sociales se cimentaban en la subjetividad del revisor, sin la implementación de normas que dirigieran la objetividad de la investigación, lo que llevaba a una acumulación pobre de conocimiento científico, generando así, bases no sólidas sobre las evidencias científicas extraídas (Rosenthal, 1991).

En la década de 1970, estas dificultades basadas en la subjetividad se incrementaron en la medida que se produjo un alto volumen de estudios empíricos en diferentes ámbitos, lo que requirió el uso de estrategias objetivas y sistemáticas para llevar a cabo el proceso de revisión

(Sánchez-Meca, 2010), originando de esta manera, el *meta-análisis* (MA) y las *revisiones sistemáticas* (RSs), como una metodología objetiva y rigurosa para realizar el proceso de revisión de la información en un tópico concreto y alcanzar de esta manera una eficiente acumulación de las evidencias (Hunt, 1997).

El término “metaanálisis” fue introducido en 1976 por Gene Glass sobre la eficacia de la psicoterapia, entendido como un análisis estadístico de una larga colección de resultados de análisis provenientes de estudios individuales con el objetivo de integrar los resultados, y sobre el tamaño del efecto (Smith y Glass, 1983).

Posteriormente en los años 80, surge el movimiento de la “Práctica Basada en la Evidencia”, y con él el concepto de “mejores prácticas” (*best practices*), que alude al servicio e intervención a un paciente atendido por un profesional de la salud basado en la información actualizada, teniendo en cuenta la concepción del caso, el marco cultural y valorativo del individuo. Esta idea impactó en áreas como la medicina, enfermería, trabajo social, psiquiatría y psicología (Martínez-Taboas, 2014; Martínez, Adefna y Izquierdo, 2009). La APA (2005) define la práctica psicológica basada en la evidencia como la integración de la mejor investigación en conjunto con la pericia clínica, en el marco de las características del paciente, la cultura y sus preferencias, cuyo propósito es promover una práctica psicológica efectiva y potenciar la salud pública a través de principios de evaluación, formulación de casos, relación terapéutica e intervención fundamentados empíricamente.

La Psicología Basada en la Evidencia (PBE) es un requerimiento científico que pretende modificar el modo de trabajo profesional, de tal manera que, se tome a consideración las mejores evidencias científicas en la toma de decisiones diarias en un determinado problema, por ejemplo, selección de técnicas en un tratamiento, elección de algún programa de intervención o el mejor

método para realizar un diagnóstico, lo cual, requiere de iniciar un proceso de búsqueda de la información y su respectivo análisis crítico, sin embargo, dada la constante acumulación de información de estudios empíricos, estos pueden bloquear el abordaje de la PBE al hacer casi imposible la selección de estudios relevantes y su respectiva lectura crítica, teniendo en cuenta, el corto período con el que cuenta el profesional y su elevada carga laboral; en este sentido, es donde entra en función las RSs y los MA (Sánchez-Meca y Botella, 2010).

Una revisión sistemática es un tipo de investigación que tiene como objetivo reunir toda la evidencia científica referente a un tema en particular, parte de una pregunta formulada, clara y objetiva, llevando a cabo métodos sistémicos y explícitos, a través de criterios de elegibilidad previamente establecidos, que permitan localizar, seleccionar y valorar críticamente las investigaciones primarias relevantes entorno a la pregunta planteada y aplicar protocolos sistemáticos para la recolección de datos e información, de tal manera que, se logre alcanzar conclusiones válidas y objetivas sobre las evidencias del tema de estudio, y a su vez, implementar estrategias para reducir sesgos y errores aleatorios (Fuensanta, 2015; Torres-Fonseca y López-Hernández, 2014; Higgins y Green, 2011; Sánchez-Meca, 2010; Urra y Barría, 2010; Moher, Liberati, Tetzlaff y Altman, 2009).

Por su parte, el metaanálisis es una técnica estadística que tiene como función combinar los resultados de un conjunto de estudios individuales para sintetizar los resultados y dar una estimación global (Fuensanta, 2015; Higgins y Green, 2011; Moher et al., 2009). En este sentido, el metaanálisis no es sinónimo de revisión sistemática, es una parte del proceso de revisión. De esta manera, el uso de las RSs y MA le permite al profesional ahorrar tiempo y a su vez, ofrece información de las mejores evidencias científicas (Sánchez-Meca y Botella, 2010).

Para el proceso de implementación y elaboración de RSs, en el plano internacional, han surgido unas series de organizaciones que han desarrollado manuales y lineamientos ampliamente aceptados y adoptados por la comunidad científica, que funcionan como guías para la preparación y publicación de una RS en el marco de altos estándares de calidad, en el que han sido partícipes grupos interdisciplinarios de expertos (metodólogos, clínicos, editores, etc.).

Entre los más utilizados se encuentran:

- La Colaboración Cochrane se ha constituido como la principal organización que concreta las RSs, cuyo propósito es ayudar a la toma de decisiones sanitarias bien fundamentadas a través de su preparación, mantenimiento y promoción del acceso a las RSs basadas en la evidencia. El *Manual Cochrane de Revisiones Sistemáticas de Intervenciones* sirve como guía para la preparación sistemática de las revisiones Cochrane sobre intervenciones, la cual constituye una serie de pasos para la respectiva elaboración (Higgins y Green, 2011). Así mismo, la Colaboración Cochrane desarrolló una guía para incluir las revisiones en pruebas diagnósticas bajo el *Manual Cochrane de revisiones sistemáticas de validez de pruebas diagnósticas* (Deeks, Bossuyt & Gatsonics, 2013).
- La declaración *Quality of Reporting of Meta-analyses* -QUOROM (Calidad de los Informes de Meta-análisis) (Moher, Cook, Eastwood, Olki, Rennie & Stroup, 1999), luego su posterior revisión y ampliación en la declaración *Preferred Reporting Items for Systematic Reviews and Meta-analyses* -PRISMA (Elementos de Informes Preferidos para Revisiones Sistemáticas y Meta-análisis) es una lista de verificación de 27 ítems y un diagrama de flujo de 4 fases, cuyo objetivo es brindar una guía a los autores para mejorar la presentación de los informes de revisiones sistemáticas y metaanálisis, se centra en los ensayos aleatorizados, pero puede ser implementado en otros tipos de investigaciones, en

particular, evaluaciones de intervención y no es un instrumento de evaluación de la calidad de una revisión sistemática (Moher et al., 2009).

- La Colaboración Campbell es una red internacional que apoya la preparación y extensión de las RSs de alta calidad sobre evidencia de investigaciones en torno a la efectividad de programas, políticas y prácticas sociales, cuya misión consiste en promover un cambio social positivo para contribuir a mejores decisiones y mayor efectividad en servicios públicos y privados alrededor del mundo, cuenta con una serie de políticas que informan los requisitos, pautas de ejecución y fuentes seleccionadas de información para las RS (The Campbell Collaboration, 2014).

Desde sus inicios, las RSs se han centrado en estudios experimentales de preferencias los Ensayos Clínicos Aleatorizados (ECA) enfocándose en la efectividad de sus intervenciones (Urra y Barría, 2010; Sánchez-Meca y Botella, 2010). Sin embargo, las RSs en instrumentos de medición son cada vez más reconocidos como unas herramientas útiles para la selección de instrumentos en la investigación y la práctica clínica (Trujols et al., 2015; Terwee, Mokkink, Knol, Ostelo, Bouter, de Vet, 2012) y para evaluar su calidad (Mokkink, Terwee, Stratford, Alonso, Patrick, Riphagen, Knol, Bouter & de Vet, 2009).

Aunque existen otras listas o normas de verificación, estas no constan como herramientas de calificación de la calidad para RSs, por lo tanto, la lista de verificación *Consensus-Based Standards for the selection of Health Measurement Instruments- COSMIN* (Estándares basados en el censo para la selección de instrumentos de medición de la Salud), es la única herramienta actual que evalúa la calidad metodológica de los estudios de propiedades de medición en el contexto de revisiones sistemática y de manera estandarizada, su implementación permitirá

aumentar la solidez de las RSs y los MA de las propiedades métricas de un instrumento de medición (Trujols et al., 2015).

El enfoque inicial de la metodología COSMIN se basa en la evaluación de *Health Related Patient- Reported Outcomes*- HR PROs (Resultados Informados por el Paciente relacionados con la Salud) (Mokkink, Terwee, Knol, Stratford, Alonso, Patrick, Bouter, de Vet, 2006), pero puede implementarse en la evaluación de la calidad metodológica de estudios en otros instrumentos de medición dado que, las mismas propiedades de medición pueden ser relevantes para otro tipo de instrumentos relacionados con la salud, por ejemplo, basados en el rendimiento o escalas de calificación clínica (Mokkink et al., 2012).

Para la medición de los resultados informados de los pacientes en el área de salud, es vital evaluar la calidad metodológica de los estudios en los que se evalúan sus propiedades psicométricas, dado que, al reportar una alta calidad metodológica, sus conclusiones son más confiables, y para esto, se necesitan estándares para los requisitos de diseño y análisis estadísticos.

### **Metodología**

El presente estudio es documental (Páramo, 2020), de alcance descriptivo (Letelier, Manríquez y Rada, 2005) y el diseño corresponde a una revisión sistemática, mediante la cual se efectúa la recolección, organización, evaluación y sintetización de la mejor evidencia disponible, mediante la implementación de una metodología explícita y rigurosa (Argimón & Jiménez, 2013).

### **Población y muestra**

La presente investigación seleccionó y sistematizó los estudios instrumentales que cumplieron con los *criterios de elegibilidad*, mediante su búsqueda teniendo en cuenta los *criterios de búsqueda* y se tomó la información relevante con base a los *criterios de extracción*. La población correspondió a los estudios preseleccionados en la fase 1 del protocolo de estrategia de búsqueda y la muestra incluyó los estudios que cumplieron con los criterios de elegibilidad de la fase 2 del proceso de selección.

### **Variables y/o indicadores del estudio**

La información se organizó entre sub-apartados (variables) que corresponden a los aspectos evaluados de manera general y los elementos que permitieron la evaluación (indicadores). El primero corresponde a las características generales de los estudios. Algunos indicadores que se extrajeron fueron: autor(es), año, instrumento asociado, tipo de estudio, etc. El segundo al análisis de los indicadores bibliométricos y reporte de las características de los instrumentos de medición, tercero análisis de la adecuación conceptual de estos instrumentos y finalmente, reporte de las propiedades psicométricas de las medidas identificadas. Las variables e indicadores del estudio se encuentran a detalle en la tabla 3.

Tabla 3

*Matriz de variables e indicadores del estudio*

Variables	Indicador(es)	Tabla asociada
Características generales de los estudios	<ul style="list-style-type: none"> <li>▪ Autor(es)</li> <li>▪ Año</li> <li>▪ Instrumento(s) asociado(s)</li> <li>▪ Tipo de estudio</li> <li>▪ País: <ul style="list-style-type: none"> <li>Autor de correspondencia</li> <li>Institución de autor de correspondencia</li> <li>Coautoría internacional</li> </ul> </li> <li>▪ Muestra: <ul style="list-style-type: none"> <li>Tipo</li> <li>Nacionalidad</li> <li>Número</li> <li>Edad (M)</li> </ul> </li> </ul>	Tabla 4. Caracterización de los estudios
Potencial impacto de los estudios instrumentales sobre el análisis de los indicadores bibliométricos	<ul style="list-style-type: none"> <li>▪ Autor(es)</li> <li>▪ Año</li> <li>▪ Instrumento(s) asociado(s)</li> <li>▪ Citas (<i>n</i>)</li> <li>▪ Revista <ul style="list-style-type: none"> <li>Nombre</li> <li>País de origen</li> <li>Institución/compañía editorial</li> </ul> </li> <li>▪ Idioma</li> <li>▪ Categoría: <ul style="list-style-type: none"> <li>Scimago journal &amp; country rank (índice H, dominio y cuartil)</li> <li>Journal Citation Report (Factor de impacto, dominio, cuartil y frecuencia de publicación)</li> </ul> </li> </ul>	-
Caracterización de los instrumentos	<ul style="list-style-type: none"> <li>▪ Autor(es)</li> <li>▪ Aplicabilidad: <ul style="list-style-type: none"> <li>Tipo de administración</li> <li>Modalidad</li> <li>Duración</li> </ul> </li> <li>▪ ítems (<i>n</i>)</li> <li>▪ Subescalas (<i>n</i>)</li> <li>▪ Componentes (<i>n</i>)</li> <li>▪ Nombre de las subescalas</li> <li>▪ Nombre de los componentes</li> <li>▪ Escala de medida</li> <li>▪ Etiquetas verbales</li> <li>▪ Tipo del banco de reactivos</li> </ul>	Tabla 5. Caracterización de los instrumentos
Análisis de la adecuación conceptual de las medidas	<ul style="list-style-type: none"> <li>▪ Instrumento</li> <li>▪ Dimensiones sintomáticas del constructo: <ul style="list-style-type: none"> <li>Cognitivo</li> <li>Conductual</li> <li>Afectivo</li> <li>Somático/fisiológico</li> </ul> </li> <li>▪ Tipo de reporte: <ul style="list-style-type: none"> <li>Gravedad</li> <li>Intensidad</li> <li>Cronicidad</li> </ul> </li> </ul>	-
Análisis de las propiedades psicométricas de los instrumentos	<ul style="list-style-type: none"> <li>▪ Autor(s)</li> <li>▪ Media (DT)</li> <li>▪ Correlación ítem-escala</li> <li>▪ Varianza explicada</li> <li>▪ Autor(es)</li> <li>▪ Análisis Factorial</li> </ul>	<p>Tabla 6. Propiedades psicométricas de los instrumentos: diferenciación entre sujetos</p> <p>Tabla 7. Propiedades psicométricas de los instrumentos: Reportes de análisis factorial y confiabilidad</p>



*Matriz de variables e indicadores del estudio*

Variables	Indicador(es)	Tabla asociada
	Análisis Factorial Exploratorio (estadístico de adecuación de los datos, método de extracción, método de rotación y factores) Análisis Factorial Confirmatorio (índice de bondad de ajuste, modelo corroborado y factores)	
	<ul style="list-style-type: none"> <li>▪ Confiabilidad</li> <li>▪ Autor(es)</li> <li>▪ Validez externa</li> </ul>	Tabla 8. Propiedades psicométricas de los instrumentos: validez externa
	Validez de criterio (validez concurrente, validez predictiva, punto de corte, sensibilidad y especificidad) Validez convergente Validez discriminante	

*Fuente:* Autoría propia

**Instrumentos y técnicas**

En el marco del proceso de revisión sistemática, se tuvo en cuenta las consideraciones establecidas en el *Manual Cochrane de Revisiones Sistemáticas de Intervenciones* (Higgins & Green, 2011), el *Manual Cochrane de Revisiones Sistemáticas de Validez de Pruebas Diagnósticas* (Deeks et al., 2013), *Elementos de Informes Preferidos para revisiones sistemáticas y meta-análisis* (PRISMA por sus siglas en inglés) (Moher et al., 2009), las Directrices para revisiones sistemáticas de los *Estándares basados en el censo para la selección de instrumentos de medición de la Salud* (COSMIN por sus siglas en inglés) (Prinsen et al., 2018), y las directrices establecida en las *Normas para el Desarrollo y Revisión de Estudios Instrumentales* (Carretero-Dios y Pérez, 2005).

**Procedimiento**

Para el desarrollo de los pasos de la revisión sistemática se tuvo en cuenta las directrices establecidas por los “Estándares basados en el censo para la selección de instrumentos de medición de la Salud” (COSMIN) (Prinsen et al., 2018), dado que se ajustan a los objetivos planteados en la presente investigación. Consta de 10 pasos subdivididos en tres partes (A, B y C), de la siguiente manera:

**Parte A. Búsqueda bibliográfica.** Consta de los pasos 1 al 4.

***Paso 1. Formulación del objetivo de la revisión.*** Previendo la gran importancia técnica que representa la pregunta y el objetivo de la revisión, en la presente investigación se delimitaron los siguientes elementos: 1. Constructo (depresión), 2. Población (adultos de 18 a 65 años), 3. Tipo de instrumento (desarrollo, uso, validación o adaptación) y 4. Propiedades de medición de interés (adecuación conceptual, aplicabilidad, construcción y evaluación cualitativa de los ítems y propiedades psicométricas).

El objetivo de la presente investigación es analizar la evidencia científica sobre la adecuación conceptual, aplicabilidad, construcción y evaluación cualitativa de los ítems y propiedades psicométricas de los instrumentos de medición de sintomatología depresiva empleados en la práctica clínica e investigativa a adultos de 18 a 65 años.

***Paso 2. Formulación de los criterios de elegibilidad.*** Los criterios de elegibilidad (criterios de inclusión y exclusión) representan una combinación de aspectos de la pregunta problema y elementos de las especificaciones del tipo de diseño del estudio (Higgins y Green, 2011). Atendiendo a las especificaciones, la presente investigación seleccionó y sistematizó los estudios instrumentales que cumplieran con los siguientes *criterios de elegibilidad*:

*Criterios de inclusión.* Constructo de medición (sintomatología depresiva), tipo de estudios (estudios originales sobre el desarrollo, uso, validación o adaptación de instrumentos y estudios que den reporte de la funcionalidad del instrumento), población objetivo (adultos hispanohablantes de 18 a 65 años) e idioma (español e inglés).

*Criterios de exclusión.* Constructo de medición (sintomatología psicótica), población objetivo (pacientes geriátricos y pacientes hospitalizados).

**Paso 3. Búsqueda bibliográfica.** Para el proceso de búsqueda de estudios, se sugiere la inclusión de un diseño de búsqueda influenciado directamente por los criterios de elegibilidad (inclusión y exclusión), es decir, basándose en los principales conceptos de la revisión (Higgins y Green, 2011; de Vet, Eisinga, Riphagen, Aertgeerts & Pewsner, 2008).

*Criterios de búsqueda.* Acorde a las directrices dispuestas por Prinse et al. (2018) con relación a las bases de datos idóneas, y las consideraciones sobre el empleo de términos de búsqueda de Higgins y Green (2011) y Vet et al. (2008), se utilizaron los siguientes aspectos para el proceso de búsqueda de los estudios, agrupados de la siguiente manera:

- a) Bases de datos: Scopus, Web of Science y ScienceDirect.
- b) Términos metodológicos: propiedades psicométricas, adaptación, medición, validación, diseño, confiabilidad, validez, escala, cuestionario, test, instrumento, checklist e inventario.
- c) Términos clínicos del constructo: trastornos afectivos, trastornos del humor, depresión, trastorno depresivo, trastorno depresivo mayor, trastorno depresivo persistente y distimia.
- d) Operador booleano “And”, sugerido por la Colaboración Cochrane en cuanto a las series de términos agrupados (Higgins y Green, 2011)

Con base a lo anterior, se diseñó un banco de “Códigos de Búsqueda” que dirigió el proceso de la búsqueda inicial de los estudios instrumentales, se efectuó la asignación de números a cada término dando lugar a una serie de códigos, para de esta manera favorecer al proceso de búsqueda sistemática y evitar sesgos.

**Paso 4. Selección de resúmenes y artículos de texto completo.** El proceso de obtención de los estudios constó de dos subfases:

- **Preselección de las potenciales investigaciones:** El proceso de búsqueda inicial se realizó mediante la selección por títulos y resúmenes de posibles investigaciones con el objetivo de evitar excluir potenciales artículos y descartar aquellos claramente irrelevantes. La búsqueda se registró en una tabla denominada “Preselección de los estudios (Almacenamiento inicial)” teniendo en cuenta la siguiente información: autor(es), título de la investigación, año de publicación, resumen/abstract, país, idioma, revista y DOI.
- **Selección de los estudios que conformarían la revisión:** La lectura y análisis del texto completo se llevó a cabo por la evaluación de dos jueces expertos y el investigador principal a través de una tabla denominada “Selección de los estudios identificados” para la verificación del grado de cumplimiento de los criterios de elegibilidad mediante el planteamiento de preguntas con respecto a variables de interés como: constructo de medición, tipo de estudio, muestra y tipo de artículo.

Los desacuerdos sobre la inclusión de un estudio se resolvieron mediante discusión entre el evaluador principal y los jueces expertos. La información del número de estudios identificados, incluidos y excluidos en el proceso de revisión, se sintetizaron en el *diagrama de flujo* (Moher et al., 2009).

**Parte B. Evaluación de las propiedades psicométricas.** Consta de los pasos 5, 6 y 7. Comprende el proceso de evaluación de las propiedades de medición.

En esta fase, la calidad metodológica se evaluó utilizando la lista de verificación de Riesgo de Sesgo de COSMIN, cada estudio se califica como de calidad muy buena, adecuada, dudosa o inadecuada (Mokkink, de Vet, Prinsen, Patrick, Alonso, Bouter y Terwee, 2017) y el

resultado de cada estudio individual sobre una propiedad psicométrica se calificó según los criterios actualizados.

***Paso 5. Evaluación de la validez de contenido.*** Dada la importancia de la construcción de los ítems para el constructo y su adecuada representación de cada uno de los componentes descritos en la definición semántica, la validez de contenido se evaluó teniendo en cuenta los siguientes aspectos: definición semántica de la variable, que comprende la conceptualización y operacionalización del constructo, dimensiones e indicadores (Escobar-Pérez y Cuervo-Martínez, 2008; Lord y Novick, 1968), evaluación por juicio de expertos (criterios de selección, número de jueces) estadístico empleado (valor del rango o estadístico), evaluación de la relevancia, pertinencia, coherencia y suficiencia de los ítems y componentes.

***Paso 6. Evaluación de la estructura interna.*** La estructura interna refiere a la validez estructural y la fiabilidad. La estructura interna comprende (incluida la unidimensionalidad) el uso del análisis factorial (exploratorio y confirmatorio), método de extracción y rotación, factores, índices de bondad de ajuste empleados y modelo corroborado. El análisis de la fiabilidad (estabilidad temporal, formas paralelas, división por mitades y consistencia interna) y estadístico empleado.

***Paso 7. Evaluación de las propiedades de medición restantes.*** Comprende error de medición, validez de criterio (concurrente y predictiva) y pruebas de hipótesis para validez de constructo (convergente y discriminante).

Para el proceso de extracción de la información se diseñó un formulario digital denominado “Cuadro de extracción de datos” que contiene preguntas específicas con respecto a tres aspectos importantes: descripción general del estudio, aplicabilidad del instrumento,

evaluación de propiedades psicométricas (construcción y evaluación cualitativa de los ítems y análisis psicométrico).

**Parte C. Selección de los instrumentos.** Consta de los pasos 8, 9 y 10 y se refiere a la evaluación de la interpretabilidad y viabilidad de los estudios, la formulación de recomendaciones y la presentación de informe de la revisión sistemática.

***Paso 8. Descripción de la interpretabilidad y viabilidad.*** La interpretabilidad corresponde al grado en el que se puede asignar un significado cualitativo a las puntuaciones cuantitativas obtenidas, corresponde al análisis de los resultados obtenidos en el proceso de extracción de la información. La viabilidad se refiere a la facilidad de aplicación de los instrumentos evaluados en aspectos como el tiempo, costo, longitud, tipo y facilidad de administración.

***Paso 9. Formulación de las recomendaciones.*** Las recomendaciones sobre el instrumento más adecuado para su uso en un contexto evaluativo se formularon con respecto a la sintomatología de interés, información de propiedades psicométricas (dirigido a clínicos e investigadores) en términos de avances y/o falencias en la evaluación de la depresión, datos sobre otros tipos de análisis para evaluar la calidad de otros instrumentos y revistas más favorables para consultar con mayor frecuencia, entre otros datos relevantes.

***Paso 10. Reporte de la revisión sistemática.*** Este paso consta del reporte de la información más relevante siguiendo las indicaciones de las directrices establecidas por los manuales, que se ampliarán en el apartado de resultados, que abarca: resultados de la búsqueda bibliográfica y selección de los estudios, número de estudios e instrumentos incluidos en la revisión, características de los estudios e instrumentos incluidos, característica de la muestra,

calidad metodológica de cada estudio; y resumen de hallazgos, que se utilizan para finalmente proporcionar recomendaciones para la selección del instrumento más adecuado según el propósito determinado o contexto particular.

## Resultados

En el presente apartado, se muestran los principales hallazgos derivados del proceso de revisión sistemática sobre la adecuación conceptual, aplicabilidad y propiedades psicométricas de los instrumentos de medición de sintomatología depresiva, comúnmente empleados en la práctica clínica e investigativa en población hispanohablante, en concordancia con los objetivos planteados para el presente estudio. Este apartado se encuentra dividido en 4 secciones: Inicialmente se reportan las características generales de los estudios, seguidamente del análisis de los indicadores bibliométricos y reporte de las características de los instrumentos de medición. Posteriormente, se evidencian los principales hallazgos acerca de la adecuación conceptual de estos instrumentos y finalmente, se reportan sus propiedades psicométricas de las medidas identificadas.

### Población y muestra

Para la búsqueda de los estudios, se diseñó un banco de códigos de búsqueda que generó un total de 1.155 códigos. La selección de los estudios constó de dos fases: En la fase uno, se preseleccionaron 1.793 estudios sólo por título y resumen, 1446 eran duplicados, resultando 347 investigaciones elegibles. En la fase dos, los estudios preseleccionados fueron evaluados por el investigador principal y dos jueces expertos, de los cuales 53 cumplieron a cabalidad con los criterios de selección y fueron parte del proceso de evaluación, con un total de 24 instrumentos de medición asociados (*ver figura 1*).



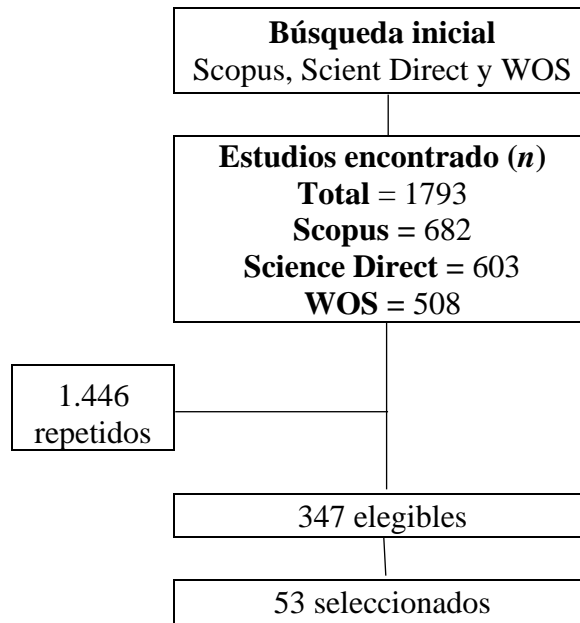


Figura 1. Diagrama de flujo. Fuente: Autoría propia.

### Características generales de los estudios e instrumentos

El rango de publicación de los estudios seleccionados fue entre 1991 y 2017, se observa un incremento del número de publicaciones en los últimos años. La mayoría de las investigaciones se realizaron en Estados Unidos ( $n=17$ ). El 62,3% de las investigaciones tuvo como objetivo principal el análisis general de propiedades psicométricas, siendo el Inventario de Depresión de Beck (BDI;  $n=10$ ) y la Escala de Depresión del Centro de Estudios Epidemiológicos (CES-D;  $n=9$ ) los instrumentos de medición más estudiados. Con respecto a la muestra, el 59% de los estudios utilizaron muestra no clínica. Los tamaños muestrales oscilaron entre 60 a 1.595 participantes para muestras clínicas y entre 54 a 16.415 participantes para las no clínicas. Las características de los estudios se describen a detalle en la tabla 4.

Tabla 4

*Caracterización de los estudios*

Autores	Año	Instrumento(s) asociado	Tipo de estudio	Muestra			
				Tipo	Nacionalidad	<i>n</i>	Edad (M)
Abello-Luque et al.	2017	EMTA	Propiedades psicométricas	Clínica	Colombia	60	37
Porter et al.	2017	HRSD	Propiedades psicométricas	Clínica	Hispanos/Latinos	364	39.7 (Muestra 1) 38.3 (Muestra 2)
Senín-Calderón et al.	2017	LEID-S	Validación	Clínica	España	103	32.66
González et al.	2017	CES-D	Propiedades psicométricas	Población general	Hispanos/Latinos	16.415	41.3
Arbona et al.	2017	CES-D	Medición	Estudiantes	Hispanos/Latinos	858	22
Ruíz et al.	2017	DASS	Propiedades psicométricas	Estudiantes	Colombia	1.273	26.74 (Muestra 1)
				Población general	España	1.707	21.16 (Muestra 2) 34.74 (Muestra 3) 29.16 (Muestra 4)
Valdés et al.	2017	BDI	Propiedades psicométricas	Clínica	Chile	99	41.31 (Muestra 1)
				Estudiantes		73	42.17 (Muestra 2)
				Comunidad		932	
Barajas et al.	2017	CBAS	Adaptación	Clínica	España	100	36.21 (Muestra 1) 20.43 (Muestra 2)
				Población general		100	35.59 (Muestra 3)
				Estudiantes		54	
Tse et al.	2017	DSSS	Propiedades psicométricas	Estudiantes	Hispanos/Latinos	491	19.26
González y Ibañez	2017	MASQ	Propiedades psicométricas	Población general	España	493	29.08
García et al.	2017	ADDI	Propiedades psicométricas	Estudiantes	Hispanos/Latinos	1024	20.83 (Mujeres)
				Comunidad		454	21.36 (Hombres)
Camacho et al.	2016	DASS	Propiedades psicométricas	Estudiantes	Hispanos/Latinos	505	23.92
Dominguez-Lara	2016	IDER	Propiedades psicométricas	Clínica	Perú	103	31.72
Williams et al.	2016	DFFS	Propiedades psicométricas	Clínica	Inglaterra	478	43.8
					España		
					Suecia		
					Francia		
					Alemania		
Londoño et al.	2016	CDH	Propiedades psicométricas	Clínica	Colombia	118	18 a 65
				No clínica		108	
Soria et al.	2016	CORE	Adaptación	Clínica	España	151	59.9
Cano et al.	2015	MADRS	Validación	Clínica	Colombia	150	50
Rizvi et al.	2015	DARS	Diseño	Comunidad	Hispanos/Latinos	229	41.7 (Estudio 1)
				Comunidad (online)		150	27.5 (Estudio 2)
				Clínica		102	
González et al.	2015	BDI	Adaptación	Estudiantes	México	410	28.62
				Comunidad		205	
Talkovsky y Norton	2015	MADRS	Propiedades psicométricas	Estudiantes	Hispanos/Latinos	922	21.95 (Muestra 1) 23.26 (Muestra 2) 21.45 (Muestra 3)

González y Jenkins	2014	CES-D BDI IDS IDAS	Propiedades psicométricas	Estudiantes	Hispanos/Latinos	218	20.65 (Muestra 4) 21.88 (Muestra 5) 20.2
Bentley et al.	2014	ODSIS	Validación	Clínica Estudiantes Comunidad	Hispanos/Latinos	100 566 189	31.59
Gili et al.	2014	IDS	Propiedades psicométricas	Clínica	España	1.595	47.7
Trujols et al.	2014	IDS	Propiedades psicométricas	Clínica	España	173	50.2
Agudelo et al.	2014	IDER	Propiedades psicométricas	Población general	Colombia	1.073	36.5
Abello-Luque et al.	2013	EMTA	Propiedades psicométricas	Estudiantes	Colombia	384	17 a 26
Trujols et al.	2013	CUDOS	Propiedades psicométricas	Clínica	España	162	52
Agüera-Ortiz et al.	2013	CUDOS	Adaptación	Clínica	España	305	51.75
Salinas-Rodríguez et al.	2013	CES-D	Propiedades psicométricas	Población general	México	301	45.27
Ruiz-Grosso et al.	2012	CES-D ZSDS	Validación	Clínica	Perú	194	32.5 (Muestra 1) 32 (Muestra 2) 34 (Muestra 3)
Terrones-González et al.	2012	CES-D	Propiedades psicométricas	Estudiantes	México	943	18 a 29
Antúñez y Vinet	2012	DDASS	Validación	Estudiantes	Chile	484	20.2
Osman et al.	2012	DASS	Propiedades psicométricas	Estudiantes	Hispanos/Latinos	1.297	19.46
Whisman et al.	2012	BDI	Medición	Estudiantes	Hispanos/Latinos	7.369	20
Vargas et al.	2011	K-10	Validación	Población general	México	280	39 (Mujeres) 41 (Hombres)
Barraca et al.	2011	BADS	Validación	Clínica No clínica	España	124 139	24.05 (Muestra 1) 38.5 (Muestra 2)
Gili et al.	2011	IDS	Propiedades psicométricas	Clínica	España	1.595	47.7
Rivera-Medina et al.	2010	CES-D	Propiedades psicométricas	Población general	Puerto Rico	3.504	39.75
Bojorquez y Salgado	2009	CES-D	Propiedades psicométricas	Mujeres	México	468	35
Herrero y Gracia	2007	CES-D	Propiedades psicométricas	Población general	España	1.051	39
Sanz y García-Vera	2007	BDI	Propiedades psicométricas	Clínica Población general Estudiantes	España	338 445 1.393	32.9 (Muestra 1) 42.5 (Muestra 2) 22.2 (Muestra 3)
De la Garza et al.	2007	IDS	Propiedades psicométricas	Clínica	Hispanos/Latinos	665	42.7
Campo et al.	2006	ZSDS	Validación	Población general	Colombia	266	37.4
Campo-Arias et al.	2006	ZSDS	Validación	Población general	Colombia	266	37.4
Agudelo et al.	2005	ST/DEP	Propiedades psicométricas	Estudiantes	España	300	21.82
Trivedi et al.	2004	IDS	Propiedades psicométricas	Clínica	Hispanos/Latinos	946	42.2 (Muestra 1) 40.3 (Muestra 2)
Bonilla et al.	2004	BDI	Propiedades psicométricas	Estudiantes	Puerto Rico	351	21.65
Lobo et al.	2002	MADRS	Validación	Clínica	España	214	47.6 (Muestra 1) 41.6 (Muestra 2)
Daza et al.	2002	DASS	Validación	Bilingües	Hispanos/Latinos	98	40
Arnau et al.	2001	BDI	Propiedades psicométricas	Atención primaria	Hispanos/Latinos	340	36.5
Lasa et al.	2000	BDI	Medición	Población general	España	1.250	18 a 29 (Muestra 1) 30 a 64 (Muestra 2)
Bonicatto et al.	1998	BDI	Propiedades psicométricas	Población general	Argentina	608	20 a 30 (Muestra 1) 31 a 45 (Muestra 2)

Polaino y Senra	1991	HRSD ZSDS BDI	Medición	Clínica	España	47	46 a 60 (Muestra 3) 60 a 75 (Muestra 4) 34.4
-----------------	------	---------------------	----------	---------	--------	----	--

---

*Nota.* EMTA = Escala Multidimensional de los Trastornos Afectivos; HRSD = Escala de Valoración de Hamilton para la Evaluación de la Depresión; LEIDS-R = Escala Leiden de Sensibilidad a la Depresión Revisada; CES-D = Escala de Depresión del Centro de Estudios Epidemiológicos; DASS = Escalas de Depresión, Ansiedad y Estrés; BDI = Inventario de Depresión de Beck; CBAS = Escala Cognitivo-Conductual de Evitación; DSSS = Depression and Somatic Symptoms Scale; MASQ = Cuestionario de Síntomas de Ansiedad y del Humor; ADDI = The Anxiety Depression Distress Inventory; IDER = Inventario de Depresión Estado Rasgo; DFFS = Depression and Family Functioning Scale; CDH = Cuestionario de Depresión para Hombres; CORE = Clinical Outcomes in Routine Evaluation; MADRS = Escala de Depresión de Montgomery-Asberg; DARS = Dimensional Anhedonia Rating Scale; IDS = Inventario de Sintomatología Depresiva; IDAS = Inventory for Depression and Anxiety Symptoms; ODSIS = Escala de gravedad y deterioro general de la depresión; CUDOS = Clinically Usefull Depression Outcome Scale; ZSDS = Escala Autoaplicada de Depresión de Zung; K-10 = Escala de malestar psicológico de Kessler; BADS = Escala de Activación Conductual para la Depresión; ST/DEP = Cuestionario de Depresión Estado-Rasgo.

En lo que respecta a los instrumentos identificados, se encontró que 11 son escalas, la mayoría de estas autoaplicadas, compuestas por bancos de ítems con hasta 90 reactivos, divididos en subescalas (de 2 a 12, a excepción del BDI) y de 1 a 13 componentes/factores. A nivel individual, el 90,6% de los ítems evalúan con escalas continuas, y las etiquetas verbales más utilizadas para medir la sintomatología depresiva fueron de intensidad y temporalidad. El 96,2% de los estudios administraron las pruebas utilizando el modo tradicional de papel y lápiz, en único momento, con tiempos de aplicación entre los 3 y 20 minutos. El estudio de Bonilla et al. (2004) tuvo una metodología de aplicación diferente, en consonancia con los análisis que propuso, en el que si bien no se subdividieron los reactivos en subescalas, si se administraron 3 secciones diferentes del BDI, a subgrupos de la muestra. La tabla 5 brinda a continuación mayor detalle de las características de los instrumentos evaluados.

Tabla 5

*Caracterización de los instrumentos*

Autores	Aplicabilidad			Ítems (n)	Subescalas (n)	Componentes (n)	Escala de medida	Etiquetas verbales	Tipo del banco de reactivos
	Tipo de administración	Modalidad	Duración						
Abello-Luque et al. (2017)	Autoaplicada	Lápiz/papel	-	85	2	-	Continuo	Temporalidad	-
Abello-Luque et al. (2013)	Autoaplicada	Lápiz/papel	-	85	2	10	Continuo	Temporalidad	No específica
Cano et al. (2015)	Heteroaplicada	Lápiz/papel	15-20 min	10	-	-	Continuo	Intensidad	Traducido
Lobo et al. (2002)	Heteroaplicada	Lápiz/papel	-	10	-	-	Continuo	Intensidad	-
Porter et al. (2017)	Heteroaplicada	Lápiz/papel	-	24	-	-	Continuo	-	-
				17					
Polaino y Senra (1991)	Heteroaplicada	Lápiz/papel	-	17	-	-	Continuo	Intensidad	-
Senín-Calderón et al. (2017)	Autoaplicada	Lápiz/papel	-	34	6	-	Continuo	Intensidad	-
Trujols et al. (2013)	Autoaplicada	Lápiz/papel	3 min	16	-	-	Continuo	Temporalidad	Traducido
Agüera-Ortiz et al. (2013)	Autoaplicada	Lápiz/papel	-	16	-	-	Continuo	Intensidad	Traducido
Arbona et al. (2017)	Autoaplicada	Lápiz/papel	-	20	-	4	Continuo	Temporalidad	No específica
Arbona et al. (2017)	Autoaplicada	Lápiz/papel	-	20	-	4	Continuo	Temporalidad	No específica
González y Jenkins (2014)	Autoaplicada	Lápiz/papel	-	20	-	-	-	-	No específica
Ruiz-Grosso et al. (2012)	Autoaplicada	Lápiz/papel	-	20	-	-	Continuo	-	No específica
Rivera-Medina et al. (2010)	Autoaplicada	Lápiz/papel	-	20	-	4	Continuo	Temporalidad	No específica
Bojorquez y Salgado (2009)	Autoaplicada	Lápiz/papel	-	20	-	-	Continuo	Temporalidad	No específica
				10					
González et al. (2017)	Autoaplicada	Lápiz/papel	-	10	-	-	Continuo	-	No específica
Terrones-González et al. (2012)	Autoaplicada	Lápiz/papel	-	23	-	-	Continuo	Temporalidad	No específica
Herrero y Gracia (2007)	-	Lápiz/papel	-	7	-	-	-	-	No específica
Salinas-Rodríguez et al. (2013)	Autoaplicada	Lápiz/papel	-	7	-	-	Dicotómico	Presencia/ausencia	No específica
Ruíz et al. (2017)	Autoaplicada	Lápiz/papel	-	21	3	5	Continuo	Temporalidad	No específica
		Informático							
Camacho et al. (2016)	Autoaplicada	Lápiz/papel	-	21	3	-	Continuo	Temporalidad	No específica
Antúnez y Vinet (2012)	Autoaplicada	Lápiz/papel	-	21	3	5	Continuo	Temporalidad	No específica
Osman et al. (2012)	Autoaplicada	Lápiz/papel	-	21	3	-	Continuo	Temporalidad	No específica
Daza et al. (2002)	Autoaplicada	Lápiz/papel	-	21	3	-	Continuo	-	Traducido
Bentley et al. (2014)	Autoaplicada	Lápiz/papel	-	5	-	1	Continuo	Intensidad	No específica
Williams et al. (2016)	Autoaplicada	Lápiz/papel	-	15	-	7	Continuo	Temporalidad	No específica

*Caracterización de los instrumentos*

Autores	Aplicabilidad			Ítems (n)	Subescalas (n)	Componentes (n)	Escala de medida	Etiquetas verbales	Tipo del banco de reactivos
	Tipo de administración	Modalidad	Duración						
Vargas et al. (2011)	Autoaplicada	Lápiz/papel	-	10	-	-	Continuo	Temporalidad	No específica
Ruiz-Grosso et al. (2012)	Autoaplicada	Lápiz/papel	-	20	-	-	Continuo	-	No específica
Campo-Arias et al. (2006)	Autoaplicada	Lápiz/papel	-	20	-	-	Continuo	Temporalidad	No específica
Polaino y Senra (1991)	Autoaplicada	Lápiz/papel	-	20	-	-	Continuo	Intensidad	No específica
Campo et al. (2006)	Autoaplicada	Lápiz/papel	-	10	-	-	Continuo	Temporalidad	No específica
Barajas et al. (2017)	Autoaplicada	Lápiz/papel	-	31	-	-	Continuo	-	Traducido
Barraca et al. (2011)	Autoaplicada	Lápiz/papel	-	25	-	4	Continuo	Temporalidad	Traducido
San Tse et al. (2017)	Autoaplicada	Lápiz/papel	-	22	-	-	Continuo	Intensidad	No específica
Rizvi et al. (2015)	Autoaplicada	Lápiz/papel Informático	5 min	17	-	4	Continuo	Temporalidad	Creado
Lasa et al. (2000)	Autoaplicada	Lápiz/papel	-	-	-	-	Continuo	Intensidad	No específica
Bonicatto et al. (1998)	Autoaplicada	Lápiz/papel	-	21	-	-	Continuo	Intensidad	No específica
Polaino y Senra (1991)	Autoaplicada	Lápiz/papel	-	21	-	-	Continuo	Intensidad	No específica
Valdés et al. (2017)	Autoaplicada	Lápiz/papel	-	21	-	-	Continuo	Intensidad	No específica
Sanz y García-Vera (2007)	Autoaplicada	Lápiz/papel	5-10 min	13	-	13	Continuo	Intensidad	No específica
Sanz y García-Vera (2007)	Autoaplicada	Lápiz/papel	5-10 min	13	-	13	Continuo	Intensidad	No específica
González et al. (2015)	Autoaplicada	Lápiz/papel	-	21	-	-	Continuo	-	Traducido
González y Jenkins (2014)	Autoaplicada	Lápiz/papel	-	21	-	-	-	-	No específica
Whisman et al. (2012)	Autoaplicada	Lápiz/papel	-	21	-	-	Continuo	Intensidad	No específica
Arnau et al. (2001)	Autoaplicada	Lápiz/papel	-	21	-	-	Continuo	Intensidad	No específica
Bonilla et al. (2004)	Autoaplicada	Lápiz/papel	-	37	3 secciones	12	Continuo	Intensidad	Traducido
González y Jenkins (2014)	Autoaplicada	Lápiz/papel	-	30	-	-	-	-	No específica
Trivedi et al. (2004)	Heteroaplicada	Lápiz/papel	-	30	-	-	Continuo	Intensidad	No específica
Trivedi et al. (2004)	Heteroaplicada	Lápiz/papel	-	16	-	-	Continuo	Intensidad	No específica
Trivedi et al. (2004)	Autoaplicada	Lápiz/papel	-	30	-	-	-	-	No específica
Trivedi et al. (2004)	Autoaplicada	Lápiz/papel	-	16	-	-	Continuo	Intensidad	No específica
Gili et al. (2011)	Autoaplicada	Lápiz/papel	-	30	-	-	Continuo	Temporalidad	No específica
Gili et al. (2014)	Autoaplicada	Lápiz/papel	5-7 min	16	-	9	Continuo	Intensidad	No específica
Trujols et al. (2014)	Autoaplicada	Lápiz/papel	-	16	-	-	Continuo	-	No específica
De la Garza et al. (2007)	Autoaplicada	Lápiz/papel	-	5	-	5	-	-	No específica
Dominguez-Lara (2016)	Autoaplicada	Lápiz/papel	-	20	2	2	Continuo	Intensidad	No específica
Agudelo et al. (2014)	Autoaplicada	Lápiz/papel	-	20	2	2	Continuo	Intensidad	No específica
González y Jenkins (2014)	Autoaplicada	Lápiz/papel	-	64	12	-	-	-	No específica

*Caracterización de los instrumentos*

Autores	Aplicabilidad			Ítems (n)	Subescalas (n)	Componentes (n)	Escala de medida	Etiquetas verbales	Tipo del banco de reactivos
	Tipo de administración	Modalidad	Duración						
García et al. (2017)	Autoaplicada	Lápiz/papel	-	27	3	-	Continuo	Intensidad	No específica
Agudelo et al. (2005)	Autoaplicada	Lápiz/papel	-	32	2	1	Continuo	Temporalidad	No específica
Talkvosky y Norton (2015)	Autoaplicada	Lápiz/papel	-	90	3	-	-	-	No específica
González y Ibañez (2017)	Autoaplicada	Lápiz/papel	-	30	3	-	Continuo	-	No específica
Londoño et al. (2016)	Autoaplicada	Lápiz/papel	-	40	-	6	Continuo	-	No específica
Soria et al. (2016)	Heteroaplicada	Lápiz/papel	-	18	3	-	Continuo	Intensidad	Traducido

*Fuente:* Autoría propia.



## **Potencial impacto de los estudios instrumentales sobre el análisis de los indicadores bibliométricos**

A continuación se relacionan los indicadores bibliométricos de los estudios seleccionados. Se encontraron 53 publicaciones en 36 revistas científicas. La revista *Salud Mental* (de México) fue la más empleada ( $n=6$ ). Tomando como referente el país de origen de las revistas, se pudo determinar que 37 estudios fueron publicados en revistas internacionales y en su mayoría los estudios fueron de publicados en inglés. Buena parte de las revistas se encuentran indexadas en el cuartil Q 1 ( $n=19$ ) del Scimago Journal Rank (SJR), y en los cuartiles Q1 ( $n= 12$ ) y Q 4 ( $n= 13$ ) en el Journal Citation Report (JCR). El número de citas oscilaron entre 0 a 70 y 0 a 768, en revistas latinoamericanas e internacionales, respectivamente.

En cuanto a redes de colaboración, 9 investigaciones colaboraron con países como: Estados Unidos, Inglaterra, Francia, España, Australia, Singapur y Argentina.

## **Análisis de la adecuación conceptual de las medidas**

En el proceso de construcción/adaptación de una prueba, la validez de contenido puede obtenerse a través de tres etapas claves (Delgado-Rico, Carretero-Dios y Ruch, 2012). Inicialmente se define el dominio de contenido a evaluar mediante la *definición semántica de la variable* (Lord y Novick, 1968), que expone la conceptualización y operacionalización del constructo, y se procede con la construcción de los ítems, o la traducción de los mismos (si aplica) y por último, la valoración por un panel o juicio de grupos de expertos (Escobar-Pérez y Cuervo-Martínez, 2008). En la revisión realizada de los estudios, solo se identificaron 3 estudios con análisis relativos a esta etapa (Abello-Luque et al., 2017; Senín-Calderón et al., 2017 y Rizvi et al., 2015).

La traducción de un instrumento es también un paso importante en el proceso de validación que al no realizarse de manera adecuada puede facilitar diversos errores (Ramada-Rodilla, Serra-Pujadas y Delclós-Clanchet, 2012). Es necesario seguir procedimientos que permitan asegurar la equivalencia semántica, idiomática, experiencial y conceptual entre la versión original y la traducida (Arafat, Rahman, Qusar y Hafez, 2016; Gordon, 2004; Hambleton y Jong, 2003; Hambleton, 1996, 1994). En la revisión ejecutada, se pudo identificar que 9 investigaciones llevaron a cabo el proceso de adaptación/traducción de inglés a español e implementaron el método de traducción inversa (Barajas et al., 2017; Soria et al., 2016; Cano et al., 2015; González et al., 2015; Trujols et al., 2013; Agüera-Ortiz et al., 2013; Barraca et al., 2011; Bonilla et al., 2004 Daza et al., 2002). De ellos 8 estudios reportaron evaluación por grupos de comité (Soria et al., 2016; Cano et al., 2015; González et al., 2015; Trujols et al., 2013; Agüera-Ortiz et al., 2013; Barraca et al., 2011; Bonilla et al., 2004; Daza et al., 2002), y tan solo solo 6 estudios describieron cómo se resolvieron las diferencias entre las versiones originales y la traducida (Soria et al., 2016; Cano et al., 2015; González et al., 2015; Agüera-Ortiz et al., 2013; Barraca et al., 2011; Daza et al., 2002).

Por otra parte, se realizó la identificación del modelo teórico que los estudios reportan como el soporte del instrumento, la operacionalización de los constructos de medida y el tipo de valoración de los instrumentos. Se encontró que algunos instrumentos reportan basarse en el DSM ( $n=8$ ; González et al., 2017; Gili et al., 2014; González y Jenkins, 2014; Agüera-Ortiz et al., 2013; Trujols et al., 2014; Trujols et al., 2013; Gili et al., 2011; Trivedi et al., 2004), en sus diferentes versiones (APA, 2014; 1994; 1968); otros estudios ( $n=3$ ; García et al., 2017; González y Jenkins, 2017; Talkvosky y Norton, 2015;) tomaron como referencia el modelo tripartito (Clark y Watson, 1991), el estudio de Agudelo et al. (2005) reporta como base el

modelo de afectividad positiva y afectividad negativa, PANAS (Watson, Clark y Tellegen, 1988), y el estudio de Agudelo et al. (2014) toma como referencia teórica el modelo tripartito y el PANAS.

Respecto al constructo de medida de los instrumentos analizados en función de las dimensiones sintomáticas de la depresión (cognitivo, conductual, afectivo y somático/fisiológico), se encontró que la dimensión afectiva es la sintomatología más evaluada por los instrumentos y en su mayoría se miden gravedad de la depresión. Cabe resaltar que las Escalas de Depresión, Ansiedad y Estrés (DASS-21), el Inventario de Depresión Estado Rasgo (IDER) y el Anxiety Depression Distress Inventory (ADDI-27) no dan evidencia del tipo de medición (gravedad, intensidad o cronicidad).

El objetivo de la medición de los instrumentos varió según su finalidad, siendo las más recurrentes:

- La detección de los síntomas (algunos aplicados en atención primaria),
- La evaluación de los efectos de la psicoterapia y/o los tratamientos antidepresivos en la variabilidad de los síntomas,
- El monitoreo epidemiológico,
- La evaluación del impacto de la depresión en el deterioro disfuncional (algunos a nivel familiar y de pareja) y,
- La valoración de aspectos dimensionales del constructo (reactividad cognitiva, componente afectivo de la depresión, estados negativos emocionales, depresión melancólica, conductas de evitación y activación, síntomas somáticos y anhedonia).

### **Análisis de las propiedades psicométricas de los instrumentos**

Una vez se ha establecido el banco de ítems y su respectiva valoración por expertos, el siguiente paso sugerido por los metodólogos para la construcción de instrumentos psicométricos, consiste en su aplicación a una muestra con características similares a la muestra objetivo, cuyo tamaño adecuado debe oscilar entre 50 a 100 participantes, según Osterlind (1989). Los resultados indican que siete de los 53 estudios, reportan la realización del proceso de pilotaje. De estos siete estudios, dos se centraron en la funcionalidad de los ítems (Abello et al., 2013; y Bonilla et al., 2014), mientras que los cinco restantes realizaron este procedimiento para evaluar el proceso de traducción de los ítems (Soria et al., 2016; Cano et al., 2015; González et al., 2015; Agüera-Ortíz et al., 2013 y Daza et al., 2002).

Un criterio importante para la selección de los ítems adecuados, es la identificación de aquellos reactivos que mejor distingan diferencias inter-sujetos, para ello, se recomienda procurar por un conjunto de ítems que maximicen la varianza, con alta desviación típica, con puntuaciones de medias de respuestas ubicados alrededor de la media de la escala y coeficientes de correlación ítem-escalas con valores mayores o iguales a 0,25-0,30 (Nunnally y Berstein, 1995). Al respecto, en la revisión se identificó que el estudio de Abello-Luque et al. (2017) reportaron que los ítems 18, 19 y 20 con índices de correlación bajos y en su mayoría inversamente proporcionales, y la investigación de Bojorquez y Salgado (2009) reporta correlaciones negativas con los ítems 4 y 8 que al ser eliminados aumentó el índice de confiabilidad de la escala. La tabla 6 expone los resultados relativos al análisis descriptivo de los ítems reportados por 44 de los 53 estudios revisados, el excedente no dio reporte alguno de este aspecto.

Tabla 6

*Propiedades psicométricas de los instrumentos: diferenciación entre sujetos*

Autores	Media (DT)	Correlación ítem-escala	Varianza explicada
Escala Multidimensional de los Trastornos Afectivos [EMTA]			
Abello-Luque et al. (2017)	88.23 (6.15) (Sub. de Depresión)	0.4 - .78 (Sub. de Depresión)	-
	54.36 (12.06) (Sub. de Manía)	.22 - .62 (Sub. de Manía)	-
Abello-Luque et al. (2013)	71.12 (13.21) (Sub. de Depresión)	> .50	68% (Sub. de Depresión)
	48.74 (9.67) (Sub. de Manía)		65% (Sub. de Manía)
Escala de Depresión de Montgomery-Asberg [MADRS]			
Cano et al. (2015)	-	-	91%
Escala de Valoración de Hamilton para la Evaluación de la Depresión [HRSD]			
Polaino y Senra (1991)	20.58 (2.35) Episodio único	-	-
	21.22 (1.87) Episodio recurrente		
Escala Leiden de Sensibilidad a la Depresión Revisada [LEIDS-R]			
Senín-Calderón et al. (2017)	49,58 (Muestra 1)	> 0.40	42%
	52,39 (Muestra 2)		
Clinically Usefull Depression Outcome Scale [CUDOS]			
Trujols et al. (2013)	16.4 (8.1)	> 0.549	55.7%
Agüera-Ortiz et al. (2013)	29.86 (12.12)	> 0.15	38.2%
Escala de Depresión del Centro de Estudios Epidemiológicos [CES-D]			
González y Jenkins (2014)	20.18 (10.39)	-	-
Escala de Depresión del Centro de Estudios Epidemiológicos [CES-D-10]			
Bojorquez y Salgado (2009)	-	> 0.20	50.6% (CESD-20)
			46.5% (CESD-10)
González et al. (2017)	6.99 (.08)	-	-
Escala de Depresión del Centro de Estudios Epidemiológicos/ Ideación Suicida [CES-D/IS]			
Terrones-González et al. (2012)	-	-	45.5%
Escala de Depresión del Centro de Estudios Epidemiológicos [CESD-7]			
Herrero y Gracia (2007)	12.23 (4.01) (Muestra 1)	> 40	51% (Muestra 1)
	12.85 (4.26) (Muestra 2)		54% (Muestra 2)
Salinas-Rodríguez et al. (2013)	4.64	-	-
Escala de Depresión, Ansiedad y Estrés [DASS]			
Ruiz et al. (2017)	-	-	70%
Camacho et al. (2016)	-	-	53.5%
Antúnez y Vinet (2012)	-	-	49.99%
Osman et al. (2012)	-	-	61.9%
Daza et al. (2002)	-	> .40	83% (Escala de depresión)
			75% (Escala de ansiedad)
			84% (Escala de estrés)
Overall Depression Severity and Impairment Scale [ODSIS]			
Bentley et al. (2014)	5.16 (4.81)	-	-
Depression and Family Functioning Scale [DFFS]			
Williams et al. (2016)	27.83 (10.5) Línea base	-	46%
	23.29 (11.1) 2 meses		
Escala de malestar psicológico de Kessler [K-10]			
Vargas et al. (2011)	-	-	53.4%
Escala Autoaplicada de Depresión de Zung [ZSDS]			

*Propiedades psicométricas de los instrumentos: diferenciación entre sujetos*

Autores	Media (DT)	Correlación ítem-escala	Varianza explicada
Campo et al. (2006)	-	-	36.6%
Campo-Arias et al. (2006)	37.4 (9.1)	-	-
Polaino y Senra (1991)	51.00 (2.58) Episodio único	-	-
	51.33 (2.74) Episodio recurrente		
Escala de Activación Conductual para la Depresión [BADS]			
Barraca et al. (2011)	86.72 (25.44)	-	-
Depression and Somatic Symptoms Scale [DSSS]			
Tse et al. (2017)	36.93 (10.87)	> 0.30	45%
Dimensional Anhedonia Rating Scale [DARS]			
Rizvi et al. (2015)	55.8 (10.6) Estudio 1	-	35.4 % (Factor 1)
	50.0 (12.6) Estudio 2		10.2% (Factor 2)
			8.4% (Factor 3)
			5.2% (Factor 4)
Inventario de Depresión de Beck [BDI]			
Bonicatto et al. (1998)	8.46 (7.34)	> 0.32	28.7%
Polaino y Senra (1991)	25.3 (3.76)	-	-
Inventario de Depresión de Beck/versión corregida [BDI-IA]			
Valdés et al. (2017)	-	-	37%
Sanz y García-Vera (2007)	BDI-IA-SF =	> 0.23	-
	10.5 (7.6) (Muestra clínica)		
	4.1 (4.1) (Muestra población general)		
	3.6 (3.7) (Muestra estudiantes)		
	BDI-IA-SCA =		
	10.5 (7.8) (Muestra clínica)		
	4.3 (4.3) (Muestra población general)		
	3.6 (3.8) (Muestra estudiantes)		
Inventario de Depresión de Beck- Segunda edición [BDI-II]			
González et al. (2015)	9.31 (7.84) (Muestra estudiantes)	> .39	-
	9.82 (7.70) (Muestra comunidad)		
González y Jenkins (2014)	14.66 (9.62)	-	-
Whisman et al. (2012)	9.27 (8.07)	-	-
Arnau et al. (2001)	8.74 (9.7)	> .54	53.5%
Inventario de Depresión de Beck- Spanish [BDI-S]			
Bonilla et al. (2004)	-	-	49%
Inventario de Sintomatología Depresiva [IDS]			
González y Jenkins (2014)	21.98 (10.84)	-	-
Inventario de Sintomatología Depresiva- Clinician Rating [IDS-C]			
Trivedi et al. (2004)	-	≥ 0.65	-
Inventario de Sintomatología Depresiva -Autoreporte [IDS-SR]			
Gili et al. (2011)	34.85 (15) (Primera evaluación)	> 0.20	49.87%
	15.34 (11) (Segunda evaluación)		
Gili et al. (2014)	-	-	46.80%
Inventario Rápido de Sintomatología Depresiva -Autoreporte [QIDS-SR]			
Trujols et al. (2014)	-	> 0.369	53.1%
El Inventario muy rápido de sintomatología depresiva - Autoinforme [VQIDS]			
De la Garza et al. (2007)	8.5 (3.0)	-	-

*Propiedades psicométricas de los instrumentos: diferenciación entre sujetos*

Autores	Media (DT)	Correlación ítem-escala	Varianza explicada
<b>Inventario de Depresión Estado Rasgo [IDER]</b>			
Agudelo et al. (2014)	16.32 (4.86) (Estado) 16.08 (4.53) (Rasgo)	-	51.57%
<b>Inventory for Depression and Anxiety Symptoms [IDAS]</b>			
González y Jenkins (2014)	49.26 (12.23)	-	-
<b>Cuestionario de Depresión Estado-Rasgo [ST/DEP]</b>			
Agudelo et al. (2005)	11.77 (4.35) Mujeres: Distimia 18.95 (2.07) Mujeres: Eutimia 11.22 (4.38) Hombres: Distimia 23.74 (5.55) Hombres: Eutimia	> 0.57	54.12% (Mujeres) 53.25% (Hombres)
<b>Cuestionario de Síntomas de Ansiedad y del Humor [MASQ]</b>			
González y Ibañez (2017)	-	-	39.7%
<b>The Anxiety Depression Distress Inventory [ADDI]</b>			
García et al. (2017)	19.21 (7.34) Estudio 1: Distrés general 27.47 (8.08) Estudio 1: Afecto positivo 14.13 (5.59) Estudio 1: Ansiedad somática 21.30 (8.07) Estudio 2: Distrés general 30.90 (8.20) Estudio 2: Afecto positivo 16.02 (7.25) Estudio 2: Ansiedad somática 21.60 (7.62) Estudio 3: Distrés general 31.57 (7.39) Estudio 3: Afecto positivo 16.34 (6.98) Estudio 3: Ansiedad somática	-	-
<b>Cuestionario de Depresión para Hombres [CDH]</b>			
Londoño et al. (2016)	73.5 (22.2)	-	56.6%
<b>Clinical Outcomes in Routine Evaluation [CORE]</b>			
Soria et al. (2016)	14.85 (9.96)	> 0.32	-

*Fuente:* Autoría propia

Una vez evaluado teórica y estadísticamente el banco de ítems, se sugiere proceder con el análisis de la dimensionalidad (estructura interna) del instrumento para verificar si empíricamente se agrupan tal y como la teoría lo predice, es decir, si la estructura interna representa los componentes o dimensiones del constructo (Carreto-Dios y Pérez, 2007). Tradicionalmente el análisis factorial es la técnica que permite contrastar estadísticamente las hipótesis con base a la agrupación de los ítems. La revisión realizadas permitió identificar, en consonancia con lo anterior, que la mayoría de los estudios implementaron la técnica multivariada del Análisis Factorial Exploratorio (AFE;  $n= 43$ ), y 22 de estos realizaron además el Análisis Factorial Confirmatorio (AFC). Para constatar la adecuación de los datos, 18 investigaciones recurrieron a la Prueba de Esfericidad de Bartlett y el Kaiser-Meyer-Olkin (KMO). El método de extracción de análisis de componentes principales y el método de rotación Promax fueron los más utilizados en los estudios analizados.

Con respecto al AFC, los índices de bondad de ajuste más implementados fueron el Chi Cuadrado escalado entre los grados de libertad (S-B  $\chi^2/df$ ; Satorra y Bentler, 2001), el Índice de Ajuste Comparado (CFI; Bentler, 1990), el Índice de Tucker-Lewis (TLI; Tucker, 1961), el Índice de Bondad de Ajuste (GFI), Error Cuadrático Medio de Aproximación (RMSEA; Steiger, 1980) y Standardized root mean square residual (SRMR).

Con respecto al índice de fiabilidad, el análisis más empleado fue la consistencia interna mediante la aplicación del estadístico Alfa de Cronbach ( $\alpha$ ; Cronbach, 1951). De acuerdo con George y Mallery (2003) un coeficiente  $>.9$  es excelente (pero puede entrañar un exceso de ítems que sobrevalúan un aspecto de la variable),  $>.8$  es bueno,  $>.7$  es aceptable,  $>.6$  es cuestionable,  $>.5$  es pobre, y  $<.5$  es inaceptable. Los estudios de Londoño et al. (2016), Rizvi et al. (2015) y Daza et al., (2002) reportan mayor puntaje con un coeficiente de  $.96$ , seguido de Abello-Luque et



al. (2017), Ruíz, et al. (2017), Barajas et al. (2017) y Talkvozky y Norton (2015), con un puntaje  $>.95$ ; mientras que, los estudios de Arbona et al. (2017), Porter et al. (2017), Antúnez y Vinet (2012), Bojorquez y Salgado (2009) y Sanz y García-Vera (2007) reportan coeficientes  $>.7$  (.73-.79) y el estudio de De la Garza et al. (2007) reporta un puntaje mínimo de .67 en sus resultados. La tabla 7 resume los hallazgos obtenidos de los análisis factoriales y la confiabilidad.

Tabla 7

*Propiedades psicométricas de los instrumentos: Reportes de análisis factorial y confiabilidad*

Autores	Análisis Factorial							Confiabilidad
	Análisis Factorial Exploratorio (AFE)				Análisis Factorial Confirmatorio (AFC)			
	Estadístico de adecuación de los datos	Método de extracción	Método de rotación	Factores (n)	Índice de Bondad de ajuste	Modelo corroborado	Factores (n)	
Escala Multidimensional de los Trastornos Afectivos [EMTA]								
Abello-Luque et al. (2017)	-	-	-	-	-	-	-	$\alpha$ .958 (Sub. de Depresión)
Abello-Luque et al. (2013)	<ul style="list-style-type: none"> <li>▪ KMO .856 (Sub.de depresión)</li> <li>▪ .845 (Sub.de manía)</li> <li>▪ Esfericidad de Barlett <math>\chi^2 = 5809 p &lt; .000</math> (Sub.de depresión)</li> <li>▪ <math>\chi^2 = 2917 p &lt; .000</math> (Sub.de manía)</li> </ul>	ACP	Varimax	-	-	-	-	$\alpha$ .870 (Sub. de Manía) $\alpha$ .942 (Sub. de Depresión) $\alpha$ .864 (Sub. de Manía)
Escala de Depresión de Montgomery-Asberg [MADRS]								
Cano et al. (2015)	<ul style="list-style-type: none"> <li>▪ KMO = 0,892</li> <li>▪ Esfericidad de Barlett = <math>\chi^2 = 1.000,867; gl = 45; p &lt; 0,001</math></li> </ul>	ACP	Promax	1	-	-	-	$\alpha$ 0.91 CCI 0.98
Lobo et al. (2002)	<ul style="list-style-type: none"> <li>▪ KMO = 0,886</li> <li>▪ Esfericidad de Barlett = 473,3; <math>p &lt; 0,01</math></li> </ul>	-	Varimax	1	-	-	-	$\alpha$ 0.88 CCI 0.98
Escala de Valoración de Hamilton para la Evaluación de la Depresión [HRSD]								
Porter et al. (2017)	-	-	Oblimin directo	3 (HRSD-24)	-	-	-	$\alpha$ .77 (HRSD-17) $\alpha$ .81 (HRSD-24)
Polaino y Senra (1991)	-	-	-	-	-	-	-	Entre evaluadores .89
Escala Leiden de Sensibilidad a la Depresión Revisada [LEIDS-R]								
Senín-Calderón et al. (2017)	<ul style="list-style-type: none"> <li>▪ KMO = 0,91</li> <li>▪ Esfericidad de Barlett = <math>\chi^2 (378, 300) = 3420,03; p &lt; 0,001</math></li> </ul>	Mínimos cuadrados no ponderados	Oblimin directo	4	<ul style="list-style-type: none"> <li>▪ <math>\chi^2 = 281,34</math></li> <li>▪ CFI = 0,956</li> <li>▪ TLI = 0,941</li> <li>▪ GFI = 0,929</li> <li>▪ RMSEA = 0,035 [0,024, 0,045]</li> <li>▪ AIC = 471,34</li> </ul>	Multidimensional	4	$\alpha$ 0.87
Clinically Usefull Depression Outcome Scale [CUDOS]								
Trujols et al. (2013)	<ul style="list-style-type: none"> <li>▪ KMO = 0,912</li> <li>▪ Esfericidad de Barlett = <math>\chi^2(78) = 1415.24; p &lt; 0.001</math></li> </ul>	ACP	-	1	-	-	-	$\alpha$ 0.93 CCI 0.84
Agüera-Ortiz et al. (2013)	<ul style="list-style-type: none"> <li>▪ KMO = 0.86, <math>p &lt; 0.001</math></li> </ul>	ACP	Oblimin directo	1	-	-	-	$\alpha$ 0.88
Escala de Depresión del Centro de Estudios Epidemiológicos [CES-D]								
Arbona et al. (2017)	-	-	-	-	<ul style="list-style-type: none"> <li>▪ CFI = <math>\geq .90</math></li> <li>▪ TLI = <math>\geq .90</math></li> <li>▪ SRMR = <math>\leq .10</math></li> <li>▪ RMSEA = <math>\leq .08</math></li> </ul>	Unidimensional	3	$\alpha$ .74 (Preocupaciones somáticas) $\alpha$ .87 (Afecto negativo)

*Propiedades psicométricas de los instrumentos: Reportes de análisis factorial y confiabilidad*

Autores	Análisis Factorial							Confiabilidad
	Análisis Factorial Exploratorio (AFE)				Análisis Factorial Confirmatorio (AFC)			
	Estadístico de adecuación de los datos	Método de extracción	Método de rotación	Factores (n)	Índice de Bondad de ajuste	Modelo corroborado	Factores (n)	
González y Jenkins (2014)	-	-	-	-	-	-	-	$\alpha$ .78 (Falta de afecto positivo/anhedonia) $\alpha$ .89
Ruiz-Grosso et al. (2012)	-	-	-	-	-	-	-	$\alpha$ 0.93 $\omega$ 0.90
Rivera-Medina et al. (2010)	-	-	-	2	<ul style="list-style-type: none"> <li>▪ CFI = .955</li> <li>▪ RMSEA = .048</li> <li>▪ TLI = .948</li> </ul>	-	Resultados no adecuados	-
Escala de Depresión del Centro de Estudios Epidemiológicos [CES-D-10]								
Bojorquez y Salgado (2009)	-	ACP	Varimax	4 (CESD-20) 2 (CESD-10)	-	-	-	$\alpha$ 0.88 (CESD-20) $\alpha$ 0.79 (CESD-10)
González et al. (2017)	-	-	-	-	<ul style="list-style-type: none"> <li>▪ <math>\chi^2 = 1,197.46, df = 34, p &lt; .001</math></li> <li>▪ CFI = .986</li> <li>▪ RMSEA = .047</li> </ul>	Unidimensional	1	$\alpha$ .80-.86 Test-retest .41-.70
Escala de Depresión del Centro de Estudios Epidemiológicos/ Ideación Suicida [CES-D/IS]								
Terrones-González et al. (2012)	<ul style="list-style-type: none"> <li>▪ KMO = 0.92</li> </ul>	ACP	Varimax	4	-	-	-	$\alpha$ 0.89
Escala de Depresión del Centro de Estudios Epidemiológicos [CESD-7]								
Herrero y Gracia (2007)	-	ACP	Varimax	1	-	-	-	$\alpha$ 0.82
Salinas-Rodríguez et al. (2013)	-	-	-	-	-	-	-	$\alpha$ 0.86 (Primera medición) $\alpha$ 0.81 (Segunda medición)
Escalas de Depresión, Ansiedad y Estrés [DASS]								
Ruiz et al. (2017)	-	-	-	-	<ul style="list-style-type: none"> <li>▪ RMSEA = .048 - .076 (Muestra 1) .051 - .070 (Muestra 2) .049 - .092 (Muestra 3)</li> <li>▪ CFI = .98 - .99 (Muestra 1) .98 - .99 (Muestra 2) .97 - .99 (Muestra 3)</li> <li>▪ NNFI = .98 - .99 (Muestra 1) .98 - .99 (Muestra 2) .97 - .99 (Muestra 3)</li> <li>▪ ECVI = .52 - 1.20 (Muestra 1)</li> </ul>	Multidimensional	3	$\alpha$ .92 - .95

*Propiedades psicométricas de los instrumentos: Reportes de análisis factorial y confiabilidad*

Autores	Análisis Factorial							Confiabilidad
	Análisis Factorial Exploratorio (AFE)			Análisis Factorial Confirmatorio (AFC)				
	Estadístico de adecuación de los datos	Método de extracción	Método de rotación	Factores (n)	Índice de Bondad de ajuste	Modelo corroborado	Factores (n)	
					0.65 -1.10 (Muestra 2) .049 -.092 (Muestra 3) ▪ SRMR = .038 -.053 (Muestra 1) .043 -.051 (Muestra 2) .041 -.062 (Muestra 3)			
Camacho et al. (2016)	▪ KMO = 0.95 ▪ Esfericidad de Barlett: 4960.9, (df = 210); p < 0.01	-	-	1	-	-	Resultado s no adecuados	$\omega$ 0.95
Antúñez y Vinet (2012)	▪ KMO = .909 ▪ Esfericidad de Barlett = .000	ACP	Oblimin directo	3	-	-	-	$\alpha$ .85 (Escala de depresión) $\alpha$ .83 (Escala de estrés) $\alpha$ .73 (Escala de ansiedad)
Osman et al. (2012)	-	ACP	Promax	3	▪ CFI = .830 ▪ TLI = .812 ▪ RMSEA = .072 ▪ SRMR = .067 ▪ AIC = 17,373.32	Multidimensional	3	$\omega$ .89
Daza et al. (2002)	-	-	-	-	▪ RMSEA = .12 ▪ NFI = .73 ▪ NNFI = .80 ▪ PNFI = .60	Multidimensional	3	$\alpha$ .96
Overall Depression Severity and Impairment Scale [ODSIS]								
Bentley et al. (2014)	-	-	-	-	▪ $\chi^2(5) = 7.956, p = .159$ ▪ RMSEA = .077 ▪ SRMR = .028 ▪ NNFI = .973 ▪ CFI = .987	Unidimensional	1	$\alpha$ .94 (Clínica) $\alpha$ .91 (Estudiantes) $\alpha$ .92 (Comunidad)
Depression and Family Functioning Scale [DFFS]								
Williams et al. (2016)	-	ACP	-	1	-	-	-	$\alpha$ 0.85 (Línea base) $\alpha$ 0.89 (2 meses) CCI 0.75
Escala de malestar psicológico de Kessler [K-10]								
Vargas et al. (2011)	▪ KMO = 0.916 ▪ Esfericidad de Barlett = < 0.000	ACP	-	1	-	-	-	$\alpha$ .90
Escala Autoaplicada de Depresión de Zung [ZSDS]								
Ruiz-Grosso et al. (2012)	-	-	-	-	-	-	-	$\alpha$ 0.89 $\omega$ 0.75
Campo et al. (2006)	▪ KMO = 0,835 ▪ Esfericidad de Barlett = $\chi^2 = 628,897$ ; gl = 45, p < 0,001	ACP	-	1	-	-	-	$\alpha$ .80
Campo-Arias et al. (2006)	-	-	-	-	-	-	-	$\alpha$ .832
Polaino y Senra (1991)	-	-	-	-	-	-	-	Entre evaluadores .80

*Propiedades psicométricas de los instrumentos: Reportes de análisis factorial y confiabilidad*

Autores	Análisis Factorial							Confiabilidad
	Análisis Factorial Exploratorio (AFE)			Análisis Factorial Confirmatorio (AFC)				
	Estadístico de adecuación de los datos	Método de extracción	Método de rotación	Factores (n)	Índice de Bondad de ajuste	Modelo corroborado	Factores (n)	
Escala Cognitivo-Conductual de Evitación [CBAS] Barajas et al. (2017)	-	-	-	-	<ul style="list-style-type: none"> <li>▪ CFI = .929</li> <li>▪ RMSEA = .057</li> <li>▪ SRMR = .051</li> <li>▪ <math>\chi^2(333) = 603.28, p &lt; .001, \chi^2/df = 1.81</math></li> </ul>	Multidimensional	4	$\alpha$ .95
Escala de Activación Conductual para la Depresión [BADS] Barraca et al. (2011)	-	-	-	-	<ul style="list-style-type: none"> <li>▪ CFI = .850</li> <li>▪ TLI = .833</li> <li>▪ RMSEA = .070</li> <li>▪ SRMR = .079</li> </ul>	Multidimensional	5	$\alpha$ .91
Depression and Somatic Symptoms Scale [DSSS] Tse et al. (2017)	<ul style="list-style-type: none"> <li>▪ KMO = 0.91</li> <li>▪ Esfericidad de Barlett = <math>\chi^2 [231] = 4637.72, p &lt; 0.001</math></li> </ul>	ACP	Promax	2	-	-	-	$\alpha$ 0.91
Dimensional Anhedonia Rating Scale [DARS] Rizvi et al. (2015)	-	ACP	Promax	4	-	-	-	$\alpha$ 0.92 (Estudio 1) $\alpha$ 0.92 (Estudio 2) $\alpha$ 0.96 (Estudio 3)
Inventario de Depresión de Beck [BDI] Lasa et al. (2000)	-	-	-	-	-	-	-	-
Bonicatto et al. (1998)	-	ACP	Varimax	2	-	-	-	$\alpha$ 0.87 Corrección Spearman-Brown 0.80
Polaino y Senra (1991)	-	-	-	-	-	-	-	Entre evaluadores .93
Inventario de Depresión de Beck/versión corregida [BDI-IA] Valdés et al. (2017)	-	ACP	-	1	<ul style="list-style-type: none"> <li>▪ <math>\chi^2 (189, n = 569) = 463.53, p = 0.00</math></li> <li>▪ RMSEA = 0,051</li> <li>▪ CFI = 0,979</li> </ul>	Unidimensional	1	$\alpha$ 0.92
Sanz y García-Vera (2007)	-	ACP	Promax	2	-	-	-	BDI-IA-SF = $\alpha$ 0.89 (Clínica) $\alpha$ 0.80 (Población general) $\alpha$ 0.79 (Estudiantes) BDI-IA-SCA = $\alpha$ 0.89 (Clínica) $\alpha$ 0.76 (Población general) $\alpha$ 0.79 (Estudiantes)
Inventario de Depresión de Beck- Segunda edición [BDI-II] González et al. (2015)	-	ACP	Oblimin directo	2	<ul style="list-style-type: none"> <li>▪ <math>\chi^2 = 415.41</math></li> <li>▪ CFI = .958</li> <li>▪ TLI = .952</li> <li>▪ RMSEA = .053</li> </ul>	Multidimensional	3	$\alpha$ .92 (Estudiantes) $\alpha$ .87 (Comunidad)

*Propiedades psicométricas de los instrumentos: Reportes de análisis factorial y confiabilidad*

Autores	Análisis Factorial							Confiabilidad
	Análisis Factorial Exploratorio (AFE)			Análisis Factorial Confirmatorio (AFC)				
	Estadístico de adecuación de los datos	Método de extracción	Método de rotación	Factores (n)	Índice de Bondad de ajuste	Modelo corroborado	Factores (n)	
González y Jenkins (2014)	-	-	-	-	-	-	-	$\alpha$ .91
Whisman et al. (2012)	-	-	-	-	<ul style="list-style-type: none"> <li>▪ S-B<math>\chi^2 = 2095.06, p &lt; .001</math></li> <li>▪ CFI = .923</li> <li>▪ SRMR = .033</li> <li>▪ RMSEA = .037, 90% CI = [.036, .039.]</li> </ul>	Multidimensional	3	$\alpha$ .91
Arnau et al. (2001)	-	ACP	Promax	2	-	-	-	$\alpha$ .94
Inventario de Depresión de Beck- Spanish [BDI-S]								
Bonilla et al. (2004)	-	ACP	Varimax	3	-	-	-	$\alpha$ .88
Inventario de Sintomatología Depresiva [IDS]								
González y Jenkins (2014)	-	-	-	-	-	-	-	$\alpha$ .86
Inventario de Sintomatología Depresiva- Clinician Rating [IDS-C]								
Trivedi et al. (2004)	-	-	-	1	<ul style="list-style-type: none"> <li>▪ GFI = 0.93</li> </ul>	Unidimensional	1	$\alpha$ 0.85 (QIDS-C) $\alpha$ 0.90 (IDS-C)
Inventario de Sintomatología Depresiva -Autoreporte [IDS-SR]								
Trivedi et al. (2004)	-	-	-	1	<ul style="list-style-type: none"> <li>▪ GFI = 0.93</li> </ul>	Unidimensional	1	$\alpha$ 0.86 (QIDS-SR) $\alpha$ 0.92 (IDS-SR)
Gili et al. (2011)	<ul style="list-style-type: none"> <li>▪ KMO = 0.96</li> <li>▪ Esfericidad de Barlett = 21669.61; <math>p &lt; 0.0001</math></li> </ul>	ACP	Promax	3	<ul style="list-style-type: none"> <li>▪ <math>\chi^2 = 2774.2</math></li> <li>▪ TLI = 0.95</li> <li>▪ CFI = 0.96</li> <li>▪ RMSEA = 0.067 (0.064-0.069)</li> <li>▪ AIC = 69926.9</li> </ul>	Multidimensional	3	$\alpha$ .94
Gili et al. (2014)	<ul style="list-style-type: none"> <li>▪ KMO = 0.894</li> <li>▪ Esfericidad de Barlett = <math>\chi^2 = 11691.39; p \leq 0.001</math></li> </ul>	-	-	1	-	-	-	$\alpha$ 0.871 Corrección Spearman-Brown 0.850 Estabilidad temporal > 0.70
Inventario Rápido de Sintomatología Depresiva -Autoreporte [QIDS-SR]								
Trujols et al. (2014)	<ul style="list-style-type: none"> <li>▪ KMO = 0.921</li> <li>▪ Esfericidad de Barlett = <math>\chi^2(36) = 708.67; p &lt; 0.001</math></li> </ul>	ACP	-	1	<ul style="list-style-type: none"> <li>▪ <math>\chi^2(27) = 30.345, p = 0.299</math></li> <li>▪ GFI = 0.962</li> <li>▪ RMSEA (90% CI) = 0.027 (0.000 - 0.067)</li> <li>▪ CFI = 0.995</li> </ul>	Unidimensional	1	$\alpha$ .881
El Inventario muy rápido de sintomatología depresiva - Autoinforme [VQIDS]								
De la Garza et al. (2007)	-	ACP	-	1	-	-	-	$\alpha$ 0.67 - 0.81
Inventario de Depresión Estado Rasgo [IDER]								
Dominguez-Lara (2016)	-	-	-	-	<ul style="list-style-type: none"> <li>▪ RMSEA = .075 (estado), .000 (rasgo)</li> <li>▪ CFI = .975 (estado), 1.000 (rasgo)</li> <li>▪ SRMR = .092 (estado), .052 (rasgo)</li> </ul>	Bidimensional	2 (estado) 2 (rasgo)	$\alpha > .80$ $\omega > .80$ Coeficiente H > .80

*Propiedades psicométricas de los instrumentos: Reportes de análisis factorial y confiabilidad*

Autores	Análisis Factorial							Confiabilidad
	Análisis Factorial Exploratorio (AFE)			Análisis Factorial Confirmatorio (AFC)				
	Estadístico de adecuación de los datos	Método de extracción	Método de rotación	Factores (n)	Índice de Bondad de ajuste	Modelo corroborado	Factores (n)	
Agudelo et al. (2014)	<ul style="list-style-type: none"> <li>▪ KMO = .848</li> <li>▪ Esfericidad de Barlett = <math>\chi^2 = 2871.893</math> <math>gl = 45</math>; <math>p &lt; .001</math></li> </ul>	ACP	Promax	2	<ul style="list-style-type: none"> <li>▪ <math>\chi^2 = 73.309</math>, <math>gl = 29</math>; <math>p &lt; .001</math></li> <li>▪ CFI = .981</li> <li>▪ TLI = .968</li> <li>▪ RMSEA = .042</li> <li>▪ SRMSR = .029</li> </ul>	Bidimensional	2	<ul style="list-style-type: none"> <li>α .84 (Estado)</li> <li>α .81 (Rasgo)</li> </ul>
Inventory for Depression and Anxiety Symptoms [IDAS] González y Jenkins (2014)	-	-	-	-	-	-	-	α .87
Cuestionario de Depresión Estado-Rasgo [ST/DEP] Agudelo et al. (2005)	-	FEP	Promax	2	-	-	-	α 0.90
Cuestionario de Síntomas de Ansiedad y del Humor [MASQ] González y Ibañez (2017)	-	ACP	Promax	3	<ul style="list-style-type: none"> <li>▪ CFI = &gt; 0.9</li> <li>▪ TLI = &gt; 0.9</li> <li>▪ RMSEA = &lt; 0.08</li> <li>▪ SRMR = &lt; 1</li> </ul>	Multidimensional	3	<ul style="list-style-type: none"> <li>α 0.88 - 0.92</li> <li>Test-retest 0.47 - 0.68</li> </ul>
Talkovsky y Norton (2015)	-	-	-	-	Modelo 1 <ul style="list-style-type: none"> <li>▪ <math>\chi^2 (6,843) = 10235.47</math></li> <li>▪ RMSEA = .046</li> <li>▪ CFI = .944</li> <li>▪ TLI = .942</li> </ul> Modelo 2 <ul style="list-style-type: none"> <li>▪ <math>\chi^2 (7,641) = 10072.08</math></li> <li>▪ RMSEA = .037</li> <li>▪ CFI = .960</li> <li>▪ TLI = .963</li> </ul>	Multidimensional	3	<ul style="list-style-type: none"> <li>α .958 (MASQ-GD)</li> <li>α .925 (MASQ-AD)</li> <li>α .916 (MASQ-AA)</li> </ul>
The Anxiety Depression Distress Inventory [ADDI] García et al. (2017)	-	-	Oblimin directo	3	Solución de un factor <ul style="list-style-type: none"> <li>▪ <math>\chi^2(324) = 5,495.67</math></li> <li>▪ CFI = .69</li> <li>▪ TLI = .67</li> <li>▪ SRMR = .20</li> <li>▪ RMSEA = .17</li> </ul> Solución de dos factores <ul style="list-style-type: none"> <li>▪ <math>\chi^2(298) = 2,377.86</math></li> <li>▪ CFI = .90</li> <li>▪ TLI = .88</li> <li>▪ SRMR = .08</li> <li>▪ RMSEA = .10</li> </ul> Solución de tres factores <ul style="list-style-type: none"> <li>▪ <math>\chi^2(273) = 1,176.45</math></li> <li>▪ CFI = .96</li> <li>▪ TLI = .94</li> <li>▪ SRMR = .04</li> </ul>	Multidimensional	3	<ul style="list-style-type: none"> <li>α .92 (Estudio 1)</li> <li>α .88 - .92 (Estudio 2)</li> <li>α .87 - .89 (Estudio 3)</li> </ul>

*Propiedades psicométricas de los instrumentos: Reportes de análisis factorial y confiabilidad*

Autores	Análisis Factorial							Confiabilidad
	Análisis Factorial Exploratorio (AFE)				Análisis Factorial Confirmatorio (AFC)			
	Estadístico de adecuación de los datos	Método de extracción	Método de rotación	Factores (n)	Índice de Bondad de ajuste	Modelo corroborado	Factores (n)	
					■ RMSEA = .07			
Cuestionario de Depresión para Hombres [CDH]								
Londoño et al. (2016)	-	-	-	-	-	-	-	α .96
Clinical Outcomes in Routine Evaluation [CORE]								
Soria et al. (2016)	-	-	-	-	-	-	-	α 0.93 CCI 0.96 (inter-rater) k = 0.35 - 0.82 CCI 0.94 (test-retest)

*Nota.* KMO = Kaiser-Meyer-Olkin; ACP = Análisis de Componentes Principales; α = Estadístico Alfa de Cronbach; CCI = Coeficiente de Correlación Intraclase; S-B  $\chi^2$  = Chi Cuadrado escalado entre los grados de libertad; CFI = Índice de Ajuste Comparado; TLI = Índice de Tucker-Lewis o NNFI = Índice de Ajuste No Normalizado; AIC = Criterio de información de Akaike; GFI = Índice de Bondad de Ajuste; RMSEA = Error Cuadrático Medio de Aproximación; SRMR = Standardized root mean square residual; ECVI = Índice de validación cruzada esperada; NFI = Índice de ajuste normalizado; PNFI = Índice de ajuste normalizado de parsimonia; ω = Coeficiente Omega; FEP = Factorización de ejes principales; k = kappa de Cohen.



Una vez el constructo se ha definido teóricamente y respaldado empíricamente, se espera que los estudios procedan con la evaluación de las evidencias de la validez externa, que involucra la validez de criterio, convergente y discriminante. Respecto a los resultados, se encontró que la correlación de Pearson fue el estadístico más implementado para la evaluación de la validez concurrente. La validez predictiva fue el análisis menos estudiado, solo los estudios de Senín-Calderón et al. (2017) y González y Ibañez (2017) la implementaron, usando análisis de regresión para su estimación.

Por su parte, la validez convergente fue evaluada por 30 investigaciones (Barajas et al., 2017; García et al., 2017; González et al., 2017; González y Ibañez, 2017; Porter et al., 2017; Senín-Calderón et al., 2017; Tse et al., 2017; Dominguez-Lara et al., 2016; Soria et al., 2016; Williams et al., 2016; González et al., 2015; Rizvi et al., 2015; Talkvosky y Norton, 2015; Agudelo et al., 2014; Bentley et al., 2014; Gili et al., 2014; González y Jenkins., 2014; Trujols et al., 2014; Antúnez y Vinet, 2012; Trujols et al., 2013; Agüera-Ortíz et al., 2013; Gili et al., 2011; De la Garza et al., 2007; Herrero y Gracia, 2007; Bonilla et al., 2004; Trivedi et al., 2004; Daza et al., 2002; Lobo et al., 2002; Arnau et al., 2001; Polaino y Senra, 1991), y la validez divergente (o discriminante) implementada por 15 investigaciones (Barajas et al., 2017; Porter et al., 2017; García et al., 2017; Camacho et al., 2016; Williams et al., 2016; Rizvi et al., 2015; Agudelo et al., 2014; Bentley et al., 2014; Trujols et al., 2014; Agüera-Ortíz et al., 2013; Antúnez y Vinet, 2012; Barraca et al., 2011; Bonilla et al., 2004; Daza et al., 2002; Lobo et a., 2002;). La tabla 8 contiene los datos encontrados en la revisión respecto a la validez de criterio (predictiva y concurrente), punto de corte, sensibilidad y especificidad, y las correlaciones de la validez convergente y discriminante. Los estudios que no dieron reporte alguno no fueron incluidos en la

tabla ( $n = 5$ ; Arbona et al., 2017; Ruíz et al., 2017; Whisman et al., 2012; Rivera-Medina et al., 2010; Agudelo et al., 2005).

Tabla 8

*Propiedades psicométricas de los instrumentos: validez externa*

Autores	Validez externa					Validez convergente	Validez discriminante
	Validez de criterio		Punto de corte	Sensibilidad (%)	Especificidad (%)		
	Validez concurrente	Validez predictiva					
Escala Multidimensional de los Trastornos Afectivos [EMTA]							
Abello-Luque et al. (2017)	▪ $r = \sigma = .816$ (BDI) $\sigma = .165$ (CPI)	-	-	-	-	-	-
Abello-Luque et al. (2013)	▪ $r = a = .715$ (BDI) $r = .025$ (ASMS)	-	-	-	-	-	-
Escala de Depresión de Montgomery-Asberg [MADRS]							
Cano et al. (2015)	-	-	20 moderada 34 grave	98 98	96 92	-	-
Lobo et al. (2002)	-	-	-	-	-	▪ Análisis de regresión multivariante = $p < 0,05$ y $0,01$	▪ Análisis de regresión multivariante = $p < 0,001$
Escala de Valoración de Hamilton para la Evaluación de la Depresión [HRSD]							
Porter et al. (2017)	-	-	-	0.01 - 1.00	0.17 - 1.00	$r = 0.49, p < 0.001$ (HARS y HRSD-17) $r = .50, p < 0.001$ (HARS y HRSD-24)	$r = 0.03,$ $\rho = 0.60$ (HRSD-R-II y HARS-R-II)
Polaino y Senra (1991)	-	-	-	-	-	$r = .50$ (BDI) $r = .47$ (SZDS)	-
Escala Leiden de Sensibilidad a la Depresión Revisada [LEIDS-R]							
Senín-Calderón et al. (2017)	-	Análisis de regresión lineal múltiple: ▪ Modelo 1= Constante 0,049 $\beta = 0,57$ (desesperanza) 0,23 (rumiación) -0,10 (agresión) -0,06 (aceptación) ▪ Modelo 2 = Constante 0,036 $\beta = 0,45$ (logro) 0,23 (autonomía) ▪ Modelo 3 = Constante 0,042 $\beta = 0,43$ (desesperanza) 0,26 (logro) 0,11 (autonomía) 0,17 (rumiación) -0,10 (agresión) -0,79 (aceptación)	48 puntos	75	70	$\rho = 0,49$ (BDI-II) $\rho = 0,48$ (DASS)	-
Clinically Usefull Depression Outcome Scale [CUDOS]							

*Propiedades psicométricas de los instrumentos: validez externa*

Autores	Validez externa					Validez convergente	Validez discriminante
	Validez de criterio		Punto de corte	Sensibilidad (%)	Especificidad (%)		
	Validez concurrente	Validez predictiva					
Trujols et al. (2013)	0.962 (área bajo la curva ROC)	-	≤ 20	91.09	91.67	$r = 0.77$ (HRSD) $r = 0.73$ (CGI-S) $r = 0.89$ (BDI)	-
Agüera-Ortiz et al. (2013)	-	-	-	-	-	$r = -0.42$ (SOFAS) $r = 0.45$ (HAM-D-17) $r = -0.22$ (PCS) $r = -0.65$ (MCS)	▪ Corrección de Bonferroni ( $p < 0.03$ ): No depresivo Depresión leve-moderada Depresión grave-muy grave
Escala de Depresión del Centro de Estudios Epidemiológicos [CES-D]							
González et al. (2017)	-	-	-	-	-	$r = .72$ (STAI) $r = -.65$ (SF-12 MHC) $r = .25$ (SF-12 PHC) $r = .80$ (PHQ-9)	-
González y Jenkins (2014)	-	-	-	-	-	.77 (BDI-II) .79 (IDAS- depresión general) .73 (IDS)	-
Salinas-Rodríguez et al. (2013)	$r_t = 0.9015$ (ICD-10) $r_t = 0.8717$ (DSM-IV) $r_t = 0.7879$ (BDI) Valor $p < 0.01$	-	≥ 9	90.2	86	-	-
Ruiz-Grosso et al. (2012)	-	-	≥ 29 (Muestra 1) ≥ 24 (Muestra 2)	77.1 91.4	79.4 96.7	-	-
Terrones-González et al. (2012)	$r = 0.40$ (Escala de Expectativas de Vivir-Morir)	-	-	-	-	-	-
Bojorquez y Salgado (2009)	$\rho = 0.94$ (CESD-20 y CESD-10)	-	-	79.8	97.3	-	-
Herrero y Gracia (2007)	-	-	-	-	-	Modelos de Regresión lineal: ▪ Modelo 1: $\beta = 0.56, p < .001$ (CESD-20) $\beta = 0.51, p < .001$ (CESD- 7) ▪ Modelo 2: $\beta = 0.12, p < .001$ (CESD-20) $\beta = 0.13, p < .001$ (CESD- 7) ▪ Modelo 3: $\beta = -0.09, p < .001$ (CESD-20) $\beta = -0.13, p < .001$ (CESD- 7)	-
Escalas de Depresión, Ansiedad y Estrés [DASS]							
Camacho et al. (2016)	-	-	-	-	-	-	MANOVA = Wilks' $\lambda = 0.99, F(3,500) = 2.41, p = 0.07$ ANOVA =

*Propiedades psicométricas de los instrumentos: validez externa*

Autores	Validez externa					Validez convergente	Validez discriminante
	Validez de criterio		Punto de corte	Sensibilidad (%)	Especificidad (%)		
	Validez concurrente	Validez predictiva					
Antúnez y Vinet (2012)	-	-	-	-	-	$r = .70$ (BDI) (Depresión) $r = .63$ (SCL-90) (Depresión) $r = .62$ (SCL-90-R) (Depresión)	$F(2, 499) = 8.55, p < 0.01$ (Depresión) $F(2, 499) = 9.25, p < 0.01$ (Ansiedad) $r = .57$ (BAI) (Depresión) $r = .47$ (SCL-ansiedad) (Depresión) $r = .55$ (SCL-Índice Malestar Positivo) (Depresión) $r = .56$ (SCL- Total Síntomas Positivos) (Depresión)
Osman et al. (2012)	▪ Análisis bivariado: $r = .80, p < .001$ (BDI-II) $r = .73, p < .001$ (MASQ) $r = .73, p < .001$ (Escala de estrés percibido)	-	-	-	-	-	-
Daza et al. (2002)						$r = .79, p < .01$ (DASS-21) $r = .86, p < .01$ (BDI)	$r = .60, p < .01$ (BAI) (Depresión)
Overall Depression Severity and Impairment Scale [ODSIS] Bentley et al. (2014)	-	-	$\geq 8$	74	85	Correlación .71 - .74 (Muestra 1) .62 - .76 (Muestra 2) -.63 - .88 (Muestra 3)	Correlación .25 - .55 (Muestra 1) -.11 - .59 (Muestra 2) .10 - .50 (Muestra 3)
Depression and Family Functioning Scale [DFFS] Williams et al. (2016)	-	-	-	-	-	$r$ línea base = 0.35 (SF-12) $r$ 2 meses = 0.51 (SF-12) $r$ línea base = .46 (SDS) $r$ 2 meses = 0.61 (SDS)	$r$ línea base = 0.05 (SF-12 PCS) $r$ 2 meses = 0.33 (SF-12 PCS) $r$ línea base = 0.28 (SDS) $r$ 2 meses = 0.45 (SDS)
Escala de malestar psicológico de Kessler [K-10] Vargas et al. (2011)	-	-	21	78.7	79.0	-	-
Escala Autoaplicada de Depresión de Zung [ZSDS] Ruiz-Grosso et al. (2012)	-	-	$\geq 47$ (Muestra 1) $\geq 45$ (Muestra 2)	85.7 91.4	71.4 91.8	-	-
Campo et al. (2006)	-	-	20	95.5	70.3	-	-
Campo-Arias et al. (2006)	-	-	40	88.6	74.8	-	-
Polaino y Senra (1991)	-	-	-	-	-	$r = .60$ (BDI) $r = .47$ (HRSD)	-
Escala Cognitivo-Conductual de Evitación [CBAS]							

*Propiedades psicométricas de los instrumentos: validez externa*

Autores	Validez externa					Validez convergente	Validez discriminante
	Validez de criterio		Punto de corte	Sensibilidad (%)	Especificidad (%)		
	Validez concurrente	Validez predictiva					
Barajas et al. (2017)	-	-	53	86	86	$\rho = .819 - .861$ (Muestra 1) $\rho = .658 - .787$ (Muestra 2) $\rho = .591 - .735$ (Muestra 3)	$\rho = -.470$ (MAAS-A) $\rho = -.416$ (MAAS-B)
Escala de Activación Conductual para la Depresión [BADS]							
Barraca et al. (2011)	$r = -.73$ (BDI-II)	-	-	-	-	-	$t(261) = 6.36; p < .001; d = .78$ (Muestra) $t(261) = -6.12; p < .001; d = .76$ (Evitación/Rumiación) $t(261) = -2.45; p < .05; d = .30$ (Deterioro laboral/escolar) $t(223.13) = -8.03; p < .001; d = 1.00$ (Deterioro social) MANOVA = $F(3, 237) = 10.33; p < .001; \eta^2 = .12$ (Evitación/Rumiación) $F(3, 237) = 20.88; p < .001; \eta^2 = .21$ (Deterioro social) $F(3, 237) = 1.55; p = .20; \eta^2 = .02$ (Activación) $F(3, 237) = 1.86; p = .14; \eta^2 = .02$ (Deterioro laboral/escolar) $F(3, 237) = 11.32; p < .001; \eta^2 = .12$ (Total)
Depression and Somatic Symptoms Scale [DSSS]							
Tse et al. (2017)	-	-	-	-	-	$r = 0.93$ (Sub. de depresión) $r = 0.84$ (Sub.somática) $r = 0.58$ (Sub. entre sí) $r = 0.78; p < 0.001$ (CESD y Sub.de depresión) $r = 0.49; p < 0.001$ (CESD y Sub. somática)	-
Dimensional Anhedonia Rating Scale [DARS]							
Rizvi et al. (2015)	-	-	-	-	-	$r_s = 0.58, p < 0.001$ (SHAPS) $r_s = 0.53$ (BAS-reward) $r_s = 0.35, p < 0.001$ (BAS-drive) $r_s = 0.18, p < 0.034$ (BAS-fun) $r_s = 0.35, p < 0.001$ (DARS-social y BAS-fun)	$r_s = -0.47, p < 0.001$ (CESD) $r_s = 0.47, p < 0.001$ (CESD y DARS-hobbies) $r_s = 0.47, p < 0.001$ (CESD y DARS-social) $r_s = 0.29, p < 0.001$ (CESD y DARS-comida/bebidas) $r_s = 0.31, p < 0.001$ (CESD y DARS-experiencia sensorial)
Inventario de Depresión de Beck [BDI]							
Valdés et al. (2017)	-	-	10 13.8 11.5	92.86 90.81 86.73	80 85.72 85.73	-	-

*Propiedades psicométricas de los instrumentos: validez externa*

Autores	Validez externa					Validez convergente	Validez discriminante
	Validez concurrente	Validez predictiva	Punto de corte	Sensibilidad (%)	Especificidad (%)		
González et al. (2015)	-	-	-	-	-	$r = .65, p < .001$ (HADS-escala de depresión) $r = .71, p < .001$ (HADS-escala de ansiedad)	-
González y Jenkins (2014)	-	-	-	-	-	.77 (CES-D)	-
Sanz y García-Vera (2007)	-	-	15 (BDI-IA-SCA) 14 (BDI-IA-SF)	-	-	.72 (IDAS-depresión general) .76 (IDS)	-
Bonilla et al. (2004)	-	-	-	-	-	Matriz de correlación= .62 - .69 (Depresión)	Matriz de correlación= .13 - .52 (Cognición, apoyo social y estrés)
Arnau et al. (2001)	-	-	18	94	92	$r = -.65$ (Sub.de Salud Mental) -.42 (Sub.de Percepción de Salud) -.24 (Sub.de Dolor Físico) -.19 (Sub. de Funcionalidad Física) -.35 (Sub. de Función del Rol)	-
Lasa et al. (2000)	-	-	12 / 13	100	99	-	-
Bonicatto et al. (1998)	$r = 0.68; p < 0.000$ (Sub depresión del SCL90)	-	-	-	-	-	-
Polaino y Senra (1991)	-	-	-	-	-	$r = .60$ (Escala de Zung) $r = .50$ (HRSD)	-
Inventario de Sintomatología Depresiva [IDS]							
Gili et al. (2014)	-	-	$\geq 6$	98	63	$r_{xy} = 0.845; p \leq 0.001$ (HRSD)	-
Trujols et al. (2014)	0.926 (área bajo la curva ROC)	-	$\geq 9$	80.37	89.09	$r = 0.775, p < 0.001$ (HRSD) $r = 0.766, p < 0.001$ (CGI)	$r = 0.781, p < 0.001$ (EQ-5D VAS) $r = 0.724, p < 0.001$ (EQ- Felicidad) $r = 0.382, p < 0.001$ (MSPSS)
González y Jenkins (2014)	-	-	-	-	-	.76 (BDI-II) .73 (CES-D)	-
Gili et al. (2011)	-	-	-	-	-	.69 (IDAS-depresión general) $r = 0.85$ (HRSD- primera medición) $r = 0.86$ (HRSD- segunda medición)	-
De la Garza et al. (2007)	-	-	-	-	-	$r = 0.90$ (QIDS-SR16) $r = 0.57$ (IDS-C30)	-
Trivedi et al. (2004)	-	-	-	-	-	$c = 0.83$ (QIDS-SR y IDS-SR- pacientes con TDM)	-

*Propiedades psicométricas de los instrumentos: validez externa*

Autores	Validez externa					Validez convergente	Validez discriminante
	Validez de criterio		Punto de corte	Sensibilidad (%)	Especificidad (%)		
	Validez concurrente	Validez predictiva					
						$c = 0.82$ (QIDS-C y IDS-C-pacientes con TDM) $c = 0.81$ (QIDS-C y IDS-C-pacientes con TB)	
Inventario de Depresión Estado Rasgo [IDER]							
Dominguez-Lara (2016)	-	-	-	-	-	$\phi = -.436$ (Estado) $\phi = -.657$ (Rasgo)	-
Agudelo et al. (2014)	-	-	-	-	-	.53 (CBD y Rasgo) .43 (CBD y Estado)	.38 (STAI y Rasgo) .36 (STAXI y Rasgo) .31 (STAI y Estado) .31 (STAXI y Estado)
Inventory for Depression and Anxiety Symptoms [IDAS]							
González y Jenkins (2014)	-	-	-	-	-	.72 (BDI-II) .79 (CES-D) .69 (IDS)	-
Cuestionario de Depresión Estado-Rasgo [ST/DEP]							
Cuestionario de Síntomas de Ansiedad y del Humor [MASQ]							
González y Ibañez (2017)	-	Análisis de regresión jerárquica: $\beta = -.13 - 0.49$ (BDI y AP, AN, AS) $\beta = 0.31 - 0.46$ (BAI y AS, AN)	-	-	-	$r = -.24$ (AS y AP) $r = 0.6$ (AS y AN) $r = -0.38$ (AP y AN)	-
Talkovsky y Norton (2015)	-	-	-	-	-	Correlación= Muestra 1= .454 - .666 (DASS) -.34 - .538 (PANAS) .518 - .782 (BAI) .615 - .749 (BDI) Muestra 2= .454 - .743 (DASS) -.141 - -.608 (PANAS) .657 - .783 (BAI) .620 - .783 (BDI) Muestra 3= .352 - .736 (DASS) -.059 - -.657 (PANAS) .353 - .811 (BAI) .566 - .756 (BDI) Muestra 4= .411 - .720 (DASS) -.026 - .551 (PANAS) .429 - .727 (BAI) .600 - .714 (BDI)	
The Anxiety Depression Distress Inventory [ADDI]							
García et al. (2017)	-	-	-	-	-	Correlación= -.48 [-.37, -.59] ADDI-GD y	Correlación= -.15 [-.01, -.30] (ADDI-PA y ADDI-SA)



*Propiedades psicométricas de los instrumentos: validez externa*

Autores	Validez externa					Validez convergente	Validez discriminante
	Validez de criterio		Punto de corte	Sensibilidad (%)	Especificidad (%)		
	Validez concurrente	Validez predictiva					
						ADDI-PA .46 [.33, .58] (ADDI-GD y ADDI-SA)	
Cuestionario de Depresión para Hombres [CDH] Londoño et al. (2016)	Correlación de Pearson: $r = .813$ ; $p \leq .001$ (MINI)		70 y 99 (leve) 100 y 119 (moderada) 120 y más (severa)	0.98	-	-	-
Clinical Outcomes in Routine Evaluation [CORE] Soria et al. (2016)	-	-	> 6	81.7	80	$\rho = 0.64$ ; $p < 0.001$ (Escala Newcastle)	-

*Nota.*  $r$  = Correlación de Pearson;  $\rho$  = Correlación de Spearman;  $\sigma$  = Desviación típica; CPI = Inventario Chino de Polaridad; ASMS = Escala Autoadministrada de Altman; HARS = Escala de Ansiedad de Hamilton;  $\beta$  = Error beta; ROC = Característica Operativa del Receptor; CGI-S = Escala de Impresión Clínica Global;  $t$  = Prueba  $t$  de Student;  $r_s$  = Correlación por rangos de Spearman;  $\phi$  = correlaciones interfactoriales; SOFAS = Social and Occupational Functioning Assessment Scale; HAMD-17 = Hamilton Depression Rating Scale; PCS = Physical Component Summaries; MCS = Mental Component Summaries; STAI = Cuestionario de Ansiedad Estado Rasgo; SF-12 MHC = The Mental Component of the short-form 12 health survey; PHQ-9 = Cuestionario sobre la Salud del Paciente-9;  $r_t$  = Correlación tetracórica; MANOVA = Análisis multivariante de la varianza; ANOVA = Análisis de la varianza; SCL = Symptom Checklist ; SCL-90 = Symptom Checklist-90; SCL-90 = Symptom Checklist-90-Revised; BAI = Inventario de Ansiedad de Beck; SDS = Sheehan Disability Scale; MAAS = Mindfulness Attention Awareness Scale-Adolescents; AAQ = Acceptance and Action Questionnaire; ATQ = Automatic Thought Questionnaire; MCQ-30 = Short form of the Metacognitions Questionnaire; EROS = Environmental Reward Observation Scale;  $r$  = Correlación producto-momento de Pearson; SHAPS = Snaith-Hamilton Pleasure Scale; EQ-5D = EuroQol; EQ-5D VAS = EuroQol-5D Visual Analogue Scale; MSPSS = Multidimensional Scale of Perceived Social Support; CBD = Cuestionario Básico de Depresión;  $\phi$  = Coeficiente Phi; STAXI = Inventario de Expresión de Ira Estado Rasgo; AP = Afecto Positivo; AN = Afecto Negativo; AS = Ansiedad Somática.

### Discusión

La depresión se ha convertido en un diagnóstico frecuente en la atención primaria de salud (Ministerio de Salud y Protección Social, 2017) y en la principal causa de discapacidad a nivel mundial (OMS, 2020). La población hispana representa un porcentaje significativo de esta prevalencia (Kohn, Levav, Caldas de Almeida, Vicente, Andrade, Caraveo-Anduaga, Saxena y Saraceno, 2005), incluso en un contexto estadounidense (Santiago-Rivera, Benson-Flórez, Santos y López, 2015).

La literatura evidencia que las normas y creencias culturales de una comunidad modelan la manera en que se experimenta y expresa la depresión (Chentsova-Dutton y Tsai, 2009). Un estudio reciente sobre la expresión lingüística de la depresión entre grupos culturales, encontró que los hispanos/latinos tienden a expresar gran cantidad de emociones positivas y negativas, y a evitar aspectos como la muerte, la autodestrucción, la religión o la salud (Loveys, Torrez, Fine, Moriarty y Coppersmith, 2018). Sin embargo, el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-IV TR; APA, 2000) sugiere que los hispanos son propensos a expresar síntomas somáticos como “nervios” o “dolores de cabeza” en lugar de síntomas afectivos. Por lo que resulta necesario mejorar la detección de la depresión en esta población para la adecuada selección de un tratamiento (Benuto, 2012). A menudo, uno de los principales desafíos que enfrenta un clínico es encontrar una medida adecuada para evaluar la depresión dado que es difícil establecer equivalencias por las diferencias culturales en el significado y expresión de los síntomas (Santiago-Rivera et al. 2015). Loveys et al. (2018) resaltan la importancia de crear instrumentos clínicos culturalmente adaptados, sensibles para medir la depresión, y que brinden evidencias de su calidad psicométrica.

El objetivo de este estudio fue analizar la evidencia científica disponible sobre la adecuación conceptual, los modos de aplicabilidad, y propiedades psicométricas de estudios originales sobre el desarrollo, uso, validación y adaptación de instrumentos de medición de sintomatología depresiva empleados en la práctica clínica e investigativa en adultos hispanohablantes entre 18 a 65 años. Un total de 53 estudios fueron incluidos en la revisión.

Respecto al objetivo específico del análisis de los indicadores bibliométricos, los resultados señalan un aumento de publicaciones de estudios instrumentales de depresión en los últimos años en población hispana (con mayor frecuencia en muestra española e hispano/latina en Estados Unidos) y con énfasis en el análisis de propiedades psicométricas y validación de las medidas. Lillo y Martini (2013) señalan que la variable de investigación más estudiada en Iberoamérica es la depresión (9%), así como una fuerte tendencia a la validación y uso de instrumentos psicométricos.

Se encontró que Estados Unidos (EE.UU.) es el país con mayor porcentaje de estudios publicados con población hispana, seguido de España y Colombia, otras revisiones hallaron resultados similares (Reilly et al., 2014; Reuland et al., 2009 y Limón et al., 2015). Dichas prevalencia de publicación puede explicarse por dos razones. Primero, el rápido crecimiento de la población hispana en EE.UU con alta prevalencia de depresión (Olfson, Shea, Feder, Fuentes, Nomura, Gameroff, Myma y Weissman, 2000) y segundo, por la influencia de la localización geográfica de las revistas de países de alto nivel de desarrollo científico (Cañedo, Hernández y Fresno, 2003).

La revista *Salud Mental* (de México) fue la más empleada para la publicación de los estudios. Pese a esto, se encontró una tendencia de publicación en revistas internacionales (EE.UU y Europa), teniendo en cuenta que en ellas los científicos adquieren mejores

puntuaciones y mayor visibilidad (Cano, 1995), resultados congruentes con otras investigaciones (Osca-Lluch, González-Salas, Fonseca y Civera, 2017; Martín, Molina y Martín, 2010; Alcain y Carpintero, 2001). Otras investigaciones hallaron resultados opuestos, en la medida que hay mayor tendencia a publicar en revistas locales (Vera-Villaruel, López-López, Lillo y Silva, 2010). Otros autores resaltan la importancia de apostarle a la publicación en las revistas locales para el desarrollo científico del país y la facilidad que brinda a la comunicación entre investigadores (Gómez y Bordons, 1996).

Por su parte, se halló que el factor de impacto de las revistas latinoamericanas es generalmente bajo (Téllez-Zenteno, Morales-Buenrostro y Estañol, 2007; Luna-Morales y Collazo-Reyes, 2006), debido a las formas de financiamiento, la calidad de la impresión, diseño y edición, los costos de distribución y difusión (Ugarte, 2004), y el escaso ajuste a los estándares internacionales (Licea de Arenas, Castaños y Arenas, 2002). Ahora bien, las perspectivas de desarrollo parecen favorecedoras. Vera-Villaruel et al. (2010) manifiestan que ha habido un incremento del factor de impacto de las revistas en Latinoamérica, sus ingresos a las bases de datos más prestigiosas demuestra un trabajo con alta calidad y buena gestión editorial. Por su parte, Sierra y Gómez (2019) señalan que existen otros aspectos que favorecen a la visibilidad de las revistas latinoamericanas como el uso del soporte digital, el rol del consejo editorial, el posicionamiento Web, inclusión de un segundo idioma, el acceso abierto, el prestigio de los autores y el apoyo de la editorial.

Respecto a las redes de colaboración, se encontró que 9 investigaciones tuvieron filiación extranjera, principalmente con EE.UU y países europeos (Ruíz et al., 2017; Williams et al., 2016; Soria et al., 2016; González et al., 2015; Talkvosky y Norton, 2015; De la Garza et al., 2007; Agudelo et al., 2005; Lasa et al., 2000; Bonicatto et al., 1998). Según muestran algunos trabajos,

los investigadores latinoamericanos no han logrado articular prácticas de colaboración en la investigación y publicación, lo que impacta significativamente en la visibilidad y el alcance de la producción científica (Humaní, González, Curioso y Pacheco-Romero, 2012; López, Silva, García-Cepero, Aguilar y Aguado, 2011). En Colombia, tres estudios publicaron en revistas internacionales con impacto significativo, según el número de citas (Ruíz et al., 2017; Cano et al., 2015 y Campo-Arias et al., 2006). Sin embargo, el estudio de Campo et al. (2006) evidencian un número elevado de citas. Frente a esto, Vera-Villaruel et al. (2010) encontraron que Colombia publica menos artículos que Brasil, México, Chile y Argentina, pero estos reciben tantas citas como los publicados en este grupo. La producción científica y el impacto de las publicaciones en los últimos años en Colombia ha ido en aumento, pero tal contribución se ha visto limitada por la tendencia de publicar en revistas de bajo impacto (Departamento Administrativo de Ciencia, Tecnología e Innovación [Colciencias], 2016).

En lo que respecta a la búsqueda y selección de las medidas, se identificaron 24 instrumentos. El Inventario de Depresión de Beck (BDI; Beck, Ward, Mendelson, Mock y Erbaugh, 1961) y la Escala de Depresión del Centro de Estudios Epidemiológicos (CES-D; Radloff, 1977) fueron las medidas más evaluadas entre los estudios, principalmente en población general. Otras revisiones hallaron que estas dos medidas son ampliamente evaluadas en hispanohablantes (Limón, Lamson, Hodgson, Bowler y Saeed, 2015; Reuland, Cherrington, Watkins, Bradford, Blanco, Gaynes, 2009) y en otras poblaciones (Vilagut, Forero, Barbaglia y Alonso, 2016; Mutumba, Tomlinson y Tsai, 2014).

El BDI es el instrumento de autoevaluación más conocido y más utilizado por su fácil administración y la exploración que brinda de los síntomas cognitivos relacionados con la depresión (Colasanti, Marianetti, Micacchi, Amadio y Mina, 2010). Sus resultados son

clínicamente útiles y se usan frecuentemente para valorar el progreso del paciente en la terapia (Benuto, 2012). Por su parte, el CES-D fue desarrollado originalmente para evaluar síntomas somáticos y afectivos de la depresión en población general (Santiago-Rivera et al., 2015). Santor (2006) menciona la amplia gama de puntajes que se pueden obtener del CES-D al incluir en su evaluación aspectos relacionados a las enfermedades y la salud en poblaciones no clínicas, pero también puede explicar la variabilidad de sus puntos de corte.

La mayoría de los estudios utilizaron muestra clínica para su investigación con mayor concentración de media de edades entre 40 a 65 años de España y países latinos. Belló, Puentes-Rosas, Medina-Mora y Lozano (2005) identificaron mayor prevalencia de síntomas depresivos a medida que se incrementa la edad. Otras investigaciones relacionan los factores personales y sociales con los síntomas, independientemente de la edad (Cerquera y Meléndez, 2010; Balanza, Morale y Guerrero, 2009) y otros atribuyen las experiencias de edades tempranas a la predisposición a la depresión en la adultez (Vázquez, 2016). Por lo cual, se concluye que existen diversos factores que influyen en la aparición de los síntomas depresivos, como la edad.

En lo que respecta a los modos de aplicabilidad, se encontró que la mayoría de los instrumentos son autoadministrados, diseñados para ser efectuados por el propio sujeto y se consideran como la más amplia categoría metódica en la evaluación psicológica (Fernández-Ballesteros, 2013). En la práctica clínica, el autoinforme puede ahorrar tiempo y costos en comparación con una entrevista clínica (Lako, Bruggeman, Knegtering, Wiersma, Schoevers, Slooff y Taxis, 2012). Sin embargo, la depresión presenta síntomas característicos en la capacidad de concentración, que en caso más extremos, puede invalidar los instrumentos autoadministrados (Ojeda, Silva, Cabrera, Tchimino, de la Carrera y Molina, 2003). Por su parte, las medidas heteroadministradas son valoradas por el examinador y requieren niveles de

capacitación profesional para llevarse a cabo (Salvador, Velázquez y González, 2000). En la evaluación de la depresión, la escala HDRS es considerada como la medida estándar de oro (Colasanti et al., 2010).

Se presenta un predominio de la aplicación tradicional de papel y lápiz, tan solo dos investigaciones que implementaron administraciones “en línea”, pese a que la revisión adelantada por van Ballegooijen, Riper, Cuijpers, van Oppen y Smit (2016) identificó una gran cantidad de instrumentos en línea para los síntomas depresivos, siendo la CES-D el instrumento mejor estudiado con adecuadas propiedades psicométricas para la población general (ej. Herrero y Meneses, 2006), seguido de la MADRS-S (ej. Holländare, Andersson y Engström<sup>1</sup>, 2010).

Fernández-Ballesteros (2013) argumenta que la técnica de lápiz y papel es utilizada con mayor frecuencia en los instrumentos por su sencillez y aplicabilidad. Mientras que Muñoz et al. (2015) aseguran que la tecnología ha impactado en la evaluación, pero, para que su uso sea satisfactorio se requiere establecer su fiabilidad y validez (Buchanan, 2002). Sin embargo, se presenta un incremento de dificultades respecto a la seguridad de las medidas debido a la proliferación de la evaluación informatizada (muchas veces sin vigilancia) (Muñoz et al., 2015).

Respecto al objetivo específico relacionado con la caracterización de la adecuación conceptual, se halló que la mayoría de los estudios evaluaron la validez de constructo mediante análisis factorial y la mayoría no evaluó validez de contenido. La revisión de Sanz, Izquierdoe y García-Vera (2013) hallaron deficiencias y limitaciones en algunos instrumentos en cuanto a la validez de contenido.

La validez de constructo es el grado en que un instrumento mide la variable objetivo y esta subsume todas las categorías de validez (Haynes et al., 1995; Messick, 1980). La validez de contenido refiere al grado en que los ítems son relevantes y representativos al constructo

(Mitchell, 1986; Anastasi, 1988). Kane (2009) la considera como una condición necesaria (aunque no suficiente) para las interpretaciones de las puntuaciones. Su estimación puede hacerse mediante el método de juicios de expertos y el uso de estadísticos (Sireci, 1998). Carretero-Dios y Pérez (2005) resaltan la importancia de la tabla de especificaciones para el establecimiento de los ítems y su respectiva evaluación mediante jueces (Escobar-Pérez, y Cuervo-Martínez, 2008). Frente a esto, surgen una serie de observaciones. Primero, De los Santos-Roig y Pérez-Meléndez (2013) argumentan que es difícil evidenciar la presentación de la definición semántica en un artículo debido a las limitaciones del formato, pero se puede hacer el intento de resumir el proceso de los ítems. Otro aspecto, es la adecuada selección de los jueces expertos, cuyos criterios pueden ser variados pero que influyen en el análisis de los ítems (Lawshe, 1975). Y por último, el análisis del juicio de expertos involucra elementos subjetivos, por lo tanto, para determinar su confiabilidad, se debe conocer el grado de acuerdo mediante análisis estadísticos (Pedrosa, Suárez-Álvarez y García-Cueto, 2014; Aiken, 2003).

La traducción de un test a un idioma distinto al original es una práctica común (Carretero-Dios y Pérez, 2005) con mayor frecuencia del idioma inglés a contextos de habla no inglesa (Hyrkäs, Appelqvist-Schmidlechner y Oksa, 2003). El español es uno de los idiomas a los que los instrumentos se adaptan con mayor frecuencia (Vallejo-Medina, Gómez-Lugo, Marchal-Bertrand y Saavedra-Roa, 2017). Sin embargo, Benuto (2012) argumenta que los instrumentos en español son limitados. En un contexto hispano/latino es importante tener en cuenta la heterogeneidad del idioma por los distintos dialectos de palabras, frases y expresiones que varían según el país (Benuto, 2012). Por esto, el proceso de traducción requiere de adaptaciones culturales, lingüísticas y métricas (Muñiz, Elousa y Hambleton, 2013). Sin embargo, múltiples autores argumentan que la traducción suele ser una elección rápida y arriesgada que puede llevar



a resultados inadecuados (Osterlind, 1989; Smith, Fischer y Fister, 2003) dado que la sola adaptación o traducción del instrumento no conduce a una adecuación conceptual ni a la equivalencia cultural debido a las diferencias culturales y del idioma (Escobar-Pérez y Cuervo-Martínez, 2008; Carretero-Dios y Pérez, 2005).

Los estudios revisados que realizaron el proceso de adaptación/traducción, emplearon el método de traducción inversa del habla inglesa a español y la mayoría no reportó aspectos relevantes del proceso. En línea con este hallazgo, la revisión de Limón et al., (2015) detectó un bajo número de instrumentos traducidos de detección de la depresión en la atención primaria en hispanohablantes y sólo un estudio reportó evidencias cegada en la administración, y el estudio de Reilly et al. (2014) encontraron que la mayoría de las traducciones de la escala QIDS no presentan evaluación psicométrica. La revisión de Arafat et al. (2016) concluyó que la mayoría de las investigaciones sobre adaptación transcultural siguen metodologías similares basadas en Beaton (2000) pero no hay un consenso en la literatura sobre cómo debe realizarse el proceso.

La mayoría de los estudios no registraron el modelo teórico en que sustentan sus medidas, los pocos que incluyeron este reporte mencionan que los ítems responden en los riterios del DSM el DSM, modelo tripartito o el modelo PANAS. La medición de la depresión representa desafíos conceptuales, clínicos (Botto, Acuña y Jiménez, 2014) y metodológicos (Ezquiaga, García, Díaz y García, 2011). De acuerdo con Ezquiaga et al. (2011) el concepto de depresión es confuso, ambiguo y equívoco y esto radica en gran parte por la controversia nosológica de los trastornos afectivos. Winokur (1997) propone que la depresión unipolar es un trastorno clínicamente homogéneo pero etiológicamente heterogéneo. No obstante, para Ezquiaga et al. (2011) los trastornos depresivos son un grupo grande heterogéneo de cuadros clínicos. Los manuales diagnósticos (DSM y CIE) clasifican los síntomas de la depresión mediante criterios

diferenciados por su gravedad, número de síntomas y los tiempos de duración (Vallejo y Leal 2005). Sin embargo, su utilidad ha sido altamente cuestionada (Alarcón, 2009). Moscoso (2014) argumenta que la depresión es un problema clínico, epidemiológico y de salud en general que amerita un marco conceptual más amplio, claro y coherente, es decir, debe ser comprendida desde una perspectiva dimensional (Botto et al., 2014). Por otra parte, Ezquiaga et al. (2011) atribuye las altas tasas de prevalencia de la depresión a problemas metodológicos en aspectos como la aplicación de instrumentos epidemiológicos de screening sin contrastar con otros instrumentos diagnósticos, el uso indebido de los términos depresión y/o síntomas depresivos y las diferencias de los criterios diagnósticos. Mientras que, Silverstone (1994) señala la necesidad de utilizar instrumentos adecuados, aplicar diagnósticos estandarizados y excluir el deterioro cognitivo.

En lo concerniente al objetivo específico del análisis de las propiedades psicométricas, se encontró que la mayoría de los estudios no proporcionan información suficiente sobre el proceso. Similar a esto, la revisión de Lako et al. (2012) de las medidas de síntomas depresivos en pacientes con esquizofrenia, descartaron ocho estudios por describir información incompleta sobre las propiedades psicométricas. Sin embargo, una revisión sobre la detección de síntomas depresivos en personas con discapacidades intelectuales, hallaron que todos los estudios que utilizaron nuevos instrumentos proporcionaron informes completos sobre las propiedades psicométricas (Pérez-Achiaga et al., 2009). De acuerdo con Argibay (2006), la verificación de las propiedades psicométricas como la confiabilidad y la validez son importantes porque permite obtener un instrumento que mide lo más exacto posible la variable objetivo. No obstante, Cardoso et al. (2010) aclara que las propiedades no aplican similarmente para todas las medidas

y por lo tanto, cada investigador debe determinar las propiedades adecuadas para su estudio acorde a la escala que esté evaluando.

Un primer aspecto para realizar en un análisis adecuado en la construcción/adaptación de una medida, es la aplicación del grupo de ítems en una muestra similar a la objetivo, cuyo propósito es detectar ítems problemáticos, dificultades en la comprensión de la instrucción y errores en el formato (Carretero-Dios y Pérez, 2005). Luján-Tangarife y Cardona-Arias (2015) añaden otros parámetros como el tiempo de aplicación, la claridad de las preguntas, requerimiento de entrenamiento del personal (o no) e identificación de si el registro, codificación, interpretación y evaluación es simple. Sin embargo, la mayoría de los estudios no realizaron este procedimiento.

Por otro lado, para el análisis de la dimensionalidad de los ítems, los estudios aplicaron el análisis factorial exploratorio y confirmatorio. La prueba de esfericidad de Bartlett y el KMO fueron ampliamente utilizados para la adecuación de los datos. El método de extracción y rotación más implementadas fueron el análisis de componentes principales y promax, respectivamente.

El análisis factorial es una de las técnicas más implementadas en el desarrollo y evaluación de las medidas psicológicas (Floyd y Widaman, 1995). Representa las correlaciones entre el grupo de ítems, dando lugar al establecimiento de variables no observables o factores (Abad et al., 2006). Argibay (2006) menciona que este análisis sirve para medir la validez factorial del instrumento, pero aclara que las estructuras factoriales tienden a ser inestables y depende en gran parte del tamaño de la muestra. Tabachnick y Fidell (2001) recomiendan aplicar el análisis en muestras grandes de aproximadamente 300 participantes. Sin embargo, Nunnally y Bernstein (1995) especifican que se debe contar con 10 participantes por variable y mínimo 5 por

ítems. Pero cuando se trata de muestras muy grandes es recomendable diferenciarse por cada sexo (Kline, 2000). Por su parte, Floyd y Widaman (1995) mencionan que pese a la gran popularidad de estos análisis, la complejidad y flexibilidad de los procesos analíticos factoriales han mostrado ambigüedad en la práctica investigativa.

El análisis factorial presenta dos modalidades diferentes: Análisis Factorial Exploratorio (AFE) y el Análisis Factorial Confirmatorio (AFC) cuyas diferencias son de tipo teórica y matemática. El AFE es una técnica multivariada que se usa para explorar un conjunto de factores subyacentes a un grupo amplio de datos (Pérez y Medrano, 2010). El número de factores es menor al número de variables y constituyen un significado teórico (Kline, 2000). Lloret-Segura, Ferreres-Traver, Hernández-Baeza y Tomás-Marco (2014) mencionan que existen dos tendencias de publicación científica muy marcadas con este tipo de análisis y que ha generado discrepancia respecto a su uso. Primero, los metodólogos buscan mejores criterios para la aplicabilidad del AFE mediante el estudio y comparación de los criterios comúnmente utilizados con otros nuevos o no tan nuevos pero menos populares, y segundo (el más utilizado), la identificación de la estructura subyacente a los ítems en donde los investigadores siguen una línea de los procesos clásicos, pero, desconocen lo inadecuados y peligrosos que pueden ser (Lloret-Segura, et al., 2014). Pérez y Medrano (2010) señalan que el AFE supone una definición débil del constructo debido a que su medición es meramente empírica al no exigir un modelo teórico previo. Algunos autores señalan que el AFE debe complementarse con un AFC (Pérez y Medrano, 2010; Bryne, 2001).

Por otra parte, el AFC resulta adecuado cuando los investigadores tienen conocimiento de la estructura teórica subyacente y permite definir el número de factores que se espera, las relaciones entre sí y los ítems relacionados a cada factor (Lloret-Segura et al., 2014). Sin

embargo, pese a la importancia que tiene la teoría previa en el análisis factorial, Ferrando y Anguiano-Carrasco (2010) argumentan que esta no es suficiente para especificar un modelo confirmatorio, por esto, proponen implementar un modelo no restringido (exploratorio) pero con finalidad confirmatoria para la estimación del número de factores y si son o no independientes. Pero si el estudio está lo suficientemente avanzado como para plantear una solución restringida o si el grupo de ítems son excepcionalmente simples, el AFC sería un modelo viable de usar (Ferrando y Anguiano-Carrasco, 2010). No obstante, Pérez y Medrano (2010) mencionan que el AFC se encuentra inmerso en una teoría que guía la definición del constructo. Floyd y Widaman (1995) mencionan que este análisis se usa para confirmar una hipótesis a priori que pueden estar basada en la teoría o a partir de los resultados de estudios empíricos anteriores. Por otra parte, Batista-Foguet et al. (2004) mencionan que actualmente existe una gran variedad de programas accesibles y sencillos para estimar los modelos del AFC. Sin embargo, Pérez Gil et al. (2000) señalan que pese a que estos procedimientos están muy desarrollados, aún se continúa implementando el AFE con fines confirmatorios.

Por su parte, en la construcción/adaptación de un instrumento un aspecto esencial es determinar la confiabilidad del test. Refiere a la precisión con la que el instrumento mide lo que pretende medir en una población específica y en un contexto normal de aplicación (Anastasi, 1986; Aiken, 2003). La falta de confiabilidad radica en la relación con la intervención del error, que pueden ser de dos tipos: error constante (sistemático) y el error causal, al azar o no sistemático, (Aliaga, 2007). Cardoso et al (2010) menciona que existen una serie de factores que pueden afectar la precisión de un test y por ende la fiabilidad, como son: mayor variabilidad de la muestra (Muñiz, 1994), mayor número de ítems (Cortada, 1999; Muñiz, 1994), mayor tiempo de responder, menor amplitud de la dificultad de los ítems, mayor homogeneidad del contenido del

instrumento, mayor experiencia con el formato, mayor motivación y menos preguntas de contenido emocional (Cortada, 1999). Existen diversos estadísticos para estimar empíricamente la fiabilidad: el test-retest, formas paralelas y la consistencia interna (Prieto y Delgado, 2010). Esta última fue la más implementada entre los estudios mediante el estadístico Alfa de Cronbach ( $\alpha$ ; Cronbach, 1951).

Diversas revisiones reportan el alfa de Cronbach como el estadístico más utilizado para la estimación de la consistencia interna (Arafat et al., 2016; da Silva, Goncalves, Arancibia, Bento, Castro, Hernandez y da Silva, 2015; Mutumba et al., 2014; Reilly et al., 2014). Algunos autores consideran la consistencia interna aceptable cuando se encuentra entre 0,70 y 0,90 (Terblanche y Boshoff, 2006; Cervantes, 2005; Sánchez y Gómez, 1998; Bland y Altman, 1997), mientras que, para otros es adecuada si se encuentra entre 0,80 y 0,90 (Roberts y Priest, 2006; Streiner, 2003; Clark y Watson, 1995, Cortina, 1993). Otra revisión halló mayor reporte de la fiabilidad entre evaluadores (Pérez-Achiaga et al., 2009)

Algunos autores mencionan que pese a la popularidad del coeficiente alfa de Cronbach, este puede verse afectado por una serie de aspectos como el número de ítems, el número de alternativa de respuestas y la proporción de la varianza del test (Domínguez-Lara y Merino-Soto, 2015). Ante esto, McDonald (1999) propone el uso del Coeficiente Omega ( $\omega$ ) como alternativa para la estimación de la confiabilidad. Este coeficiente trabaja a partir de las cargas factoriales (Gerbing y Anderson, 1988) que son la suma ponderada de las variables estandarizadas permitiendo la estabilidad de los cálculos y refleja el verdadero nivel de la confiabilidad, no depende del número de ítems y es adecuada sino se cumple el principio de equivalencia (McDonald, 1999). Pese a esto, su poco uso se debe a que no se encuentra disponible en los

paquetes estadísticos como el Statistical Package for the Social Sciences (SPSS) (Ventura-León y Caycho-Rodríguez, 2017).

Por otra parte, la validez externa (criterio, convergente y discriminante) fue el procedimiento menos estudiado entre los estudios. La correlación de Pearson fue el estadístico más recurrente. Una revisión encontró un uso frecuente de la correlación de Pearson para este tipo de análisis (Reilly et al., 2014). Pero otra revisión encontró que los estudios instrumentales se basaron en entrevistas diagnósticas para el establecimiento del criterio de diagnóstico de referencia (Matumba et al., 2014).

La validez de criterio (concurrente y predictiva) se usa para hacer inferencias de las puntuaciones del test con otras variables (Argibay, 2006) o prueba de referencia o criterio (*gold standard*; Israni, 2010) que pretende medir lo mismo. Por lo tanto, entre más se relacionen los resultados del test con los del criterio, mayor será la validez (Hernández et al., 2015). No obstante, Arafat et al. (2016) argumentan que en el caso de las validaciones, muchas culturas carecen de estándares de oro validados, lo que genera una situación problemática para los investigadores.

Por su parte, la validez convergente y discriminante son propiedades que correlacionan los puntajes a través de escalas diferentes (Luján y Cardona, 2015). Se evalúan mediante las matrices multirrasgos-multimétodo (Abad et al., 2006). Para la validez convergente, Suárez (2014) argumenta que los procedimientos estadísticos dependerán del número de predictores y criterios. Para el caso de una única escala y un solo criterio, se aplica los procedimientos de correlación o regresión lineal simple y para varios predictores y un solo criterio se utiliza la regresión lineal multivariante y correlación canónica (Martínez, 1995). Para el caso de la validez discriminante, Martínez y Martínez (2009) mencionan que se debe tener precaución al usar un

criterio, por esto, presentan otros métodos como comparación entre las correlaciones de los indicadores, comparación entre la varianza compartida y la varianza extraída, intervalo de confianza entre las correlaciones, correlaciones entre el método común y diferencia entre valores medios.



### **Conclusiones y recomendaciones**

A partir de los hallazgos de la presente revisión sistemática se concluye que la mayoría de los estudios no proporcionaron suficiente información acerca del proceso de construcción/adaptación de los instrumentos. Una revisión determinó que la limitación metodológica más común y de gran importancia es la falta de documentación (Reuland et al., 2009). La revisión de Clark y Watson (2003) señala la falta de coherencia entre los artículos publicados en aspectos como el formato de presentación, el rigor y la disparidad de los datos. La Comisión Internacional de Test refiere que los investigadores que decidan construir un nuevo instrumento de medida, deben proporcionar información completa sobre sus características y propiedades psicométricas (Hambleton, Merenda y Spielberger, 2005; Muñiz, Elosua y Hambleton, 2013). Otras organizaciones resaltan que la creación y adaptación es una cuestión de reflexión y que esta debe seguir las recomendaciones generales sobre el proceso (AERA, APA y NCME, 1999). Carretero-Dios y Pérez (2007) argumentan que la mayoría de los estudios publicados carecen de información relevante que permitan enjuiciar su calidad. Buela-Casal, Carretero-Dios y De los Santos-Roig (2002) señalan la falta de similitud y de contenido en los artículos con objetivos y metodologías similares, lo que dificulta, por ejemplo, llevar a cabo meta-análisis sobre análisis de metodologías o la comunicación científica.

#### **Limitaciones de la revisión**

Si bien esta es, a conocimiento de la autora, la primera revisión sistemática instrumental en depresión en adultos hispanohablantes enfocada en la adecuación conceptual, aplicabilidad y propiedades psicométricas, reconoce en ella algunas limitaciones que deben considerarse. Primero, no se contó con el equipo suficiente para llevar a cabo el proceso de preselección, selección y extracción de datos de los estudios incluidos en la revisión, de manera más eficiente,

teniendo en cuenta, el número de códigos de búsqueda que surgieron a partir de la combinación de los términos. Segundo, es posible que no se hayan incluido estudios potenciales en el proceso de revisión porque no fue posible acceder a los documentos completos, aun cuando se contactó con los autores originales. Tercero, el nivel de análisis se realizó desde una perspectiva meramente metodológica, dada la limitada experticia en psicometría inherente al nivel formativo de la autora de esta revisión. Seguramente, un análisis más profundo permitiría realizar inferencias estadísticas y con ello la obtención de más evidencias claves. Pese a las limitaciones manifiestas, los resultados hallados dan cuenta que la mayoría de las investigaciones no reportaron aspectos importantes del proceso de creación/adaptación de los instrumentos de medida, lo que pone en juicio la calidad de los mismos.

### **Recomendaciones finales**

El diagnóstico de la depresión constituye uno de los trastornos más frecuentes a nivel mundial y se estima su aumento en los próximos años. Su detección oportuna permite brindar un tratamiento óptimo y de calidad. Sin embargo, parte del proceso de evaluación lo constituyen los instrumentos de medidas válidos y confiables que el clínico debe implementar con precaución y de forma adecuada.

A continuación se relaciona los instrumentos de medida hallados en la presente revisión, y su respectivo uso recurrente con el objetivo de facilitar a los clínicos e investigadores el análisis y adecuada selección de las medidas y sus oportunas acciones profesionales:

- El BDI y la escala LEIDS son medidas de autoevaluación que identifican aspectos cognitivos de la depresión. El primero es un inventario ampliamente implementado que estima la gravedad y funciona para monitorear los síntomas de la depresión durante la terapia. La

LEIDS mide la reactividad cognitiva sin utilizar una tarea de inducción del estado de ánimo negativo.

- Las escalas MADRS, HRSD y DASS suele emplearse en pacientes diagnosticados con depresión. Las dos primeras son heteroaplicadas. La MADRS mide síntomas depresivos y los efectos del tratamiento antidepresivo; la HRSD mide intensidad en pacientes y la DASS evalúa los componentes dimensionales de la ansiedad y los trastornos depresivos.

- Los instrumentos EMTA, ODSIS, DSSS, CUDOS, IDER, ST/DEP, el IDAS, el y el IDS son medidas diseñadas para medir los síntomas de la depresión. La EMTA evalúa los síntomas de los trastornos del estado de ánimo en aspectos cognitivos, conductual, afectivo y fisiológicos. La ODSIS evalúa severidad y deterioro disfuncional de la depresión. La DSSS es un autoinforme que mide los síntomas somáticos de la depresión en muestras clínicas. La escala CUDOS es un instrumento de detección de casos y una medida sensible al cambio de la severidad de los síntomas depresivos. El IDER y el ST/DEP evalúan el componente afectivo de la depresión. El IDAS es un instrumento de 64 ítems que se centra en medir síntomas cognitivos y emocionales. Y por último, el IDS es un inventario diseñado para medir los síntomas de depresión y subtipos del DSM, consta de una versión heteroaplicada y autoaplicada.

- El CORE y el DARS se centran en medir la anhedonia.

- El DFFS evalúa el impacto de la depresión en el funcionamiento familiar desde la perspectiva del paciente y la pareja.

- La K-10 y el ZSDS son instrumentos que permiten identificar síntomas depresivos en atención primaria.

- El CBAS y el BADS son medidas que evalúan la evitación. La primera es una medida multidimensional que evalúa la dimensión de la evitación cognitiva, conductual,

social/no social y activa/pasiva y la segunda medida mide los cambios en la evitación y activación sobre el curso de la terapia de Activación Conductual.

- La escala CES-D es una escala autoaplicada diseñada para medir síntomas depresivos en población general. Es implementada en estudios epidemiológicos.
- La MASD y el ADDI están diseñados para medir las dimensiones del modelo tripartido de la ansiedad y la depresión.
- El CDH mide síntomas característicos de la depresión en hombres.

Para las futuras investigaciones, se ofrecen una serie de recomendaciones de tipo metodológico y procedimental. Primero, se sugiere contar con listas de verificaciones y/o manuales centradas en variables psicológicas. Si bien, en la actualidad existen referencias altamente respaldadas, éstas se enfocan en variables alejadas del estudio psicológico. Segundo, teniendo en cuenta los cambios tecnológicos en la actualidad, se recomienda implementar más análisis de test en línea, cuyas herramientas pueden ser útiles en un contexto terapéutico como investigativo, pero siguiendo todas las precauciones necesarias para su adaptación y uso. Tercero, por sugerencias de autores, es importante apostarles a las publicaciones en contextos nacionales contribuye a los procesos científicos del país y facilita la comunicación entre investigadores. Y por último, se recomienda a los investigadores seguir los estándares de calidad sobre la adaptación/construcción de un instrumento de medida y dar reporte de los mismos. Esto permite ofrecer una medida válida, confiable y adecuada para su uso clínico e investigativo.

### Referencias

- Abad, F. J., Olea, J., Ponsoda, V. y García, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid: Síntesis.
- Abad, F.J., Garrido, J., Olea, J. y Ponsoda, V. (2006). *Introducción a la psicometría. Teoría Clásica de los Test y Teoría de la Respuesta al ítem*. Madrid: Universidad Autónoma de Madrid.
- Abadal, E. y Rius, L. (2006). Revistas científicas digitales: características e indicadores. *Revista de Universidad y Sociedad del Conocimiento*, 6-20.
- Abello-Luque, D. y Cortés Peña, O.F. (2012). Diseño de la Escala Multidimensional de Trastornos Afectivos (EMTA): Análisis psicométrico de confiabilidad y validez. *Psicología desde el Caribe*, 29(3), 546- 591
- Abramson, L. Y.; Metalsky, G. I., y Alloy, L. Y. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, 96, 358-372.
- Abramson, L. Y.; Seligman, M. E. P., y Teasdale, J. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology*, 87, 49-74.
- AERA, APA y NCME, (1999). *Standards for educational and psychological tests*. Washington DC: American Psychological Association, American Educational Research Association, National Council on Measurement in Education.
- Aiken, L. (2003). *Test psicológicos y evaluación*. Mexico: Pearson Education.
- Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement* 40, 955–959
- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52(3), 317-332.
- Alarcón R. (2009). Culture, cultural factors and psychiatric diagnosis: review and projections. *World Psychiatry*, 8, 131-139.
- Alcain Partearroyo, M.D. y Carpintero, H. (2001). La psicología en España a través de las revistas internacionales: 1981-1999. *Papeles del Psicólogo* (78), 11-20

- Aliaga Tovar, J. (2007). Psicometría: Test psicométricos, confiabilidad y validez. Recuperado de [https://s3.amazonaws.com/academia.edu.documents/38260625/1U2LibroEAPAliaga.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1526678696&Signature=52Uw%2B0YIPzH3KuoIR%2BmXNaaKXxQ%3D&response-content-disposition=inline%3B%20filename%3DPsicometria\\_Tests\\_Psicometricos\\_Confiabi.pdf](https://s3.amazonaws.com/academia.edu.documents/38260625/1U2LibroEAPAliaga.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1526678696&Signature=52Uw%2B0YIPzH3KuoIR%2BmXNaaKXxQ%3D&response-content-disposition=inline%3B%20filename%3DPsicometria_Tests_Psicometricos_Confiabi.pdf)
- Alvarado Izquierdo, J.M. y Santisteban Requena, C. (2011). *La validez en la medición psicológica*. Madrid, España: UNED Ediciones.
- Álvarez, E. (2016). Bases neuroquímicas de la depresión. Un modelo médico de la enfermedad depresiva. *Psiquiatría biológica*, 23(1), 9-15. Doi: 10.1016/S1134-5934(17)30048-9
- Álvarez, N., y Londoño, C. (2012). Diseño y validación del Cuestionario de Depresión en Hombres. *Tesis de Maestría*. Universidad Católica de Colombia.
- Amador Muñoz, L., Monreal Gimeno, M.C., y Marco Macarro, M.J. (2001). El adulto: etapas y consideraciones para el aprendizaje. *Eúphoros*, 3, 97-112
- American Psychiatric Association [APA] (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington, DC: Author.
- American Psychiatric Association [APA] (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Association: Washington, DC.
- American Psychiatric Association [APA] (2000). *Diagnostic and statistical manual of mental disorders, DSM-IV*. 4th edition, 4th text rev. [DSM-IV-TR] Washington, DC: American Psychiatric Association.
- American Psychiatric Association [APA] (2014). *Manual Diagnóstico y Estadístico de los Trastornos Mentales [DSM-5]*. España: Editorial Médica Panamericana.
- American Psychological Association [APA] (2005). Report of the 2005 Presidential Task Force on Evidence-Based Practice. Recuperado de <https://www.apa.org/practice/resources/evidence/evidence-based-report.pdf>
- Anastasi, A. (1988). *Psychological testing* (6.ª ed.). New York: Macmillan

- Andresen, E.M., Malmgren, J.A., Carter, W.B., Patrick, D.L. (1994). Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). *American Journal of Preventive Medicine*, 10, 77-84.
- Arafat, S. Y., Chowdhury, H. R., Qusar, M. S., y Hafez, M. A. (2016). Cross cultural adaptation y psychometric validation of research instruments: A methodological review. *Journal of Behavioral Health*, 5(3), 129-136.
- Aragón Borja, L.E. (2004). Fundamentos psicométricos en la Evaluación Psicológica. *Revista Electrónica de Psicología Iztacala*, 7(4), 23-43
- Aragón Borja, M.E. (2015). Evaluación Psicológica: Historia, fundamentos teórico-conceptuales y psicometría. México, D.C: Editorial El Manual Moderno
- Argibay, J.C. (2006). Técnicas psicométricas. Cuestiones de validez y confiabilidad. *Subjetividad y procesos cognitivos*, 15-33
- Argimón Pallas, J.M. y Jiménez Villa, J. (2013). *Métodos de investigación clínica y epidemiológica: Cuarta Edición*. Madrid: Elsevier España, S.L
- Arnett, J. J. (2000) Emerging adulthood: a theory of development from the late teens through the twenties. *American Psychologist*, 55(5), 469-480
- Attoresi, H.F., Lozzia, G.S., Abal, F.J.P., Galibert, M.S. y Aguerri, M.E. (2009). Teoría de Respuesta al Ítem. Conceptos básicos y aplicaciones para la medición de constructos psicológicos. *Revista Argentina de Clínica Psicológica*, 10(2), 179-188
- Balanza Galindo, S., Morales Moreno, I., y Guerrero Muñoz, J. (2009). Prevalencia de ansiedad y depresión en una población de estudiantes universitarios: factores académicos y sociofamiliares asociados. *Clínica y salud*, 20(2), 177-187.
- Batista-Foguet, J. M., Coenders, G., & Alonso, J. (2004). Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud. *Medicina clínica*, 122(1), 21-27.
- Beato, M.S. (1993). Explicaciones teóricas de la depresión: del conductismo al cognitivismo. *Revista de Historia de la Psicología*, 14(3-4), 385-394

- Beaton, D. E., Bombardier, C., Guillemin, F., y Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186-3191.
- Beck, A. T. (1996). Beyond belief: A theory of modes, personality, and psychopathology. In P. M. Salkovskis (Ed.), *Frontiers of cognitive therapy* (pp. 1-25). New York: The Guilford Press
- Beck, A. T. y Beck, R. W. (1972). Screening depressed patients in family practice. A rapid technic. *Postgraduate Medicine*, 52 (6), 81-85.
- Beck, A. T., Rush, A. J., Shaw, B. F., y Emery, G. (1979). *Cognitive therapy of depression*. New York, NY: Guilford Press.
- Beck, A. T., Steer, R. A., y Brown, G. K. (1996). *Beck Depression Inventory-II manual*. San Antonio, TX: Psychological Corporation
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., y Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.  
Doi:10.1001/archpsyc.1961 .01710120031004
- Beck, A. T., y Clark, D. A. (1988). Anxiety and depression: An information processing perspective. *Anxiety Research*, 1, 23-36.
- Beck, A. T., y Steer, R. A. (1987). *Manual for the Beck Depression Inventory*. San Antonio, TX: Psychological Corporation
- Beck, A.T. (1963). Thinking and depression: 1, idiosyncratic content and cognitive distortions. *Archives of General Psychiatric*, 9, 324-333.
- Beck, A.T. (1964). Thinking and depression: 2, theory and therapy. *Archives of General Psychiatric*, 10, 561-571.
- Beck, A.T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.
- Beck, A.t. (1987). Cognitive models of depression. *Journal of Cognitive Psychotherapy: An International Quarterly*, 1, 5-37.



- Beck, D.A y Koenig, H.G. (1996). Minor depression: a review of the literature. *International Journal of Psychiatry Medical*, 26(2), 177-209.
- Belló, M., Puentes-Rosas, E., Medina-Mora, M. E., y Lozano, R. (2005). Prevalencia y diagnóstico de depresión en población adulta en México. *Salud pública de México*, 47(1), S4-S11.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246
- Bentley, K.H. y Gallagher, M.W. (2014). Development and Validation of the Overall Depression Severity and Impairment Scale. *Psychological Assessment*, 26(3), 815-830
- Benuto, L. T. (Ed.). (2012). *Guide to psychological assessment with Hispanics*. New York: Springer Science y Business Media.
- Birnbaum, A. (1968) Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. En: Lord, F.M. and Novick, M.R., Eds., *Statistical Theories of Mental Test Scores*, Addison-Wesley, Reading, 397-479.
- Bland, J. M., y Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *Bmj*, 314(7080)
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley y Sons
- Bordons, M. y Zulueta, M.A. (1999). Evaluación de la actividad científica a través de indicadores bibliométricos. *Revista española de cardiología* 52, 790-800
- Botto, A., Acuña, J., y Jiménez, J. P. (2014). La depresión como un diagnóstico complejo: Implicancias para el desarrollo de recomendaciones clínicas. *Revista médica de Chile*, 142(10), 1297-1305.
- Bower, G. H. (1981). Mood and memory. *American Psychologist*, 36(2), 129-148. Doi: 10.1037//0003-066X.36.2.129
- Brioso Díez, A., Contreras Felipe, A., Corran Íñigo, A., Delgado Egido, B., Díaz Mardomingo, M.C., Giménez Dasí, M., Gómez Veiga, I., Oliva Delgado, A. y Sánchez Queija, I.

- (2009). *Psicología del desarrollo, desde la infancia a la vejez*. Madrid, España: McGraw-Hill/Interamericana de España.
- Brown, G. F. (1980). *Principios de la medición en psicología y educación*. México: El Manual Moderno.
- Buchanan, T. (2002). Online assessment: Desirable or dangerous?. *Professional psychology: Research and practice*, 33(2), 148.
- Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum
- Campagne, D.M. (2012). Causas orgánicas y comórbidas de la depresión: el primer paso. *Semergen*, 38(5), 301-311. Doi: 10.1016/j.semereg.2011.12.010
- Cano, V. (1995). Characteristics of the publishing infrastructure of peripheral countries : a comparison of periodical publications from Latin America with periodicals from the US and UK. *Scientometrics*, 34, 121-138.
- Cañedo, R. A., Hernández, A. S. J. y Fresno, C. C. H. (2003). Iberoamérica a 500 años del descubrimiento: la producción científica de una región en ciencias biomédicas en la década de los ochenta. *ACIMED*, 11(1).
- Cardoso Ribeiro, C., Gómez-Conesa, A. y Hidalgo Montesinos, M.D. (2010). Metodología para la adaptación de instrumentos de evaluación. *Fisioterapia*, 32(6), 264-270
- Carretero-Dios, H. y Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales. *International Journal of Clinical and Health Psychology*, 5(3), 521-551
- Carretero-Dios, H. y Pérez, C. (2007). Normas para el desarrollo y revisión de estudios instrumentales: consideraciones sobre la selección de tests en la investigación psicológica. *International Journal of Clinical and Health Psychology*, 7(3), 863-882
- Casal, G., Carretero-Dios, H. y De los Santos-Roig, M. (2002). Estudio comparativo de las revistas de Psicología en castellano con factor de impacto. *Psicothema*, 14, 837-852.

- Cerquera Córdoba, A.M. y Meléndez Merchán, C. (2010). Factores culturales asociados a las características diferenciales de la depresión a través del envejecimiento. *Pensamiento psicológico*, 7(14), 63-72
- Cervantes, V.H. (2005). Interpretaciones del coeficiente de alpha de Cronbach. *Avances en Medicion*, 3, 9-25.
- Chentsova-Dutton, Y.E. y Tsai, J.L. (2009). Understanding depression across cultures. En Gotlib, I.H. y Hammen, C.L. (Ed). *Handbook of depression*. Londres: The Guilford Press
- Claeys, C., Nève, J., Tulkens, P. M. y Spinewine, A. (2012). Content validity and inter-rater reliability of an instrument to characterize unintentional medication discrepancies. *Drugs Aging*, 29, 577-591
- Clark, L. A., y Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*, 100, 316–336. Doi: 10.1037/0021-43X.100.3.316
- Clark, L.A. y Watson, D. (1995). Constructing validity: basic issues in objective scale development. *Psychol Assess*, 7, 309-319
- Cohen, R.J. y Swerdilk, M.E. (2006). *Pruebas y evaluación psicológicas. Introducción a las pruebas y a la medición* (Sexta edición). México: The McGraw-Hill Company
- Colasanti, V., Marianetti, M., Micacchi, F., Amabile, G. A., y Mina, C. (2010). Tests for the evaluation of depression in the elderly: a systematic review. *Archives of Gerontology and Geriatrics*, 50(2), 227-230.
- Cortina, J.M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98-104.
- Cosco, T.D., Doyle, F., Ward, M. y McGee, H. (2012). Latent structure of the Hospital Anxiety and Depression Scale: A 10-year systematic review. *Journal of Psychosomatic Research*, 72, 180-184. Doi: 10.1016/j.jpsychores.2011.06.008
- Costello, C. G. (1972). Depression: Loss of reinforcers or loss of reinforcer effectiveness? *Behavior Therapy*, 3, 240-247. Doi: 10.1016/S0005-7894(72)80084-4

- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of test. *Psychometrika*, 16, 297-334
- CronbachPacker, Abel L. (2001). *The SciELO Model for electronic publishing and measuring of usage and impact of Latin American and Caribbean scientific journals*. En: Second Icsu-Unesco International Conference electronic publishing in science. Paris, Unesco, 19-23 febrero 2001. Session III: responses from the scientific community.
- Currin, D., Haschke, C., Kluge, D., Niebuhr-Timpe, P. y Stüttgen, M. (Ed.). (2001). *Todas las medicinas*. Bogotá, Colombia: Círculo de lectores.
- da Silva FC, Gonçalves E, Arancibia BA, Bento, G.G., Castro, T.L., Hernandez, S.S. y da Silva, R. (2015). Estimators of internal consistency in health research: the use of the alpha coefficient. *Revista Peruana de Medicina Experimental y Salud Publica*, 32(1),129-138.
- Dávila Hernández, A., González González, R., Liangxiao, M. y Xin, N. (2016). Estudio sinomédico de la fisiopatología de la depresión. *Revista Internacional de Acupuntura*, 10(1), 9-15. Doi: 10.1016/j.acu.2016.01.003
- De Dios, C., Goikolea, J.M., Colom, F., Moreno, C. y Vieta, E. (2014). Los trastornos bipolares en las nuevas clasificaciones: DSM-5 y CIE-11. *Revista de Psiquiatría y Salud Mental*, 7(4), 179-185. Doi: 10.1016/j.rpsm.2014.07.005
- De los Santos-Roig, M., y Pérez-Meléndez, C. (2013). De la definición semántica a los ítems a través de las tablas de especificaciones: elaboración de la Escala sobre Representación de la Enfermedad, ERCE. *Anales de Psicología/Annals of Psychology*, 29(2), 345-359.
- de Vet, H.C.W., Eisinga, A., Riphagen II., Aertgeerts. B., Pewsner D. (2008). Chapter 7: Searching for Studies. In: *Cochrane Handbook for Systematic Reviews of Diagnostic Test Version 0.4: The Cochrane Collaboration*
- Deeks, J.J., Bossuyt, P.M., Gatsonics, C. (editors) (2013). *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. The Cochrane Collaboration*
- Delgado-Rico, E., Carretero-Dios, H. y Ruch, W. (2012). Content validity evidences in test development: An applied perspective. *International Journal of Clinical and Health Psychology*, 12(3), 449-460

- Departamento Administrativo de Ciencia, Tecnología e Innovación [Colciencias] (2016). Política Nacional para mejorar el impacto de las publicaciones científicas nacionales.
- DiBenedetti, D.B., Danchenko, N., Francois, C., Lewis, S., Davis, K.H. y Fehnel, S.E. (2012). Development of a family functioning scale for major depressive disorder. *Current Medical Research and Opinion*, 28, 303–13.
- Domínguez-Lara, S. A. D. y Merino-Soto, C. M. (2015). ¿Por qué es importante reportar los intervalos de confianza del coeficiente alfa de Cronbach?. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, 13 (2), 1326-1328.
- Escobar-Pérez, J. y Cuervo-Martínez, A. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en medición*, 6, 27-36
- Ezquiaga Terrazas, E., García López, A., Díaz de Neira, M., y García Barquero, M. J. (2011). "Depresión": Imprecisión diagnóstica y terapéutica. Importantes consecuencias en la práctica clínica. *Revista de la asociación española de neuropsiquiatría*, 31(3), 457-475.
- Fernández-Ballesteros, R. (2013). *Evaluación psicológica. Conceptos, métodos y estudios de caso*. Madrid: Ediciones Pirámides
- Ferrando, P. J., y Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del psicólogo*, 31(1), 18-33.
- Ferster, C. B. (1965). Classification of behavior pathology. En U. Krasner (Ed.), *Research in behavior modification*. Nueva York: Rinehart y Winston.
- Ferster, C. B. (1973). A functional analysis of depression. *American Psychologist*, 28(10), 857-870. Doi: 10.1037/h0035605
- Fitch, K., Bernstein, S.J., Aguilar, M.D., Burnand, B., LaCalle, J.R., Lázaro, P., Loo, M., McDonnell, J., Vader, J.P. y Kahan, J.P. (2001). *The RAND/UCLA Appropriateness Method User's Manual*: RAND corporation
- Franco-López, A., Sanz-Valero, J. y Culebras, J.M. (2016). Publicar en castellano, o en cualquier otro idioma que sea inglés, negativo para el factor de impacto y citaciones. *Journal of negative y no positive results*, 1(2), 65-70

- Fuensanta, M. (2015) Lectura crítica de un metaanálisis de una revisión sistemática. Murciasalud. Recuperado a partir de <https://www.murciasalud.es>
- Gabarda, V. (s.f.). El aprendizaje en la edad adulta. Características definitorias y diferenciales. Recuperado de <http://www.orientacionandujar.es/wp-content/uploads/2015/06/Ebook-Aprendizaje-Edad-Adulta-OK.pdf>
- García López de Arenosa, A., León, Quismondo, L. Gálvez Calero, C. y Hernández Jurdado, P. (2015). Protocolo diagnóstico del paciente depresivo. *Medicine*, 11(85), 5098-5102. Doi: 10.1016/j.med.2015.08.005
- George, D. y Mallery, P. (2003). *SPSS for Windows step by step: A Simple Guide and Reference*. 11.0 Update (4.<sup>a</sup> ed.). Boston: Allyn y Bacon
- Gerbing, D. W. y Anderson J. C. (1988). An update paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, 25 (2), 186-192.
- Gibson, J., McKenzie-McHarg, K., Shakespeare, J., Price, J. y Gray, R. (2009). A systematic review of studies validating the Edinburgh Postnatal Depression Scale in antepartum and postpartum women. *Acta Psychiatrica Scandinavica*, 119, 350-364. Doi: 10.1111/j.1600-0447.2009.01363.x
- Glass, G.V. (1976). Primary, Secondary, and Meta-Analysis of Research. *Educational Researcher*, 5(10), 3-8
- Gómez Caridad, I. y Bordons Gangas, M. (1996). Limitaciones en el uso de los indicadores bibliométricos para la evaluación científica. *Política científica* 46, 21-26
- González Llana, F.M. (2007). *Instrumentos de evaluación psicológica*. La Habana, Cuba: Editorial Ciencias Médicas
- González Martínez, X., Martín Cabeza, M., Sánchez Pascual, M. y Mirapeix Bedia, I. (2015) Trastornos del humor: trastornos depresivos. *Medicine*, 11(85), 5064-5074. Doi: 10.1016/j.med.2015.08.002

- Gordon, J. (2004). Developing and improving assessment instruments. *Assessment in Education: Principles, Policy and Practice*, 11, 243-245.
- Guadarrama, L., Escobar, A. y Zhang L. (2006). Bases neuroquímicas y neuroanatómicas de la depresión. *Revista de la facultad de Medicina*, 49(2), 66-72
- Gullion, C. M. & Rush, A. J. (1998). Toward a generalizable model of symptoms in major depressive disorder. *Biological Psychiatry* 44, 959-972.
- Hambleton, R. K. (1980). Test score validity and standard setting methods. En R. A. Berk (Ed.), *Criterion-Referenced Measurement: The State of the Art*. Baltimore: Johns Hopkins University Press.
- Hambleton, R. K. (1984). Validating the test score. En R. A. Berk (Ed.), *A Guide to Criterion-Referenced Test Construction* (pp. 199-230). Baltimore: Johns Hopkins University Press.
- Hambleton, R. K., Merenda, P. F., y Spielberger, C. D. (Eds.) (2005). *Adapting educational and psychological tests for cross-cultural assessment*. London: LEA.
- Hambleton, R.K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, 10, 229-240.
- Hambleton, R.K. (1996). Adaptación de tests para su uso en diferentes idiomas y culturas: fuentes de error, posibles soluciones y directrices prácticas. En J. Muñiz (Ed.), *Psicometría* (pp. 203-238). Madrid: Universitat.
- Hambleton, R.K. y Jong, J.H. (2003). Advances in translating and adapting educational and psychological tests. *Language Testing*, 20, 127-134.
- Hamilton, M. (1960) A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, 23, 56-62.
- Haynes, S. N., Richard, D., y Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological assessment*, 7(3), 238.

- Hermans, H. y Evenhuis, H. (2010). Characteristics of instruments screening for depression in adults with intellectual disabilities: Systematic review. *Research in Developmental Disabilities, 31*(6), 1109- 1120. Doi: 10.1016/j.ridd.2010.04.023
- Hernández -Guzmán, L., del Palacio, A., Freyre, M. y Alcázar-Olán, R. (2011). La perspectiva dimensional de la psicopatología. *Revista Mexicana de Psicología, 28*(2), 111-120
- Hernández Sampieri, R., Fernández Collado, C. y Baptista Lucio, M. (2015). *Metodología de la Investigación (6ª edición)*. México D.F: McGrawHill Education
- Hernández-Nieto, R. A. (2002), *Contributions to Statistical Analysis*. Mérida, Venezuela: Universidad de Los Andes.
- Herrero, J., y Meneses, J. (2006). Short Web-based versions of the perceived stress (PSS) and Center for Epidemiological Studies-Depression (CESD) Scales: A comparison to pencil and paper responses among Internet users. *Computers in Human Behavior, 22*(5), 830-846.
- Higgins, JPT. y Green, S. (editors) (2011). *Cochrane Handbook for Systematic Reviews of Interventions* version 5.1.0: The Cochrane Collaboration
- Hogan, T.P., Benjamin, A. y Brezinski, K.L. (2000). Reliability methods: a note on the frequency of use of various types. *Educ Psychol Meas, 60*, 523-531.
- Holländare, F., Andersson, G., y Engström, I. (2010). A comparison of psychometric properties between internet and paper versions of two depression instruments (BDI-II and MADRS-S) administered to clinic patients. *Journal of medical Internet research, 12*(5)
- Hu, L., y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.
- Huamaní, C., González, G., Curioso, W. H., y Pacheco-Romero, J. (2012). Redes de colaboración y producción científica sudamericana en medicina clínica, ISI Current Contents 2000-2009. *Revista médica de Chile, 140*(4), 466-475.



- Hung, C.I., Weng, L.J., Su, Y.J. y Liu, C. Y. (2006). Depression and somatic symptoms scale: a new scale with both depression and somatic symptoms emphasized. *Psychiatry and Clinical Neurosciences*, 60(6), 700–708, Doi: 10.1111/j.1440-1819.2006.01585.x
- Hunt, M. (1997). *How science takes stock : The story of meta-analysis*. Nueva York: Russell Sage Foundation
- Hyrkäs, K., Appelqvist-Schmidlechner, K y Oksa, L. (2003). Validating an instrument for clinical supervision using an expert panel. *International Journal of nursing studies*, 40(6), 619 -625.
- Ingram, R. E., y Wisnicki, K. (1991). Cognition in depression. En R. Magaro (Ed.), *Cognitive bases of mental disorders. Annual Review of Psychopathology* (vol. I, pp. 187-230). Newbury Park, CA: Sage.
- Israni, R. (2010). Medpage Tools Guide to Biostatistics. *Medpage today*, 1-11.
- Kane, M. T. (2009). Validating the interpretations and uses of test scores. En R. W. Lissitz (Ed.), *The concept of validity* (pp. 39-64). Charlotte, NC: Information Age Publishing.
- Kanter, J. W., Mulick, P. S., Busch, A. M., Berlin, K. S., y Martell, C. R. (2007). The Behavioral Activation for Depression Scale (BADs): Psychometric properties and factor structure. *Journal of Psychopathology and behavioral assessment*, 29, 191-202.  
Doi:10.1007/s10862-006-9038-5
- Kessler, R.C., Andrews, G., Cople, L.J., Hiripi, E., Mroczek, D.K., Normand, S.-L.T., Walters, E.E. y Zaslavsky, A.M. (2002). Short screening scales to monitor population prevalence and trends in non-specific psychological distress. *Psychological Medicine*, 32, 959-976.
- Kline R. (2005). *Principles and practice of structural equation modeling*. New York: Guilford Press.
- Kline, P. (2000). *Handbook of Psychological Testing*. London: Routledge.
- Kohn, R., Levav, I., Caldas de Almeida, J.M., Vicente, B., Andrade, L., Caraveo-Anduaga, J.J., Saxena, S. y Saraceno, B. (2005). Los trastornos mentales en América Latina y el Caribe:

- asunto prioritario para la salud pública. *Revista panamericana de Salud Pública*, 18(4), 229-240
- Kohn, R., Saxena, S. Levav, I., y Saraceno, B. (2004). The treatment gap in mental health care. *Boletín de la Organización Mundial de la Salud*, 82(11)
- Lako, I.M., Bruggeman, R., Knegterin, H., Wiersman, D., Schoevers, R.A., Slooff, C.J. y Taxis, K. (2012). A systematic review of instruments to measure depressive symptoms in patients with schizophrenia. *Journal of Affective Disorders*, 140(1), 38-47. Doi: 10.1016/j.jad.2011.10.014.
- Lara, M.A., Navarro, C., Mondragón, L., Rubí, N.A. y Lara, M.C. (2002). Validez y confiabilidad del MHI-5 para evaluar la depresión de mujeres en primer nivel de atención. *Salud Mental*, 25(6), 13-20
- Lawley, D. N. (1943). On problems connected with item selection and test construction. *Proceedings of the Royal Society of Edimburg*, 61, 273-287.
- Lawley, D. N. (1944). The factorial analysis of multiple item tests. *Proceedings of the Royal Society of Edimburg*, 62, 74-82.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28, 563-575.
- Letelier S, L.M., Manríquez M, J.J. y Rada G, G. (2005). Revisiones sistemáticas y metaanálisis : ¿son la mejor evidencia? *Boletín de la Escuela de Medicina*, 30(2), 37-39
- Levine, E. L., Maye, D. M., Ulm, R. A. y Gordon, T. R. (1997). A methodology for developing and validating minimum qualifications (MQs). *Personnel Psychology*, 50, 1009-1023.
- Lewinsohn, P. M. (1974). A behavioral approach to depression. En R. J. Friedman y M. M. Katz (Eds.). *The psychology of depression: Contemporary theory and research* (pp. 157-185). Nueva York: Wiley.
- Lewinsohn, P. M. (1976). Activity schedules in the treatment of depression. En C. E. Thoresen y J. D. Krumboltz (Eds.), *Counseling methods* (pp. 74-83). Nueva York: Holt, Rinehart y Winston.

- Lewinsohn, P. M.; Hoberman, H. M.; Teri, L., y Hautzinger, M. (1985). An integrative theory of depression. En S. Reiss y R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp.331-359). Orlando, FL: Academic Press.
- Ley 1616 (21 de enero de 2013). Por medio de la cual se expide la ley de salud mental y se dictan otras disposiciones. Recuperado de <http://wsp.presidencia.gov.co/Normativa/Leyes/Documents/2013/LEY%201616%20DEL%2021%20DE%20ENERO%20DE%202013.pdf>
- Licea de Arenas, J., Castañis Lomnitz, H. y Arenas Licea, J. (2002). Significant Mexican research in the health sciences: a bibliometric analysis. *Scientometrics*, 53, 39-48.
- Lillo, S. y Martini, N. (2013). Principales Tendencias Iberoamericanas en Psicología Clínica. Un Estudio Basado en la Evidencia Científica. *Terapia Psicológica*, 31(3), 363-371
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología/Annals of Psychology*, 30(3), 1151-1169.
- López, E. J. A. y Díaz, C. S. del. (1995). Problemas y tendencias actuales de la información científico-médica. *Revista Cubana Salud Pública*, 21(2).
- López, W. L., Silva, L. M., García-Cepero, M. C., Bustamante, M. C. A., y López, E. A. (2011). Retos para la colaboración nacional e internacional en la psicología latinoamericana: un análisis del sistema RedALyC, 2005-2007. *Estudios de Psicología (Natal)*, 16(1) 17-22.
- Lord, F. M. (1952). A theory of test scores. *Psychometric Monographs*, 7.
- Lord, F.M. y Novick, M.R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison Wesley.
- Loveys, K., Torrez, J., Fine, A., Moriarty, G., y Coppersmith, G. (2018, June). Cross-cultural differences in language markers of depression online. In Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic (pp. 78-87).

- Lovibond, P., y Lovibond, S. (1995b). The structure of negative emotional states: comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. *Behaviour Research and Therapy*, 33, 335 – 343.
- Lovibond, S., y Lovibond P. (1995a). *Manual for the depression anxiety stress scales*. Sydney, Australia: Psychology Foundation of Australia
- Luján- Tangarife, J. L., y Cardonas-Arias, J. A. C. (2015). Construcción y validación de escalas de medición en salud: revisión de propiedades psicométricas. *Archivos de medicina*, 11(3), 1.
- Luna-Morales, M. E., y Collazo-Reyes, F. (2006). Las revistas latinoamericanas y caribeñas en los rankings de la ciencia internacional. In III Congreso Internacional de Bibliometría Gilberto Sotolongo Aguilar (pp. 1-12).
- Luque, R. y Berrios, G. (2011). Historia de los trastornos afectivos. *Revista Colombiana de Psiquiatría*, 40, 130-146. Doi: 10.1016/S0034-7450(14)60199-2
- Mansilla, M.A. (2000). Etapas del desarrollo humano. *Revista de Investigación en Psicología*, 3(2), 105-116
- Martín Ruiz, E., Molina Linde, J.M. y Martín Ruiz, O. (2010). Análisis bibliométrico de la producción científica sobre la depresión unipolar en el período 2000-2005. *Psicología y Salud*, 20(2), 261-272
- Martínez Acosta, U., Adefna Pérez, R. y Izquierdo Lara, F. (2009). *Medicina Basada en la Evidencia*. Editorial El Cid Editor.
- Martínez Arias, M.R., Hernández Lloreda, M.J. y Hernández Lloreda, M.V. (2006). *Psicometría*. Madrid, España: Alianza Editorial.
- Martínez García, J.A. y Martínez Caro, L. (2009). La validez discriminante como criterio de evaluación de escalas. ¿teoría o estadística?. *Universitas Psychologica*, 8(1), 27-36.
- Martínez, M. (1995). *Psicometría. Teoría de los tests psicológicos y educativos*. Madrid: Síntesis
- Martínez-Taboas, A. (2014). Prácticas psicológicas basadas en la evidencia: beneficios y retos para Latinoamérica. *Revista Costarricense de Psicología*, 33(2), 63-78

- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah: Lawrence Erlbaum Associates, Inc.
- Meneses, J., Barrios, M., Bonillo, A., Cosculluela, A., Lozano, L.M., Turbany, J. y Valero, S. (2013). *Psicometría*. Barcelona, España: Editorial UOC
- Messick, S. (1980). Test validity and ethics of assessment. *American Psychologist*, 35, 1012-1027.
- Ministerio de Salud y Protección Social (2013). Guía de práctica clínica: Detección temprana y diagnóstico del episodio depresivo y trastorno depresivo recurrente en adultos. Atención integral de los adultos con diagnóstico de episodio depresivo o trastornos depresivo recurrente. Recuperado de [http://gpc.minsalud.gov.co/gpc\\_sites/Repositorio/Conv\\_500/GPC\\_td/GPC\\_td\\_completa.aspx](http://gpc.minsalud.gov.co/gpc_sites/Repositorio/Conv_500/GPC_td/GPC_td_completa.aspx)
- Ministerio de Salud y Protección Social (2015). Plan Estratégico Sectorial 2014-2018. Recuperado de: <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/PES/plan-estrategico-sectorial-2014-2018.pdf>
- Ministerio de Salud y Protección Social (2017). Boletín de salud mental: depresión. Tomado de <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/boletin-depresion-marzo-2017.pdf>
- Ministerio de Salud y Protección Social y Departamento Administrativo de Ciencia, Tecnología e Innovación [COLCIENCIAS] (2015). Encuesta Nacional de Salud Mental. Recuperado de: [http://www.odc.gov.co/Portals/1/publicaciones/pdf/consumo/estudios/nacionales/CO031102015-salud\\_mental\\_tomoI.pdf](http://www.odc.gov.co/Portals/1/publicaciones/pdf/consumo/estudios/nacionales/CO031102015-salud_mental_tomoI.pdf)
- Ministerio de sanidad, servicios sociales e igualdad (2014). Guía de práctica clínica sobre el manejo de la Depresión en el adulto. Recuperado de [http://www.guiasalud.es/GPC/GPC\\_534\\_Depresion\\_Adulto\\_Avaliat\\_compl.pdf](http://www.guiasalud.es/GPC/GPC_534_Depresion_Adulto_Avaliat_compl.pdf)

- Mitchell, J. V. (1986). Measurement in the larger context: Critical current issues. *Professional Psychology: Research and Practice*, 17, 544-550.
- Moher, D. Cook, D. Eastwood, S., Olki, I., Rennie, D. y Stroup, D. for the QUOROM group (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet*, 354(9193), 1896-1900
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. y the PRISMA Group, (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. *Annals of Internal Medicine*, 151(4), 264-269
- Mokkink, L.B., de Vet, H.C.W., Prinsen, C.A.C., Patrick, D.L., Alonso, J., Bouter, L.M. y Terwee, C.B. (2017). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Quality of life research*. Doi 10.1007/s11136-017-1765-4
- Mokkink, L.B., Terwee, C.B., Knol, D.L., Stratford, P.W., Alonso, J., Patrick, D.L., Riphagen, I., Knol, L.M. y de Vet, H.C.W. (2009). Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Qual Life Res*, 18, 313-333
- Mokkink, L.B., Terwee, C.B., Knol, D.L., Stratford, P.W., Alonso, J., Patrick, D.L., Bouter, L.M. y de Vet, H.C.W. (2006). Protocol of the COSMIN study: Consensus-Based Standards for the selection of health Measurement Instruments. *BMC Medical Research Methodology*, 6(2), 1-7 doi:10.1186/1471-2288-6-2
- Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M. y de Vet, H.C.W. (2012). COSMIN checklist manual. Recuperado de <http://www.cosmin.nl/images/upload/files/COSMIN%20checklist%20manual%20v9.pdf>
- Montero, I. y León, O.G. (2007). A guide for naming research studies in Psychology. *International journal of clinical and health psychology*, 7(3), 847-862
- Montgomery, S.A y Åsberg, M. (1979). A new depression rating scale designed to be sensitive to change. *British Journal of Psychiatry*, 134, 382-389.
- Moscoso, M. S. (2014). Avances en la medición psicométrica de la depresión. *Liberabit*, 20(1), 29-39.

- Muñiz, J. (1994). *Teoría clásica de los tests*. Madrid: Pirámide.
- Muñiz, J. (1998). La medición de lo psicológico. *Psicothema*, 10(1), 1-21
- Muñiz, J. (2010). Las teorías de los test: Teoría Clásica y Teoría de Respuesta a los Ítems. *Papeles del psicólogo*, 31(1), 57-66
- Muñiz, J. (2018). *Introducción a la Psicometría, Teoría clásica y TRI*. Madrid: Ediciones Pirámide.
- Muñiz, J., Elosua, P., y Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25(2), 151-157.
- Muñiz, J., Hernández, A. y Ponsoda, V. (2015). Nuevas directrices sobre el uso de los test: Investigación, Control de calidad y Seguridad. *Papeles del Psicólogo*, 36(3), 161-173
- Mutumba, M., Tomlinson, M., y Tsai, A. C. (2014). Psychometric properties of instruments for assessing depression among African youth: A systematic review. *Journal of Child and Adolescent Mental Health*, 26(2), 139-156.
- Nezu, A. M.; Nezu, C. M., y Perri, M. G. (1989). *Problem-solving therapy for depression. Theory, research, and clinical guidelines*. Toronto: John Wiley y Sons.
- Nezu, A.M. (1987). A problem-solving formulation of depression: A literature review and proposal of a pluralistic model. *Clinical Psychology Review*, 7, 121-144. Doi: 10.1016/0272-7358(87)90030-4
- Nezu, A.M. y Ronan, G.F. (1985). Life stress, current problems, problem solving, and depressive symptoms: An integrative model. *Journal of Consulting and Clinical Psychology*, 53, 693-697.
- Nolen-Hoeksema, S. (1991). Response to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology*, 100(4), 569-582.
- Nolen-Hoeksema, S. (2000). Further evidence for the role of psychosocial factors in depression chronicity. *Clinical Psychology: Science and Practice*, 7, 224-227. Doi: 10.1093/clipsy.7.2.224
- Nunnally, J.C. y Bernstein, I.J. (1995). *Teoría psicométrica*. Madrid: McGraw-Hill.

- Ojeda, C., Silva, H., Cabrera, J., Tchimino, C., de la Carrera, C., y Molina, A. (2003). Validez interna y concurrente de una escala breve autoadministrada de siete ítems para la depresión, evaluada en relación a los inventarios de Beck y Hamilton: estudio preliminar. *Revista chilena de neuro-psiquiatría*, 41(1), 31-38
- Olfson, M., Shea, S., Feder, A., Fuentes, M., Nomura, Y., Gerneroff, M., y Weissman, M. M. (2000). Prevalence of anxiety, depression, and substance use disorders in an urban general medicine practice. *Archives of family medicine* 9(9), 876.
- Organización Mundial de la Salud [OMS]. (1994). Clasificación de los trastornos mentales y del comportamiento. En, *Guía de bolsillo de la Clasificación Internacional de Enfermedades [CIE-10]*. Madrid, España: Editorial Médica Panamericana.
- Organización Mundial de la Salud [OMS]. (2004). Invertir en Salud Mental, Recuperado de: [http://www.who.int/mental\\_health/advocacy/en/spanish\\_final.pdf](http://www.who.int/mental_health/advocacy/en/spanish_final.pdf)
- Organización Mundial de la Salud [OMS]. (2008). MhGAP Programa de Acción para Superar las Brechas en Salud Mental. Mejora y ampliación de la atención de los trastornos mentales, neurológicos y por abuso de sustancias. Recuperado de: [http://www.who.int/mental\\_health/mhgap/mhgap\\_spanish.pdf](http://www.who.int/mental_health/mhgap/mhgap_spanish.pdf)
- Organización Mundial de la Salud [OMS]. (2012). 65ª Asamblea Mundial de la Salud. Recuperado de [http://apps.who.int/gb/ebwha/pdf\\_files/WHA65-REC1/A65\\_REC1-sp.pdf](http://apps.who.int/gb/ebwha/pdf_files/WHA65-REC1/A65_REC1-sp.pdf)
- Organización Mundial de la Salud [OMS]. (2013). 66ª Asamblea Mundial de la Salud. Recuperado de [http://apps.who.int/gb/ebwha/pdf\\_files/WHA66-REC1/A66\\_REC1-sp.pdf](http://apps.who.int/gb/ebwha/pdf_files/WHA66-REC1/A66_REC1-sp.pdf)
- Organización Mundial de la Salud [OMS]. (2018). Salud mental: fortalecer nuestra respuesta. Recuperado de: <http://www.who.int/mediacentre/factsheets/fs220/es/>
- Organización Mundial de la Salud [OMS]. (2020). Depresión. Recuperado de: <https://www.who.int/es/news-room/fact-sheets/detail/depression>
- Organización Panamericana de la Salud [OPS]. (2014). Plan de acción sobre salud mental 2015-2020. Recuperado de: [http://www.paho.org/hq/index.php?option=com\\_content&view=article&id=11337%3Apl](http://www.paho.org/hq/index.php?option=com_content&view=article&id=11337%3Apl)



- an-of-action-on-mental-health-2015-2020&catid=8190%3Ageneral&Itemid=41600&lang=es
- Organización Panamericana de la Salud [OPS]. (2014). Plan estratégico de la OPS 2014-2019. En pro de la salud: Desarrollo sostenible y equidad. Recuperado de: [https://issuu.com/paho2012/docs/planoestrat14-19\\_execesp\\_med](https://issuu.com/paho2012/docs/planoestrat14-19_execesp_med)
- Osca-Lluch, J., González-Sala, F., Fonseca Baeza, S. y Civera Mollá, C. (2017). Historia de la investigación española en psicología vista a través de las revistas científicas recogidas en la Web of Science durante el período 2003-2012. *Universitas Psychologica* 16(2), 1-10. <https://doi.org/10.11144/Javeriana.upsy16-2.hiep>
- Osman, A., Freedenthal, S., Gutierrez, P. M., Wong, J. L., Emmerich, A., y Lozano, G. (2011). The Anxiety Depression Distress Inventory–27 (ADDI–27): A short version of the Mood and Anxiety Symptom Questionnaire–90. *Journal of Clinical Psychology*, 67, 591–608.
- Papalia, D. Wendkos Olds, S. y Duskin Fieldman, R. (2009). *Desarrollo humano*. México, D.F: Editorial The Mcgrow Hill Companies, Inc.
- Páramo, P. (2020). Cómo elaborar una Revisión sistemática. DOI: 10.13140/RG.2.2.31465.85608
- Parker, G. y Hadzi-Pavlovic, D. (1996). *Melancholia: A Disorder of Movement and Mood*. Cambridge University Press, New York.
- Patalano, M. (2005). Las publicaciones del campo científico: las revistas académicas de América Latina. *Anales de documentación* 8, 217-235
- Pedrosa, I., Suárez-Álvarez, J., y García-Cueto, E. (2013). Evidencias sobre la validez de contenido: avances teóricos y métodos para su estimación. *Acción psicológica*, 10(2), 3-18.
- Perestelo-Pérez, L. (2012). Standards on how to develop and report systematic reviews in Psychology and Health. *International Journal of Clinical and Health Psychology*, 13, 49-57.

- Pérez-Achiaga, N., Nelson, S. y Hassiotis, A. (2009). Instruments for the detection of depressive symptoms in people with intellectual disabilities. A systematic review. *Journal of intellectual disabilities, 13*(1), 55-76
- Pocklington, C., Gilbody, S. Manea, L. y McMillan, D. (2016). The diagnostic accuracy of brief versions of the Geriatric Depression Scale: a systematic review and meta-analysis. *International Journal of Geriatric Psychiatry, 31*(8), 837-57 Doi: 10.1002/gps.4407
- Prieto, G. y Delgado, A.R. (2010). Fiabilidad y validez. *Papeles del psicólogo, 31*(1), 67-74
- Prinsen C.A.C., Mokkink, L.B., Bouter, L.M., Alonso, J., Patrick, D.L., de Vet, H.C.W. y Terwee, C.B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, Doi: doi.org/10.1007/s11136-018-1798-3
- Radloff, L.S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measure, 1*(3), 385-401.
- Ramada-Rodilla, J.M., Serra-Pujadas, C. y Delclós-Clanchet, G.L. (2012). Adaptación cultural y validación de cuestionarios de salud: revisión y recomendaciones metodológicas. *Salud pública de México, 55*(1), 57-66
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: The Danish Institute for Educational Research.
- Ratto de Sala, M. C. y Dellamesa, A. B. (2001). Difusión, acceso y visibilidad de publicaciones científicas seriadas de Iberoamérica. El sistema Latindex. *Dominguezia, 17*(1).
- Rehm, L. P. (1977). A self-control model of depression. *Behavior Therapy, 8*(5), 787-804. Doi: 10.1016/S0005-7894(77)80150-0
- Reilly, T.J., MacGillivray, S.A., Reid, I.C. y Cameron, I.M. (2014). Psychometric properties of the 16-item Quick Inventory of Depressive Symptomatology: A systematic review and meta-analysis. *Journal of Psychiatric Research, 60*, 132-40 Doi: 10.1016/j.jpsychires.2014.09.008

- Reuland, D.S., Cherrington, A., Watkins, G.S., Bradford, D.W., Blanco, R.A. y Gaynes, B.N. (2009). Diagnostic accuracy of Spanish language depression-screening instruments. *Annals of family medicine*, 7(5), 455-462.
- Ríos Gómez, C. y Herrero Solana, V. (2005). La producción científica latinoamericana y la ciencia mundial: una revisión bibliográfica (1989-2003). *Revista Interamericana de Bibliotecología*, 28(1).
- Ritterband, L.M. y Spielberger, C.D. (1996). Construct Validity of the Beck Depression Inventory as a measure of state and trait depression in nonclinical populations. *Depression and Stress*, 2, 123-145.
- Roberts, P. y Priest, H. (2006). Traynor M. Reliability and validity in research. *Nurs Stand*, 20, 41-45.
- Roberts, R. (1980). Reliability of the CES-D Scale in different ethnic contexts. *Psychiatry Research*, 2, 125-134.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (ed.rev.). Newbury Park CA: Sage
- Rovinelli, R. J. y Hambleton, R. K. (1977). On the use of content specialists in the assessment of criterion-referenced test item validity. *Dutch Journal of Educational Research*, 2, 49-60.
- Rubio, D. M., Berg-Weber, M., Tebb, S. S., Lee, E. S. y Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research*, 27(2), 94-104.
- Rubio-Stipek, M., Hicks, M.H.R y Tsuang, M.T. (2002). Cultural factors influencing the selection, use, and interpretation of psychiatric measures. En: Rush, A.J., Pincus, H.A., First, M.B., Zarin, D.A., Blacker, D. y Endicott J, et al. *Handbook of psychiatric measures*. Washington: American Psychiatric Association.
- Rush, A. J., Trivedi, M. H., Ibrahim, H. I., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S. G., Manber, R., Thase, M. E., Kocsis, J. H. y Keller, M. B. (2003a). The 16-item Quick Inventory of Depressive Symptomatology (QIDS)

- Clinician Rating (QIDS-C) and Self-Report (QIDS-SR): psychometric evaluation in patients with chronic major depression. *Biological Psychiatry* 54, 573–583.
- Rush, A.J., Gilles, D.E., Schlessner, M.A., Fulton, C.L., Weissenburger, J. y Burns, C. (1986). The Inventory for Depressive Symptomatology (IDS): preliminary findings. *Psychiatry Research*, 18, 65-87.
- Rush, A.J., Guillion, C.M., Basco, M.R., Jarret, R.B., Trivedi, M.H. (1996). The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychological Medicine*, 26, 477-86.
- Salvador Carulla, L., Romero López-Alberca, C. y González-Saiz, F. (2000): "Guías para la descripción y selección de Instrumentos de evaluación en Psiquiatría". En: Bulbena, A., Berrios, G. E. y Fdez. de Larrinoa, P. (Eds.): *Medición Clínica en Psiquiatría y Psicología*. Barcelona: Masson-Salvat, págs. 15-34
- Sánchez, R. y Gómez, C. (1998). Conceptos básicos sobre validación de escalas. *Revista Colombiana de Psiquiatría*, 27, 121-130
- Sánchez-Meca, J. (2010). Cómo realizar una revisión sistemática y un meta-análisis. *Aula abierta*, 38(2), 53-64
- Sánchez-Meca, J. y Botella, J. (2010). Revisiones sistemáticas y meta-análisis: Herramientas para la práctica profesional. *Papeles del psicólogo*, 31(1), 7-17
- Santiago-Rivera, A.Z., Benson-Flórez, G., Santos, M.M. y Lopez, M. (2015). Latinos and Depression: Measurement Issues and Assessment. En: Geisinger, K.F. (Ed). *Psychological Testing of Hispanics, Second Edition: Clinical, cultural and Intellectual Issues*.
- Santor, D. A. (2006). New Coke, Rosetta Stones, and Functional Data Analysis: Recommendations for Developing and Validating New Measures of Depression. *Measurement*, 4(3), 188-197.
- Sanz, J. y Vázquez, C. (2009). Trastornos del estado de ánimo: teorías psicológicas. En Belloch, A., Sandín, B. y Ramos, F. (Ed.). *Manual de psicopatología*, Volumen II. Madrid, España: McGrawHill.

- Sanz, J., Izquierdo, A., y García Vera, M. P. (2013). Una revisión desde la perspectiva de la validez de contenido de los cuestionarios, escalas e inventarios autoaplicados más utilizados en España para evaluar la depresión clínica en adultos. *Psicopatología Clínica, Legal y Forense*, 13, 139-175.
- Satorra, A. y Bentler, P. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507-514
- Seligman, M. E. P. (1975). *Helplessness: On depression, development and death*. San Francisco: Freeman. (Versión española: Debate, Madrid, 1981)
- Sierra Flórez, P. y Gómez Vargas, M. (2019). Prácticas editoriales en materia de visibilidad de revistas científicas latinoamericanas de Ciencias Sociales y Humanas. *Información, cultura y sociedad*, 40, 131-150
- Silverstone, P. H. (1994). Poor efficacy of the Hospital Anxiety and Depression Scale in the diagnosis of major depressive disorder in both medical and psychiatric patients. *Journal of Psychosomatic Research*, 38(5), 441-450.
- Sireci, S. G. y Geisinger, K. F. (1992). Analyzing test content using cluster analysis and multidimensional scaling. *Applied Psychological Measurement*, 16, 17-31.
- Sireci, S.G. (1998). Gathering and analyzing content validity data. *Educational Measurement*, 5, 299-321.
- Skinner, B.F. (1953). *Science and human behavior*. New York: McMillan.
- Smith, G.T., Fischer, S. y Fister, S.M. (2003). Incremental validity principles in test construction. *Psychological Assessment*, 15, 467-477.
- Smith, M.L. y Glass, G.V. (1983). Meta-analysis of research on class size and its relationship to attitude and instruction. En R.J. Light (Ed.), *Evaluation studies annual review*, Vol. 8 (pp. 433-447). Beverly Hills, CA: Sage.
- Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 15, 72-101.

- Spearman, C. (1907). Demonstration of formulae for true measurement of correlation. *American Journal of Psychology*, 18, 161-169.
- Spearman, C. (1913). Correlations of sums and differences. *British Journal of Psychology*, 5, 417-126.
- Spielberger CD, Agudelo D, Buela-Casal G. (2008). *Inventario de Depresión Estado/ Rasgo (IDER)*. Madrid: TEA Ediciones
- Spielberger, C. (1999). Evaluación de la depresión: Eutimia y Distimia. Conference presented in the International Symposium about Depression, Granada, Spain.
- Steiger, J. H. (1980). Statistically based tests for the number of common factors. *In the annual meeting of the Psychometric Society*.
- Streiner, D.L. (2003). Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess*, 80, 99-103.
- Suárez Parra, A.B. (2014). La estadística en la validación de escalas, una visión práctica para su construcción o su adaptación. *I3+*, 2(2), 46-61.
- Tabachnick, B. y Fidell, L. (2001). *Using multivariate statistics*. New York: Harper and Row.
- Teasdale, J. D. (1983). Negative thinking in depression: Cause, effect, or reciprocal relationship? *Advances in Behavior Research and Therapy*, 5, 3-25. Doi: 10.1016/0146-6402(83)90013-9
- Teasdale, J. D. (1988). Cognitive vulnerability to persistent depression. *Cognition and Emotion*, 2, 247-274. Doi: 10.1080/02699938808410927
- Téllez-Zenteno, J. F., Morales-Buenrostro, L. E., y Estañol, B. (2007). Análisis del factor de impacto de las revistas científicas latinoamericanas. *Revista médica de Chile*, 135(4), 480-487.
- Terblanche, N.S. y Boshoff, C. (2006). Improved scaled development in marketing. An empirical illustration. *Int J Mark Res*, 50, 105-119.
- Terwee, C.B., Mokkink, L.B., Knol, D.L., Ostelo, R.W.J. G., Bouter, L.M. y de Vet, H.C.W. (2012). Rating the methodological quality in systematic reviews of studies on

- measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*, 21, 651-657
- The Campbell Collaboration (2014). Campbell systematic reviews: policies and guidelines Series No.1 DOI: 10.4073/cpg.2016.1
- Thurstone, L. L. (1925). A method of scaling psychological and educational tests. *The Journal of Educational Psychology*, 16, 433-451.
- Torres, G. (2011). Una aproximación al análisis de competitividad aplicando la técnica del modelo de ecuaciones estructurales. *TESCoalt*, 13(31), 1-9.
- Torres-Fonseca, A. y López-Hernández, D. (2014). Criterios para publicar artículos de revisión sistemática. *Revista de Especialidades Médico-Quirúrgicas*, 19(3), 393-399
- Trujols, J., Ballesteros, J., Solá, I. y Portella, M.J. (2015). Improving systematic reviews on psychometric of measurements instruments: a letter to the Editor commenting on Really et al.'s (2015) review of the psychometric literature on QIDS. *Jounarl of Psychiatric Research*, 62, 136-137
- Tucker, L. R. (1961). Factor Analysis of Relevance Judgments: An Approach to Content Validity. En A. Anastasi (Ed.), *Testing Problems in Perspective* (pp. 577-586). Washington, DC.: American Council on Education
- Ugarte, F. (2004). La edición de revistas científicas en Latinoamérica. *Revista Chilena de Pediatría*, 75(6), 509-511.
- Urra Medina, E. y Barría Pailaquilén, R.M. (2010). La revisión sistemática y su relación con la práctica basada en la evidencia en salud. *Revista Latino-Americana de Enfermagem*, 18(4), 1-8
- Vallejo J. y Leal, C. (2005). *Tratado de Psiquiatría*. Barcelona: Ars Medica.
- Vallejo-Medina, P., Gómez-Lugo, M., Marchal-Bertrand, L., Saavedra-Roa, A., Soler, F., y Morales, A. (2017). Developing guidelines for adapting questionnaires into the same language in another culture. *Terapia psicológica*, 35(2), 181-194.

van Ballegooijen, W., Riper, H., Cuijpers, P., van Oppen, P., y Smit, J. H. (2016). Validation of online psychometric instruments for common mental health disorders: a systematic review. *BMC psychiatry*, *16*(1), 45.

Van der Does, W. (2002). Cognitive reactivity to sad mood: Structure and validity of a new measure. *Behaviour Research and Therapy*, *40*, 105–119. Doi: 10.1016/S0005-7967(00)00111-X

Vásquez, C., Hernangómez, L., Hervás, G. y Nieto, M. (2005). Evaluación de la depresión. En Caballo, V. (ed). *Manual para la evaluación cognitivo-conductual de los trastornos psicopatológicos*. Madrid: Pirámide, 1-80

Vázquez Machado, A. (2016). Los factores psicosociales y la depresión. *Multimed*, *20*(3), 586-602

Ventura-León, J. L., y Caycho-Rodríguez, T. (2017). El coeficiente Omega: un método alternativo para la estimación de la confiabilidad. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, *15*(1), 625-627.

Vera-Villaruel, P., López-López, W., Lillo, S., y Silva, L. M. (2011). La producción científica en psicología latinoamericana: Un análisis de la investigación por países. *Revista Latinoamericana de Psicología*, *43*(1), 95-104

Vilagut, G., Forero, C. G., Barbaglia, G., y Alonso, J. (2016). Screening for depression in the general population with the center for epidemiologic studies depression (CES-D): a systematic review with meta-analysis. *PloS one*, *11*(5).

Wancata, J. Alexandrowicz, R. Marquart, B. Weiss, M. y Friedrich, F. (2006). The criterion validity of the Geriatric Depression Scale: a systematic review. *Acta Psychiatrica Scandinavica*, *114*(6), 398- 410

Wang, Y.P. y Gorenstein, C. (2013). Assessment of depression in medical patients: A systematic review of the utility of the Beck Depression Inventory-II. *Clinics*, *68*(9), 1274- 1287. Doi: 10.6061/clinics/2013(09)15



- Watson, D., Clark, L. A., y Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070. <http://dx.doi.org/10.1037/0022-3514.54.6.1063>
- Watson, D., O'Hara, M.W., Simms, L.J., Kotov, R., Chmielewski, M., McDade-Montez, E.A., Gamez, W. y Stuart, S. (2007). Development and validation of the Inventory of Depression and Anxiety Symptoms (IDAS). *Psychological assessment*, 19, 253-268. Doi:10.1037/1040-3590.19.3.253
- Widakowich, C. (2012). El enfoque dimensional vs el enfoque categórico en psiquiatría: aspectos históricos y epistemológicos. *Alcmeon*, 17(4), 365-374
- Widiger, T.A. y Samuel, D. (2005). Diagnostic categories or dimensions? A question for the Diagnostic and Statistical Manual of Mental Disorder-Fifth Edition. *Journal of Abnormal Psychology*, 114(4), 494-504. Doi: 10.1037/0021-843X.114.4.494
- Winokur, G. (1997). All roads lead to depression: Clinically homogeneous, etiologically heterogeneous. *Journal of Affective Disorders*, 45, 97-108.
- World Federation for Mental Health [WFMH]. (2012). Depression: A global crisis. Recuperado de: [http://www.who.int/mental\\_health/management/depression/wfmh\\_paper\\_depression\\_wmhd\\_2012.pdf](http://www.who.int/mental_health/management/depression/wfmh_paper_depression_wmhd_2012.pdf)
- Zapata García, R., Cano Prous, A. y Moyá Querejeta, J. (2002). Tareas del desarrollo en la edad adulta. *Psiquis*, 23(5), 185-197
- Zarragoitía Alonso, I. (2011). *Depresión: generalidad y particularidades*. Ciudad de la Habana, Cuba: Editorial Ciencias Médicas.
- Zimmerman, M. y McGlinchey, J.B. (2008). Depressed patients' acceptability of the use of self-administered scales to measure outcome in clinical practice. *Annals of Clinical Psychiatry*, 20(3):125-9.
- Zung, W. W. K. (1965). Self-report depression scale. *Archive of General Psychiatry*, 12, 63-70.

**Referencias de los estudios incluidos**

- Abello Luque, D., Cortés Peña, O.F., Fonseca Consuegra, L., García Roncallo, P. y Meriño Buitrago, J. (2013). Escala Multidimensional de Trastornos Afectivos–emta: análisis desde la teoría clásica de los test y la teoría de respuesta al ítem. *Suma psicológica*, 20(2), 203-216
- Abello-Luque, D., Cortes-Peña, O., García-Montaño, E., García-Roncallo, P. y Nieto-Betancourt, L. (2017). Escala Multidimensional de Trastornos Afectivos, EMTA: Generación de Indicadores Psicométricos y Escalas Normativas de Uso Clínico. *Revista Iberoamericana de Diagnóstico y Evaluación*, 43(1), 187-198
- Agudelo Vélez, D. M., Gómez Maquet, Y., y Lopez, P. L. (2013). Propiedades psicométricas del Inventario de Depresión Estado Rasgo (IDER) con una muestra de población general colombiana. *Avances en Psicología Latinoamericana*, 32(1), 71-84. Doi: [dx.doi.org/10.12804/apl32.1.2014.05](https://doi.org/10.12804/apl32.1.2014.05)
- Agudelo, D., Carretero-Dios, H., Blanco Picabia, A., Pitti, C., Spielberger, C., y Buela-Casal, G. (2005). Evaluación del componente afectivo de la depresión: análisis factorial del ST/DEP revisado. *Salud mental*, 28(3), 32-41.
- Agüera-Ortiz, L., Montón, C., Cuervo, J., Medina, E., Díaz-Cuervo, H. y Maurino, J. (2013). Adaptation into spanish of the Clinically Useful Depression Outcome Scale (cudos) for assessing major depressive disorder from the patient's perspective. *Actas Españolas de Psiquiatría*, 41(5), 287-300
- Antúnez, Z., y Vinet, E. V. (2012). Escalas de depresión, ansiedad y Estrés (DASS-21): Validación de la Versión abreviada en Estudiantes Universitarios Chilenos. *Terapia psicológica*, 30(3), 49-55.
- Arbona, C., Burrridge, A., y Olvera, N. (2017). The Center for Epidemiological Studies Depression Scale (CES-D): Measurement equivalence across gender groups in Hispanic college students. *Journal of affective disorders*, 219, 112-118.

- Arnau, R. C., Meagher, M. W., Norris, M. P., y Bramson, R. (2001). Psychometric evaluation of the Beck Depression Inventory-II with primary care medical patients. *Health Psychology, 20*(2), 112.
- Barajas, S., Garra, L., y Ros, L. (2017). Avoidance in anxiety and depression: adaptation of the cognitive-behavioral avoidance scale in a Spanish sample. *The Spanish journal of psychology, 20*, 1-11
- Barraca, J., Pérez-Álvarez, M., y Bleda, J. H. L. (2011). Avoidance and activation as keys to depression: Adaptation of the Behavioral Activation for Depression Scale in a Spanish sample. *The Spanish journal of psychology, 14*(2), 998-1009.
- Bentley, K. H., Gallagher, M. W., Carl, J. R., y Barlow, D. H. (2014). Development and validation of the Overall Depression Severity and Impairment Scale. *Psychological assessment, 26*(3), 815.
- Bojorquez Chapela, L. y Salgado de Snyder, N. (2009). Características psicométricas de la Escala Center for Epidemiological Studies-depression (CES-D), versiones de 20 y 10 reactivos, en mujeres de una zona rural mexicana. *Salud mental, 32*(4), 299-307
- Bonicatto, S., Dew, A. M., y Soria, J. J. (1998). Analysis of the psychometric properties of the Spanish version of the Beck Depression Inventory in Argentina. *Psychiatry Research, 79*(3), 277-285.
- Bonilla, J., Bernal, G., Santos, A., y Santos, D. (2004). A revised Spanish version of the Beck Depression Inventory: Psychometric properties with a Puerto Rican sample of college students. *Journal of clinical psychology, 60*(1), 119-130.
- Camacho, A., Cordero, E. D., y Perkins, T. (2016). Psychometric properties of the DASS-21 among latina/o college students by the US-Mexico border. *Journal of immigrant and minority health, 18*(5), 1017-1023.
- Campo, A., Díaz, L. A., y Rueda, G. E. (2006). Validez de la escala breve de Zung para tamizaje del episodio depresivo mayor en la población general de Bucaramanga, Colombia. *Biomédica, 26*(3), 415-23.

- Campo-Arias, A., Díaz-Martínez, L. A., Rueda-Jaimes, G.E., Cadena, L. D. P., y Hernández, N. L. (2006). Validation of Zung's self-rating depression scale among the Colombian general population. *Social Behavior and Personality: an international journal*, 34(1), 87-94.
- Cano, J. F., Gomez Restrepo, C., y Rondón, M. (2016). Validación en Colombia del instrumento para evaluación de la depresión Montgomery-Åsberg Depression Rating Scale (MADRS). *Revista colombiana de psiquiatria*, 45(3), 146-155.
- Daza, P., Novy, D. M., Stanley, M. A., y Averill, P. (2002). The depression anxiety stress scale-21: Spanish translation and validation with a Hispanic sample. *Journal of Psychopathology and Behavioral Assessment*, 24(3), 195-205.
- De La Garza, N., John Rush, A., Grannemann, B. D., y Trivedi, M. H. (2017). Toward a very brief self-report to assess the core symptoms of depression (VQIDS-SR5). *Acta Psychiatrica Scandinavica*, 135(6), 548-553.
- Dominguez-Lara, S., y de Porres, M. (2016). Análisis estructural del Inventario de Depresión Estado-Rasgo en pacientes con diagnóstico de depresión de Lima. *Revista del Hospital Psiquiátrico de la Habana*, 13(1).
- García, A. F., Berzins, T., Acosta, M., Pirani, S., y Osman, A. (2017). The anxiety depression distress inventory-27 (ADDI-27): New evidence of factor structure, item-level measurement invariance, and validity. *Journal of personality assessment*, 100(3), 321-332.
- Gili, M., Lopez-Navarro, E., Homar, C., Castro, A., García-Toro, M., Llobera, J., y Roca, M. (2014). Propiedades psicométricas de la versión española de la escala QIDS-SR 16 en pacientes con trastorno depresivo. *Actas Espanolas de Psiquiatria*, 42(6), 292-299
- Gili, M., Luciano, J. V., Bauzá, N., Aguado, J., Serrano, M. J., Armengol, S., y Roca, M. (2011). Psychometric properties of the IDS-SR 30 for the assessment of depressive symptoms in Spanish population. *BMC medical research methodology*, 11(1), 131.
- González, D. A., Rodríguez, A. R., y Reyes-Lagunes, I. (2015). Adaptation of the BDI-II in Mexico. *Salud mental*, 38(4), 237-244.

- González, D. A., y Jenkins, S. R. (2014). Cross-measure equivalence and communicability in the assessment of depression: A focus on factor-based scales. *Assessment, 21*(6), 731-741.
- González, M., y Ibáñez, I. (2018). Propiedades psicométricas de una versión española breve de 30 ítems del Cuestionario de Ansiedad y Depresión (MASQE30). *Universitas Psychologica, 17*(1), 104-113.
- González, P., Nuñez, A., Merz, E., Brintz, C., Weitzman, O., Camacho, A., Penedo, F.J., Perreira, K., Choca, J., Merz, E., Navas, E. L., Buelna, C., Wassertheil-Smoller, S., Isasi, C.R., Talavera, G.A. y Gallo, L.C. (2017). Measurement properties of the Center for Epidemiologic Studies Depression Scale (ces-d 10): Findings from Hchs/sol. *Psychological assessment, 29*(4), 372.
- Herrero, J., y Gracia, E. (2007). Una medida breve de la sintomatología depresiva (CESD-7). *Salud mental, 30*(5), 40-46.
- Lasa, L., Ayuso-Mateos, J. L., Vazquez-Barquero, J. L., Díez-Manrique, F. J., y Dowrick, C. F. (2000). The use of the Beck Depression Inventory to screen for depression in the general population: a preliminary analysis. *Journal of affective disorders, 57*, 261-265.
- Lobo, A., Chamorro, L., Luque, A., Dal-Ré, R., Badia, X., Baró, E., y Grupo de Validación en Español de Escalas Psicométricas (GVEEP. (2002). Validación de las versiones en español de la Montgomery-Asberg Depression Rating Scale y la Hamilton Anxiety Rating Scale para la evaluación de la depresión y de la ansiedad. *Medicina clínica, 118*(13), 493-499.
- Londoño, C., Peñate, W., y González, M. (2016). Validación diferencial y discriminante del Cuestionario de Depresión para Hombres (CDH). *Terapia psicológica, 34*(2), 129-142.
- Osman, A., Wong, J. L., Bagge, C. L., Freedenthal, S., Gutierrez, P. M., y Lozano, G. (2012). The depression anxiety stress Scales—21 (DASS-21): further examination of dimensions, scale reliability, and correlates. *Journal of clinical psychology, 68*(12), 1322-1338.
- Polaino, A., y Senra, C. (1991). Measurement of depression: Comparison between self-reports and clinical assessments of depressed outpatients. *Journal of psychopathology and Behavioral Assessment, 13*(4), 313-324.

- Porter, E., Chambless, D. L., McCarthy, K. S., DeRubeis, R. J., Sharpless, B. A., Barrett, M. S., Milrod, B., Hollon, S. y Barber, J. P. (2017). Psychometric properties of the reconstructed Hamilton depression and anxiety scales. *The Journal of nervous and mental disease*, 205(8), 656.
- Rivera-Medina, C. L., Caraballo, J. N., Rodríguez-Cordero, E. R., Bernal, G., y Dávila-Marrero, E. (2010). Factor structure of the CES-D and measurement invariance across gender for low-income Puerto Ricans in a probability sample. *Journal of Consulting and Clinical Psychology*, 78(3), 398.
- Rizvi, S. J., Quilty, L. C., Sproule, B. A., Cyriac, A., Bagby, R. M., y Kennedy, S. H. (2015). Development and validation of the Dimensional Anhedonia Rating Scale (DARS) in a community sample and individuals with major depression. *Psychiatry research*, 229(1-2), 109-119.
- Ruiz, F. J., García Martín, M. B., Suárez Falcón, J. C. y Odriozola González, P. (2017). The hierarchical factor structure of the Spanish version of Depression Anxiety and Stress Scale-21. *International Journal of Psychology and Psychological Therapy*, 17(1), 97-105.
- Ruiz-Grosso, P., Loret de Mola, C., Vega-Dienstmaier, J. M., Arevalo, J. M., Chavez, K., Vilela, A., Lazo, M. y Huapaya, J. (2012). Validation of the spanish center for epidemiological studies depression and zung self-rating depression scales: a comparative validation study. *PloS one*, 7(10)
- Salinas-Rodríguez, A., Manrique-Espinoza, B., Acosta-Castillo, I., Téllez-Rojo, M. M., Franco-Núñez, A., Gutiérrez-Robledo, L. M., y Sosa-Ortiz, A. L. (2013). Validación de un punto de corte para la Escala de Depresión del Centro de Estudios Epidemiológicos, versión abreviada (CESD-7). *Salud pública de México*, 55(3), 267-274.
- Sanz, J., y García-Vera, M. P. (2007). Análisis psicométrico de las versiones breves del «Inventario para la Depresión de Beck» de 1978 (BDI-IA). *Psicología Conductual*, 15(2), 191-214.

- Senín-Calderón, C., Perona-Garcelán, S., Ruíz-Veguilla, M., y Rodríguez-Testal, J. F. (2017). Leiden Index of Depression Sensitivity-Revised (LEIDS-R): Spanish validation proposal. *International journal of clinical and health psychology*, 17(2), 139-150.
- Soria, V., Vives, M., Martínez-Amorós, È., Gálvez, V., Monzón, S., Crespo, J. M., Gili, M., Menchón, J.M., Roca, M., Parker, G. y Urretavizcaya, M. (2016). The CORE system for sub-typing melancholic depression: Adaptation and psychometric properties of the Spanish version. *Psychiatry research*, 239, 179-183.
- Talkovsky, A. M., y Norton, P. J. (2015). The Mood and Anxiety Symptom Questionnaire across four ethnoracial groups in an undergraduate sample. *American Journal of Orthopsychiatry*, 85(5), 431.
- Terrones-González, A., Estrada-Martínez, S., Lechuga-Quiñones, A. M., Salvador-Moysén, J., Martínez-López, Y., y La-Llave-León, O. (2012). Propiedades psicométricas de la CES-D/IS en población universitaria de la ciudad de Durango, México. *Salud mental*, 35(4), 305-313.
- Trivedi, M. H., Rush, A. J., Ibrahim, H. M., Carmody, T. J., Biggs, M. M., Suppes, T., Crismon, M.L., Shores-Wilson, K., Toprac, M.G., Dennehy, E.B., Witte, B. y Kashner, T.M. (2004). The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychological medicine*, 34(1), 73-82.
- Trujols, J., de Diego-Adeliño, J., Feliu-Soler, A., Iraurgi, I., Puigdemont, D., Álvarez, E., Pérez, V. y Portella, M. J. (2014). The Spanish version of the quick inventory of depressive symptomatology-self-report (QIDS-SR16): a psychometric analysis in a clinical sample. *Journal of affective disorders*, 169, 189-196.
- Trujols, J., Feliu-Soler, A., de Diego-Adeliño, J., Portella, M. J., Cebrià, Q., Soler, J., Puigdemont, D. Álvarez, E. y Perez, V. (2013). A psychometric analysis of the Clinically Useful Depression Outcome Scale (CUDOS) in Spanish patients. *Journal of affective disorders*, 151(3), 920-923.

- Tse, P.S., González, D. A., y Jenkins, S. R. (2018). Validating the structure of the depression and somatic symptoms scale. *Psychosomatics*, 59(3), 277-282.
- Valdés, C., Morales-Reyes, I., Pérez, J. C., Medellín, A., Rojas, G., y Krause, M. (2017). Propiedades psicométricas del Inventario de Depresión de Beck IA para la población chilena. *Revista médica de Chile*, 145(8), 1005-1012.
- Vargas Terrez, B. E., Villamil Salcedo, V., Rodríguez Estrada, C., Pérez Romero, J., y Cortés Sotres, J. (2011). Validación de la escala Kessler 10 (K-10) en la detección de depresión y ansiedad en el primer nivel de atención. Propiedades psicométricas. *Salud mental*, 34(4), 323-331.
- Whisman, M. A., Judd, C. M., Whiteford, N. T., y Gelhorn, H. L. (2013). Measurement invariance of the Beck Depression Inventory–Second Edition (BDI-II) across gender, race, and ethnicity in college students. *Assessment*, 20(4), 419-428.
- Williams, V., Francois, C., Danchenko, N., Nelson, L., Williams, N., Yarr, S., DiBenedetti, D. y Lancon, C. (2016). Psychometric validation of the depression and family functioning scale. *Current medical research and opinion*, 32(4), 639-650.