

Depresión y manía

C. Vázquez Valverde y F. Jiménez Franco

INTRODUCCIÓN

En este capítulo se revisan algunos de los principales instrumentos empleados en la evaluación de la depresión y la manía, bien conduzcan a un diagnóstico categorial (una etiqueta psiquiátrica) o una cuantificación dimensional de ambos estados anímicos. Un escrutinio minucioso de estos instrumentos delata que prácticamente todos incluyen síntomas o signos que también están presentes en instrumentos de evaluación de la ansiedad, insatisfacción de estas emociones (Gotlib, 1984; Shaver y Brennan, 1991). Esta interconexión, a pesar de sus profundas repercusiones para construir adecuadas teorías y taxonomías de las emociones, es a menudo ignorada. De hecho, la idea de encontrar instrumentos «puros» de depresión, u otra emoción, es una quimera científica carente de sentido dado que el concepto de «depresión» está intrínsecamente relacionado con otros síntomas o síndromes adyacentes y, en términos logicomatemáticos, constituye un «conjunto borroso» (*fuzzy set*) (Rosch, 1978). Por ejemplo, en la evaluación de la depresión en ancianos se han desarrollado instrumentos que cuidadosamente excluyen síntomas físicos para evitar la contaminación de dichos síntomas con la evaluación de la depresión (Vázquez y Lozoya, 1994); esa exclusión forzada sería, en nuestra opinión, una típica muestra de una confusión conceptual muy extendida: la creencia de que existen prototipos ideales de emociones, independiente unos de otros y que, por lo tanto, pueden evaluarse de modo independiente.

Con esta revisión de instrumentos y procedimientos no se pretende ser exhaustivo. Existen diversas

fuentes de instrumentos de medida en psicopatología (fundamentalmente cuestionarios, escalas e inventarios) en los que se analizan y a veces se reproducen un extenso surtido de herramientas, incluyendo algunas de depresión, que el lector interesado puede consultar tanto en inglés (Beckham y Leber, 1985; Sartorius y Ban, 1986; Thompson, 1989; Peck y Shapiro, 1990; Shaver y cols., 1991) como en castellano (Conde y Franch, 1984; Fischer y Corcoran, 1987; Fernández-Ballesteros, 1992a; Comeche y cols., 1995; Roa, 1995). Por último, además del repaso de los instrumentos dirigidos a la valoración sintomática de la depresión y la manía, se efectúa una breve descripción de instrumentos evaluadores de importantes variables ligadas a los trastornos anímicos (p. ej., ideas irracionales, inventarios de refuerzos, pensamientos automáticos, etc.).

DILEMAS EN LA EVALUACIÓN DE LAS ALTERACIONES ANÍMICAS

¿Síntomas, síndromes o categorías?

El punto de partida esencial para comprender la evaluación psicopatológica es establecer claramente la distinción entre síntoma, síndrome y cuadro clínico. Estos tres términos denotan diferentes unidades de análisis de la psicopatología (Vázquez, 1991) y la medida de cada uno de ellos puede requerir diferentes aproximaciones de medida e instrumentos diferentes.

El término «depresión» a veces se emplea simplemente para denotar un estado de ánimo bajo, alicaido. Esto es sencillamente un síntoma anímico que está

presente en muchos momentos de la vida de una persona normal y también en la mayor parte de los trastornos mentales (ansiedad, esquizofrenia, anorexia, etc.). En un sentido más estricto, la depresión puede entenderse como un síndrome o conjunto de síntomas covariantes: baja autoestima, desesperanza, apatía, desinterés por las cosas o por los demás, dolores corporales, etc. Muchas de las escalas que evalúan la depresión adoptan una perspectiva sindrómica, pues valoran un amplio espectro de síntomas potenciales. Por último, en un sentido aún más restrictivo, la depresión es un cuadro clínico, una categoría diagnóstica; en este sentido, existen determinados sistemas de clasificación (como, por ejemplo, el DSM-IV) que denominan «depresiva» sólo a aquella persona que cumple un patrón determinado de signos y síntomas pero además cumple otras condiciones (p. ej., problemas de adaptación, ausencia de trastornos esquizofrénicos y delirantes, ausencia de factores causales orgánicos, etc.). Una conclusión crucial de esta distinción es que para el diagnóstico de la depresión es deseable emplear tanto una perspectiva categorial o clasificatoria, como una perspectiva dimensional. En definitiva, es necesario conocer, respectivamente, a qué categoría diagnóstica pueden pertenecer los síntomas que manifiesta el sujeto y, por otro lado, poder cuantificar (normalmente a través de las escalas) la gravedad de los síntomas presentes.

No hay que olvidar, por lo tanto, que las escalas y otros instrumentos cuantificadores de la gravedad sirven de forma exclusiva para identificar síntomas y valorar su intensidad. Sin embargo, el diagnóstico categorial se efectúa teniendo en cuenta no sólo el número y tipo de síntomas presentes sino otros factores como, por ejemplo, la ausencia de otros cuadros clínicos potencialmente presentes (p. ej., esquizofrenia, intoxicación por drogas, etc.), la duración de los síntomas, determinadas combinaciones de inclusión y exclusión, etc. Esta distinción tiene implicaciones sustanciales en ámbitos como la epidemiología. Como ha quedado demostrado reiteradamente el empleo de un sistema de diagnóstico dimensional (empleando puntos diagnósticos de corte) suele conllevar criterios más laxos y, por consiguiente, puede inflar artificialmente las cifras de prevalencia de los trastornos afectivos en particular (Compas y cols., 1993).

¿Categorías o dimensiones?

Éste es un dilema que conlleva importantes decisiones sobre qué tipo de medida utilizar e implícitamente incorpora una serie importante de supuestos epistemológicos (p. ej., ¿existen realmente las categorías diagnósticas?, ¿son meras convenciones necesari-

as para crear un lenguaje científico?, etc.). La pregunta diagnóstica crucial ante una persona con un estado anímico depresivo es si constituye un cuadro clínico depresivo. ¿Es un trastorno depresivo o es otra categoría diagnóstica? Esta primera aproximación la efectúa todo clínico sea cual sea su orientación e incluso en aquellos clínicos que nieguen efectuar «diagnósticos» en su práctica cotidiana. La respuesta a esta pregunta debe provenir del empleo de algún sistema de clasificación categorial o, en otras palabras, de etiquetas diagnósticas.

Para la evaluación categorial es deseable proporcionar un diagnóstico de DSM, Criterios Diagnósticos de Investigación (Research Diagnostic Criteria, RDC), CIE-10 o, incluso, puede ser aconsejable efectuar un diagnóstico siguiendo las pautas de dos o más de estos sistemas dado que no son absolutamente coincidentes. La evaluación categorial de los trastornos afectivos es recomendable aun reconociendo las múltiples arbitrariedades que conlleva una formalización diagnóstica de este tipo (Wetzler y Van Praag, 1989). Estas arbitrariedades se manifiestan en la inclusión de unos criterios y no otros (p. ej., en el DSM-IV no figuran síntomas como «llorar»), en la consideración de puntos de corte predeterminados (p. ej., requerir la presencia de síntomas al menos durante 2 semanas), y en la presencia de una serie mínima de síntomas sobre un conjunto de los mismos (p. ej., la presencia de al menos cinco de un conjunto de nueve: ¿por qué no seis o cuatro?). Por otro lado, en los sistemas de clasificación operativos (como los de la APA) no se considera que ningún síntoma sea más importante que otro; es decir, es un tipo de clasificación horizontal que no otorga más peso a la presencia de determinados síntomas. Todos los síntomas, en caso de aparecer, reciben el mismo peso diagnóstico.

Dejando a un lado estas consideraciones de fondo, el diagnóstico categorial puede efectuarse simplemente siguiendo las indicaciones diagnósticas de estos sistemas de clasificación mediante un sencillo listado de síntomas (*checklist*) que cubra los criterios expuestos en los mismos, o bien cuando se trate de protocolos de investigación más rigurosos, empleando algunas de las entrevistas estructuradas que conducen a este tipo de diagnóstico. Además de esta evaluación categorial, es deseable incluir una evaluación dimensional de los síntomas y signos observables. Es decir, dado que se supone que la sintomatología depresiva varía en intensidad a lo largo de un continuo, es conveniente valorar la gravedad de los síntomas. Hay que tener en cuenta, sin embargo, que un diagnóstico categorial apenas tiene valor para el diseño terapéutico y para su seguimiento.

Como puede apreciarse, valorar un cuadro afectivo no es sencillo. No basta sólo emplear una escala de síntomas sino que se requiere el uso de unos criterios

diagnósticos precisos que permitan efectuar un diagnóstico categorial adecuado. Ahora bien, una vez establecido el diagnóstico, resulta pertinente el uso de medidas para cuantificar la gravedad de la sintomatología (Vázquez, 1986 y 1995; Vázquez y Lozoya, 1994) y a esto se dedica buena parte del capítulo.

EVALUACIÓN CATEGORIAL

Pauta de entrevista diagnóstica del trastorno depresivo mayor (DSM-IV)

La herramienta diagnóstica y de evaluación fundamental es la entrevista, aunque no siempre sea fácil obtener información de una persona deprimida. La habitual falta de interés y de motivación del paciente deprimido también puede extenderse a la entrevista, haciendo a veces frustrante la tarea. Por otro lado, hay que ser cuidadosos en la valoración de la información puesto que si bien los recuerdos biográficos de los pacientes psiquiátricos son mucho más fiables de lo que se ha pensado hasta ahora (Brewin y cols., 1993), en los deprimidos pueden operar algunos sesgos cognitivos y emocionales que pueden teñir de forma negativa la información que proporcionan: como se ha demostrado recientemente en un metaanálisis sobre memoria y estado de ánimo, tienden a recordar mejor lo negativo que lo positivo (Matt y cols., 1992).

La exploración del paciente ha de centrarse no sólo en episodios actuales sino también pasados, y naturalmente hay que explorar si hay o ha habido algún episodio de manía, en cuyo caso se clasificaría automáticamente como trastorno bipolar, sea cual sea su estado de ánimo actual (APA, 1994).

El instrumento fundamental de exploración de un paciente es la entrevista en cualquiera de sus formatos, desde las más estructuradas a las más abiertas. En cualquier caso, una entrevista con fines diagnósticos debe conllevar un guiño bien determinado. Desde nuestro punto de vista, la aproximación más razonable es seguir de cerca los criterios tipo DSM-IV (APA, 1994), independientemente de que también puedan valorarse otros síntomas (p. ej., llanto, relaciones interpersonales deterioradas, etc.) que no están presentes en estos criterios. A continuación se exponen una serie de consejos para una sencilla exploración diagnóstica siguiendo los criterios DSM-IV (Leber y cols., 1985; Vázquez, 1995). Naturalmente, además de, o en vez de, estas sencillas pautas clínicas podrían emplearse instrumentos más sofisticados como la entrevista SADS, la SCID, o la CIDI, todas ellas conducentes a diagnósticos con criterios operativos (RDC, DSM-III-R, o ICD-10) (véase más adelante).

A continuación nos ceñiremos a un breve análisis de cómo plantear la entrevista diagnóstica del trastorno depresivo mayor. Para ello, se examina cómo valorar la presencia de un episodio depresivo mayor para después analizar cómo valorar la posible existencia de un trastorno depresivo mayor.

Para la definición de episodio depresivo mayor el DSM-IV considera necesaria la presencia durante al menos 2 semanas de: a) un estado de ánimo bajo y constante durante la mayor parte del día, o b) una disminución de placer o de interés en las actividades cotidianas. Pero además de cumplir uno de estos dos criterios, se requiere también la presencia de cuatro de los siguientes síntomas:

1. *Cambios de peso y/o apetito.* La pérdida de peso y/o apetito es un síntoma tradicionalmente considerado característico de la depresión. Aunque hay más discusión sobre si el exceso de peso y/o apetito es también característico, el DSM-IV también incluye esta posibilidad. El criterio indicado en el DSM-IV es un aumento o disminución de peso del 5% sobre el peso habitual. A veces es difícil valorar este síntoma porque puede que el paciente:

- Tenga poco apetito para la comida pero coma muchos dulces o tenga estereotipias alimentarias.
- Tenga una enfermedad física o medicación que afecte a su apetito o su peso.
- Coma tanto como siempre pero diga no disfrutar nada con la comida.
- Tenga un apetito con muchos alibajos.

En todos estos casos, es conveniente preguntar sobre dos cuestiones: a) ¿es esta condición diferente a lo que sucede normalmente en esa persona?, y b) ¿lo vive el paciente como un problema? Si la respuesta a alguna de estas preguntas es afirmativa, debe considerarse que el síntoma está presente.

2. *Insomnio o hipersomnio.* Como sucede con el síntoma anterior, el insomnio se ha considerado siempre como un síntoma depresivo, pero con el hipersomnio hay más dudas. En cualquier caso, se debe explorar tanto el insomnio inicial como el medio, y el terminal, y analizar cómo es el estado de somnolencia del paciente durante las 24 horas (Vázquez, 1990c). Un criterio habitual para el insomnio inicial es tardar más de 30 minutos en quedarse dormido. El insomnio medio describe la situación de despertarse más de 30 minutos durante la noche con dificultades para volver a dormir. El insomnio tardío debe valorarse como existente cuando el paciente dice despertarse, sin poder volver a dormir, entre 1 y 3 horas antes de la hora que era habitual. Por el contrario, no existe un criterio claro para diagnosticar hipersomnio. Aunque lo que habitual-

mente se entiende por hipersomnio es la somnolencia diurna, en entrevistas como la Schedule of Affective Disorders and Schizophrenia (SADS), que se examina posteriormente se considera la presencia de hipersomnio si el paciente duerme al menos 1 hora más de lo que en él era habitual.

El diagnóstico de los problemas del sueño puede complicarse cuando el paciente está bajo medicación antidepressiva u otra, tiene dolores, turnos de trabajo rotatorios, etc. Si el paciente ha mejorado su sueño por efectos de la medicación específica (p. ej., somníferos), la entrevista diagnóstica SADS sugiere que se evalúe este ítem en función a cómo considera el paciente que sería su sueño si no estuviese medicado. Pero, como ocurre con el síntoma del apetito, cuando la decisión sea complicada, suele ser útil valorar si el patrón de sueño es diferente al de antes de estar deprimido y/o si esto resulta un problema para el paciente.

3. *Retardo o agitación.* No basta con que el paciente se sienta subjetivamente más agitado, nervioso, o lento. Su diagnóstico requiere alguna manifestación conductual que sea visible para los demás. Normalmente el retardo se expresa en un enlentecimiento del habla o de los movimientos. La agitación suele manifestarse en locuacidad verbal, a veces en hablar a gritos, no poderse estar quieto (deambular, frotarse las manos, tirarse del pelo mientras se está sentado), etc. Estas características pueden ser tan pronunciadas que a veces emplean los términos de depresión retardada o depresión agitada, respectivamente. Aunque a menudo la información obtenida durante la entrevista (incluido el comportamiento del paciente durante la misma) puede bastar para valorar el síntoma, puede ser necesario requerir información de familiares o amigos.

4. *Fatiga o pérdida de energía.* Es necesario hacer preguntas para distinguir entre pérdida de energía y pérdida de interés pues son dos síntomas diferentes. La falta de energía puede a menudo confundirse con una disminución en el interés. Se requiere hacer preguntas sobre si se cansa al hacer las cosas, o si no tiene fuerzas para acabar las tareas normales.

5. *Sentimientos excesivos de falta de valía, autorreproches o sentimiento de culpabilidad.* Para la valoración de la culpabilidad el clínico ha de situarla en el contexto cultural, religioso y social en el que se desenvuelve el paciente. El requisito general del DSM-IV de que los síntomas estén presentes a diario o casi a diario durante 2 semanas, ayuda a tomar una decisión en este sentido.

6. *Problemas de concentración o toma de decisiones.* A veces se ha considerado que esto es parte del retraso psicomotor, pero en el DSM-IV se considera aparte. Las preguntas deben explorar si el paciente es capaz de seguir una conversación o programas de televisión,

concentrarse en el trabajo, etc. Hay que hacer entender al paciente con nuestras preguntas que una cosa es tener interés en algo y otra, que es lo que se evalúa con este ítem, si es capaz de concentrarse en eso.

7. *Pensamientos recurrentes de suicidio o muerte.* Obviamente los planes o intentos de suicidio durante las dos pasadas semanas, aunque no hayan estado presentes a diario, deben contabilizarse como presencia del síntoma.

Cumplir estos síntomas no supone el diagnóstico automático de la presencia de un episodio depresivo mayor. Para ello se requiere satisfacer otra serie de requisitos adicionales (ausencia de una condición médica causante del episodio, mal funcionamiento psicossocial y ausencia de reacción de duelo). Pero, ni aún cumpliendo todos estos criterios de episodio depresivo mayor nos aseguráramos finalmente la presencia de un trastorno depresivo mayor diagnosticable. Para poder llegar a esta categoría diagnóstica a su vez hay que descartar una serie de condiciones excluyentes (manía, esquizofrenias y trastornos esquizoafectivos) (APA, 1994).

Naturalmente, para la decisión final sobre si alguno de estos ítems es diagnosticable o no en un paciente determinado, puede necesitarse recurrir a información complementaria derivada de escalas (p. ej., escalas de autoestima, desesperanza, etc.) u otros procedimientos (p. ej., información de familiares). Así mismo, la entrevista no se debe reducir a la exploración y valoración de los síntomas por su valor diagnóstico. Una entrevista característica con un paciente debe incluir el análisis inicial de otras áreas (empleo de tiempo, relaciones interpersonales, apoyo social, nivel de estrés, etc.).

Entrevistas estructurales diagnósticas

Como es bien sabido, las entrevistas estructuradas se desarrollaron a partir de la década de los setenta con el fin de ayudar a solucionar, junto con el uso de criterios diagnósticos operativos [criterios Feighner, Criterios Diagnósticos de Investigación (Research Diagnostic Criteria, RDC), criterios DSM-III, etc.] el inaceptable acuerdo interexaminador de la inmensa mayoría de las categorías diagnósticas psiquiátricas (Vázquez, 1990a y b). Dados los sesgos cognitivos y la práctica habitual del médico, el empleo de estos instrumentos formales de recogida de datos ha demostrado ser más útil que las entrevistas abiertas o semiestructuradas para recoger de modo sistemático, ordenado y eficaz, información del paciente o de sujetos en estudios comunitarios. Incluso un síntoma tan aparentemente llamativo como las ideas de suicidio,

se explora tres veces menos en las historias clínicas tradicionales que cuando se emplean entrevistas estructuradas (Robins y Helzer, 1986).

Estas entrevistas proporcionan una detallada información psicopatológica del paciente, incrementan la fiabilidad del diagnóstico efectuado, permiten la comparación de los resultados con los de otros estudios rigurosos publicados, y pueden proporcionar diagnósticos categoriales formales. Así mismo, en estudios epidemiológicos rigurosos puede ser de gran utilidad su empleo para poder establecer diagnósticos categoriales comparables con otros estudios. Las entrevistas estructuradas se han convertido en una herramienta realmente imprescindible en la investigación básica y epidemiológica (Vázquez Barquero, 1993; Rodríguez-Pulido y Montes de Oca, 1993). En el ámbito de los trastornos del estado de ánimo se señalan algunas de las entrevistas estructuradas más relevantes conducentes a diagnósticos categoriales (Sholomskas, 1993).

La irrupción de entrevistas estructuradas en el arsenal diagnóstico puede hacer que nos preguntemos con ironía si, después de todo, «son necesarios los médicos» (Spitzer, 1984). Aunque algunas están diseñadas de tal modo que incluso personas no expertas puedan administrarlas (Diagnostic Interview Schedule, DIS; o Composite International Diagnostic Interview CIDI), obteniéndose índices de fiabilidad satisfactorios (Helzer y cols., 1985), muchas otras siguen requiriendo efectuar juicios clínicos sobre la presencia o no del síntoma y sobre su gravedad como el Present State Examination (PSE), la Structural Clinical Interview DSM-III (SCID) o la SADS. Además, las entrevistas estructuradas no son tan útiles para apresar aspectos tan importantes como el estilo de interacción personal del paciente, sus actitudes, la detección de signos no verbales, etc., todos ellos inevitablemente presentes durante una entrevista (Zubin, 1989) y, en muchos trastornos, necesarios para formular un diagnóstico.

Schedule of Affective Disorders and Schizophrenia

La SADS fue un instrumento auspiciado por el National Institute of Mental Health norteamericano en la década de los setenta, con el fin de poder lograr altos índices de fiabilidad interexaminadores en protocolos de investigación (Endicott, 1978; Endicott y Spitzer, 1979). Se trata de una compleja entrevista estructurada destinada al diagnóstico, en muestras de pacientes, de trastornos afectivos y esquizofrenia. La cumplimentación de la SADS es larga (entre 1,5 y 2 horas), siempre la efectúan clínicos y requiere un entrenamiento especial.

Cada ítem está definido de un modo muy preciso, se indica exactamente la pregunta que hay que efectuar

y se señala cómo hay que valorar la gravedad de cada respuesta, por lo general en una escala de 7 puntos. Además del diagnóstico, la SADS ofrece información sobre el curso, el momento de aparición, número de episodios, etc. Existen tres versiones con fines diferentes: la SADS normal, la SADS-Life (que informa sobre episodios actuales y pasados), y la SADS-Change (que informa sobre cambios habidos durante la última semana) que cubre los problemas psicopatológicos habidos ahora y en el pasado. La SADS cubre las 25 categorías diagnósticas de los RDC, en los que se inspiraron directamente los criterios DSM-III (Vázquez, 1990a y b) y, en concreto, cubre 18 síndromes o categorías relacionadas con trastornos afectivos. Posee buenas cifras de fiabilidad interexaminadores (Sholomskas, 1993) y su uso ha sido muy extendido en protocolos rigurosos de investigación (p. ej., Vázquez y cols., 1986).

Structural Clinical Interview for DSM-III

La SCID es una entrevista estructurada de formato casi idéntico a la SADS pero cubre muchos más cuadros que ésta (Spitzer y cols., 1987). Permite obtener un diagnóstico DSM-III-R de diversas entidades diagnósticas, a través de módulos de entrevista independientes: trastornos del estado de ánimo, trastornos psicóticos, trastornos por uso de sustancias psicoactivas, trastornos de ansiedad, trastornos somatoformes, trastornos de la alimentación y trastornos de ajuste. Existen tres versiones: una para pacientes internos (SCID-P), otra para pacientes ambulatorios (SCID-OP), y otra para no pacientes (SCID-NP). Su duración media es de 60 minutos pero depende mucho de la enfermedad del paciente.

La SCID evalúa el estado clínico durante el mes pasado y en cualquier momento anterior. Al igual que la SADS, requiere un juicio clínico constante por parte de un entrevistador experto y quizá su gran ventaja es su amplia cobertura diagnóstica en términos DSM-III-R (APA, 1987).

Diagnostic Interview Schedule

La importancia del DIS reside en que fue el instrumento diagnóstico del gigantesco estudio Epidemiological Catchment Area (ECA) patrocinado por el National Institute of Mental Health (Burke y Regier, 1989). El DIS fue la primera entrevista estructurada que, cubriendo una amplia gama de categorías diagnósticas, se desarrolló para poder ser aplicada por entrevistadores legos (Vázquez, 1990a). El DIS permite formular 43 diagnósticos del DSM-III e incluye ade-

más la posibilidad de obtener diagnósticos siguiendo los criterios de Feighner y casi todos los de los RDC. En términos de cobertura específica de los trastornos del estado de ánimo, no ofrece tantos cuadros como la SADS. En términos DSM-III, cubre depresión mayor, distimia, trastornos bipolares y episodio maniaco. El tiempo medio de cumplimentación es de unos 45-90 minutos, dependiendo de la enfermedad y el estado del entrevistado. Evalúa no sólo síntomas actuales sino también los de las 2 semanas anteriores, el mes anterior, los 6 meses anteriores y todo el año pasado.

Como entrevista, el DIS tiene una estructura compleja de decisiones internas a modo de algoritmos de decisión y requiere un adecuado entrenamiento. Al formular preguntas muy cerradas y requerir básicamente respuestas de formato sí o no, no requiere el juicio clínico del entrevistador. Esto último supone una enorme ventaja respecto a la SADS o a la SCID pues, además de su uso en contextos clínicos, permite efectuar investigaciones epidemiológicas a gran escala sin necesidad de emplear personas con experiencia clínica.

Composite International Diagnostic Interview

La CIDI surgió como un producto de la colaboración de la OMS y la ADAM-HA. La CIDI es resultante de una combinación del DIS con el PSE, haciendo más énfasis que entrevistas anteriores en diferencias y peculiaridades psicopatológicas transculturales. A través de algoritmos computacionales, permite proporcionar opcionalmente diagnósticos del tipo de la OMS (CIE-10) y de la APA (DSM-III-R). Al igual que el DIS es una entrevista que pueden realizar no expertos tras un período de entrenamiento. Una aplicación de la versión oficial de la CIDI en castellano (Bravo y cols., 1991; Rubio-Stipec y cols., 1991) se puede examinar en Muñoz y cols. (1995). Una de las peculiaridades de la CIDI es que en cada síntoma existente se sondea si es de origen orgánico (en cuyo caso no computaría para determinar la presencia del trastorno mental bajo examen) o de origen mental.

En términos de trastornos del estado de ánimo, dentro del esquema DSM-III-R, cubre trastorno depresivo mayor, distimia, trastorno bipolar y trastorno bipolar atípico (bipolar II). Además, desde una perspec-

TABLA 21-1. Características de los instrumentos más usados de evaluación general de la depresión

	SADS	HDRS	CRS
Estilo	Entrevista	Heteroaplicado	Autoaplicado
Número de ítems	Más de 200 preguntas	17 + 4	52
Duración	60-120 minutos	20 minutos	20 minutos
Escala	Síntomas presentes o ausentes	Gravedad (0-2; 0-4)	Síntomas presentes o ausentes (S/No)
Marco temporal	Episodios actuales o pasados	Última semana (o días)	Últimos días
Referencia	Endicott (1978)	Hamilton (1960, 1967)	Carroll y cols. (1981)
Ventajas	Uso de criterios operativos diagnósticos (DSM, RDC)	Muy utilizada, alta fiabilidad interexaminadores, buena sensibilidad	Autoaplicada basada sólo en la HDRS Fácil cumplimentación
Inconvenientes	Larga, exige entrenamiento, validez concurrente moderada	Consistencia no elevada, peso excesivo a síntomas somáticos, poco uso con población no clínica	Consistencia no elevada. Mediocre correspondencia con la HDRS

BDI: Inventario de Depresión de Beck; CFS-D: Escala de Depresión del Centro de Estudios Epidemiológicos; CRS: Escala de Evaluación de Carroll; HDRS: Escala de Depresión de Hamilton; MMPI-D: Inventario Multifásico de Personalidad de Minnesota; SADS: Escala de Trastornos Afectivos y Esquizofrenia; SDS: Escala Autoaplicada de Depresión

tiva sindrómica permite también evaluar presencia de episodio depresivo mayor (melancólico o no) y episodio maníaco.

Tanto el DIS como la CIDI están sometidas a una fuerte controversia sobre su validez, en especial para los cuadros más importantes (p. ej., esquizofrenias o trastorno obsesivo-compulsivo) y de menor aparición en la población general (Sholomskas, 1993).

EVALUACIÓN DIMENSIONAL

Heteroaplicación o autoaplicación

Detectar la presencia de un síntoma psicopatológico determinado no es muy difícil. Como sagazmente advirtiera Hamilton (1987), el verdadero problema se plantea cuando se trata de evaluar su gravedad y su impacto en el ajuste vital del paciente. Aquí es donde las escalas, sean de observación o de autoaplicación, ofrecen su principal aportación. Casi todas evalúan la gravedad de los síntomas en varios puntos (normal-

mente utilizando entre 3, 4 o 5 niveles de intensidad). Aparte de los contenidos específicos de cada escala, no todas emplean los mismos parámetros de medida, y esto es algo que hay que considerar de forma cuidadosa al seleccionarlas para su uso, como se comenta más adelante: algunas emplean un criterio de duración de síntomas, otras de frecuencia y, finalmente, otras emplean criterios de intensidad (tabla 21-1).

Determinar la validez de una escala es siempre difícil porque en psiquiatría y en psicología clínica no se dispone de criterios dorados (*golden criteria*) de validación. Es decir, no existen criterios diagnósticos irrefutables (p. ej., pruebas radiológicas, pruebas bioquímicas, etc.) que permitan contrastar la validez de un instrumento puesto que los conceptos clínicos (esquizofrenia, depresión, ansiedad...) son en sí mismos construcciones teóricas más o menos acertadas que se proponen por la comunidad científica en un momento determinado (Seeman, 1983). Por eso, la tarea de validación es siempre aproximativa y acumulativa (Silva, 1989). Así pues, es epistemológicamente imposible obtener una escala o un instrumento diagnóstico absolutamente válido puesto que en psicopatología se ca-

BDI	MMPI-D	SDS	GES-D
Autoaplicado	Autoaplicado	Autoaplicado	Autoaplicado
21	60	20	20
15-20 minutos	15-20 minutos	10-15 minutos	10-15 minutos
Alternativas de diferente gravedad (0-3)	Verdadero-falso. Síntomas o características asociadas	Duración (0-3)	Frecuencia (0-3)
Última semana	Actitudes, conductas u opiniones que uno suele hacer «por lo general»	Última semana	Última semana
Beck y cols. (1961); Beck (1978)	Hathaway y McKinley (1942)	Zung (1965)	Radloff (1977)
Cumplimentación fácil, sensible al cambio, muy utilizada, buena discriminación	Sensible a cambios de humor	Buena validez concurrente con otras escalas, sensible a cambios	Adecuada para población general. Buena consistencia y especificidad
Alta correlación con ansiedad	Ítems demasiado heterogéneos; carece de validez discriminante con ansiedad; validez y especificidad dudosas	Validez discriminante escasa	Baja sensibilidad. Centrada demasiado en estado de ánimo

De Hamilton y Shapiro, 1990; Shaw y cols., 1985; Vázquez, 1995.

rece de criterios de referencia precisos e incuestionables. Los procedimientos habituales para validar una escala psicopatológica (de ansiedad, depresión o de cualquier otro constructo) son:

1. Compararla con los juicios clínicos de expertos (normalmente en escalas Likert de 3 a 11 puntos).
2. Comprobar si se correlaciona bien con otras escalas semejantes (validez concurrente o convergente), procedimiento que no deja de ser algo circular.
3. Ver si es capaz de diferenciar grupos supuestamente diferentes en el constructo que se pretende medir (validez discriminante).

Hay dos grandes técnicas de evaluación: escalas de heteroaplicación y escalas de autoaplicación. En las primeras, es el médico quien evalúa la gravedad del sujeto, mientras que en las segundas es el propio sujeto quien lo hace. Ningún procedimiento es en sí mismo mejor que el otro y de hecho ofrecen resultados de eficacia similar (Lambert y cols., 1986). Las heteroaplicadas presentan la ventaja de que pueden aplicarse

en personas poco colaboradoras, con dificultades de concentración o con problemas para entender el lenguaje escrito. Además, son útiles cuando, por cualquier razón, no se quiere saturar a un sujeto con tests de papel y lápiz. Así mismo, pueden evaluar mejor que los instrumentos autoaplicados algunos síntomas de valoración más dificultosa: síntomas somáticos, psicomotores, grado de *insight*, etc. Por el contrario, las escalas autoaplicadas son muy útiles porque se pueden utilizar de modo repetido, proporcionan información sobre la propia visión del paciente, y son fáciles y económicas en general (tabla 21-2). Pero es difícil construir escalas autoaplicadas adecuadas de psicopatología. Por ejemplo, hay que hacer uso de palabras comunes, inteligibles para la mayoría, lo que puede reducir el valor clínico de la escala, ya que hay ciertos matices técnicos que pueden ser difíciles de transmitir a una persona no experta.

Uno de los problemas de los instrumentos heteroaplicados es que pueden tener cabida los sesgos del clínico. Por ejemplo, puede haber problemas como la presencia de efectos de halo por la impresión general que

TABLA 21-2. Comparación entre los procedimientos de autoinforme y los basados en entrevista en la evaluación de los trastornos afectivos

Áreas evaluadas y características generales	Instrumentos autoaplicados	Instrumentos heteroaplicados
Evaluación del estado anímico subjetivo	Adecuados	Menos adecuados
Evaluación de conductas observables	No	Sí
Evaluación del <i>insight</i>	No	Sí
Especificidad	Mayor	Menor
Rango de gravedad	Menos útiles con depresión grave, características psicóticas o gran entumecimiento motor	Utilizables en cualquier rango. Quizá menos sensibles para depresiones ligeras
Utilidad en bipolares	Poco adecuadas en episodios de manía	Útiles en manía
Circunstancias de uso	Requieren motivación, colaboración y capacidad de concentración del paciente	Pueden completarse con una cooperación mínima por parte del paciente
Coste	Mínimo	Mayor. Requieren la intervención de un profesional preparado
Sesgos potenciales	Típicos sesgos de respuesta de los tests de papel y lápiz (p. ej., aquiescencia, deseabilidad social). Sesgos especiales debidos al trastorno del paciente (p. ej., distorsiones de memoria)	Sesgos de respuesta y distorsiones por parte del entrevistador (p. ej., evitar dar puntuaciones elevadas, inferir la gravedad de un síntoma a partir de lo que se evalúa en otro)

De Paykel y Norton, 1986; Hamilton, 1987; Vázquez, 1995.

ofrece el paciente, evitar dar puntuaciones elevadas, inferir de forma inadecuada la gravedad de un síntoma a partir de la gravedad observada en otro, etc. Hamilton (1987) indica incluso sesgos en función de la experiencia del clínico: los más jóvenes e inexpertos tienden a dar puntuaciones de gravedad muy altas en las primeras sesiones y luego tienden a dar puntuaciones muy bajas en sesiones posteriores, por lo que aparentemente el paciente presenta amplias mejorías. En cuanto a los problemas de las autoaplicadas, pueden ser vulnerables a los sesgos característicos de los tests de papel y lápiz (sesgos de aquiescencia, deseabilidad social, sesgos de memoria debidos al estado anímico del sujeto, etc.). Pero, en definitiva, cada instrumento tiene ventajas e inconvenientes específicos (tablas 21-1 a 21-3).

Por último, es muy importante señalar que las escalas simplemente registran lo que es común a todos los pacientes, pero no la información idiosincrática necesaria para el tratamiento (Hamilton, 1987). Para esto se necesita recurrir a otros tipos de información complementaria (entrevista, medición de variables relacionadas, análisis conductual, etc.).

Instrumentos heteroaplicados de depresión

Escala de Depresión de Hamilton

La Hamilton Depression Rating Scale (HDRS) es la escala de heteroaplicación más utilizada. Publicada

en 1960 por Max Hamilton, la forma más utilizada es la revisada en 1967, de la que se excluyeron cuatro ítems y se proporcionó una mejor definición de los mismos. La escala original de Hamilton (1960) contenía 21 ítems al incluir cuatro ítems que no miden gravedad, pero que podrían ser utilizables en circunstancias especiales (p. ej., respuesta al tratamiento): variación diurna (18), despersonalización (19), pensamiento paranoide (20) y obsesionalidad (21).

El uso de esta escala requiere la evaluación por parte del clínico de la intensidad de los síntomas del paciente a través de la entrevista clínica habitual y de todos aquellos medios de obtención de información que el clínico estime oportuno (Hamilton, 1967 y 1987).

La HDRS contiene 17 ítems que evalúan la gravedad de los síntomas durante la pasada semana, si bien se puede enmarcar un período más corto (p. ej., los 2-3 últimos días). Ocho ítems se evalúan en una escala de 0 a 2 (0, ausente; 1, dudoso; 2, presente) y nueve en una escala de 0 a 4 (0, ausente; 1, trivial o dudoso; 2, ligero; 3, moderado; 4, grave). La HDRS incorpora ciertos aspectos que son difíciles de evaluar a través de instrumentos autoaplicados: conciencia de enfermedad (*insight*), entencimiento o agitación psicomotora, y presencia de ideas delirantes depresivas. No obstante, hay ciertos síntomas (p. ej., exceso de apetito o hipersomnio) que no se valoran en la HDRS. Una de las características de esta escala es que posee muchos ítems «somáticos», lo cual a veces se le ha criticado por excesivo (tabla 21-3).

TABLA 21-3. Porcentajes con que contribuye a la puntuación teórica total cada tipo de síntoma en diferentes instrumentos generales de depresión

Síntomas	BDI	HDRS CRS	CES-D	SDS	Beck-Rafaelson	MADRS
Estado de ánimo	9	8	40	15	18	30
Vegetativos	29	28	10	35	18	30
Motores	0	12	10	5	18	0
Sociales	5	8	15	0	19	0
Cognitivos	52	28	20	35	27	30
Ansiedad	0	16	0	5	9	10
Irritabilidad	5	0	5	5	0	0

Estado de ánimo: incluye tristeza, falta de alegría, cualidad distinta del estado de ánimo, lloro y variación diurna

Vegetativos: incluye trastorno de sueño, cambio en apetito, cambio de peso, pérdida de libido, estreñimiento y fatiga

Motores: incluye agitación, entencimiento e inquietud motora

Sociales: incluye aislamiento, apartamiento e incapacidad de funcionar en el trabajo o en otras tareas

Cognitivos: incluye pensamientos de desesperanza e indefensión, de suicidio, de enfermedad, de culpa, así como falta de conciencia de enfermedad e indecisión

Ansiedad: incluye ansiedad psíquica, somática y fóbica

Irritabilidad: incluye tanto hostilidad dirigida hacia uno mismo como hacia los demás

BDI: Inventario de Depresión de Beck; CES-D: Escala de Depresión del Centro de Estudios Epidemiológicos; CRS: Escala de Evaluación de Carro-I; HDRS: Escala de Depresión de Hamilton; MADRS: Escala de Evaluación de la Depresión de Montgomery-Asberg; SDS: Escala Autoaplicada de Depresión. De Vázquez, 1995

TABLA 21-4. Puntos de corte más utilizados en los principales instrumentos de evaluación de la depresión

	Rango	Sin depresión	Depresión ligera	Depresión moderada	Depresión grave
BDI-13 ítems	0-39	0-4	5-7	8-15	16-39
BDI-21 ítems	0-63	0-9	10-18	19-29	30-63
CES-D	0-60	0-15	-	-	-
CRS	0-52	0-9	-	-	-
Escala de Melancolía de Bech-Rafaelsen	0-44	0-5	6-14	15-25	26-44
HAD	0-42	0-7	-	-	-
HDRS-17	0-52	0-6	7-17	18-24	25-52
MADRS	0-60	0-9	-	-	-
SDS (Zung)	20-100	20-35	36-51	52-67	68-100

BDI: Inventario de Depresión de Beck; CES-D: Escala de Depresión del Centro de Estudios Epidemiológicos; CRS: Escala de Evaluación de Carroll; HAD: Escala Hospitalaria de Ansiedad y Depresión; HDRS: Escala de Depresión de Hamilton; MADRS: Escala de Depresión de Montgomery Asberg; SDS: Escala Autoaplicada de Depresión. De Vázquez, 1995.

Como en la inmensa mayoría de los instrumentos clínicos, con la posible excepción del MMPI-D, no existe una estandarización de las puntuaciones en diversas muestras de sujetos o pacientes. El rango de puntuación de la HDRS varía de 0 a 52 y se emplean por lo general una serie de puntos de corte (tabla 21-4).

Fiabilidad. A pesar de la indefinición sobre los criterios de valoración, tanto la fiabilidad interexaminadores de la puntuación total como la correlación con la impresión clínica global de expertos sobre el paciente son sorprendentemente buenas: en torno a 0,90 (Bech y cols., 1975; Hedlund y Vieweg, 1979a). Incluso estudiantes con un entrenamiento no muy intensivo pueden obtener una fiabilidad interexaminadores cercana a 0,75 para la puntuación global (Rehm y O'Hara, 1985).

Validez. La escala es bastante consistente: las correlaciones entre ítem y puntuación total se sitúan en torno a 0,70, si bien hay algunos ítems muy por debajo de estos índices medios (Bech, 1981 y 1987), lo que debería ser precisado en versiones futuras.

La HDRS distingue razonablemente bien diferentes grupos de pacientes en función de su gravedad; por ejemplo, Schwab y cols. (1965) compararon las puntuaciones en la HDRS de pacientes médicos por un lado y depresivos por otro y hallaron una distribución bimodal.

En términos de validez convergente las correlaciones son moderadamente altas tanto con otros instrumentos de observación acreditados como con escalas autoaplicadas (Kearns y cols., 1982; Senra y Polaino, 1993). En concreto, Hedlund y Vieweg (1979a) en su amplia revisión de estudios, informaron de unas correlaciones medias de la HDRS con el Inventario de Depresión de Beck (Beck Depression Inventory, BDI)

la SDS y el Inventario Multifásico de Personalidad de Minnesota (Minnesota Multiphasic Personality Inventory, MMPI) de 0,58, 0,45, y 0,44 respectivamente. En cuanto al contenido factorial de la HDRS, los resultados son algo inciertos. Para Bech (1981) los análisis factoriales parecen indicar un factor unimodal de gravedad y un factor bimodal de depresión (retardada frente a agitada). Para otros autores (Rhoades y Overall, 1983) las soluciones factoriales apuntan más bien a la existencia de dos factores generales: depresión somática y depresión cognitiva.

Utilidad. La escala no pretende efectuar diagnósticos de depresión. Fue diseñada como un instrumento para cuantificar la sintomatología de pacientes una vez diagnosticados categorialmente de depresión, aunque su uso se ha extendido para valorar síntomas depresivos en cualquier otro cuadro psicológico o médico (p. ej., pacientes de Parkinson, anoréxicas, pacientes cardíacos, esquizofrénicos, etc.) (Robertson, 1989).

Un aspecto importante a la hora de la elección es que la versión de 17 ítems se comporta psicométricamente tan bien como la de 21 y, de hecho, esta versión más larga no añade información especialmente relevante para la depresión (Maier y cols., 1988), por lo que parece más recomendable utilizar la HDRS-17.

La HDRS es una escala sensible para detectar cambios debidos al tratamiento (Hedlund y Vieweg, 1979a) y, de hecho, es posiblemente la escala más utilizada para evaluar la eficacia terapéutica de los antidepressivos (Robertson, 1989). No obstante, el problema técnico de la definición de «mejoría» es de una enorme complejidad, pero rebasa la finalidad de este capítulo discutirlo extensamente (Vázquez y Sanz, 1991a). Baste señalar que, en términos de la HDRS, para algunos la eficacia terapéutica consiste en disminuir la puntuación total en un 50%, para otros consiste

en disminuirla al nivel de no depresión (p. ej., 6 puntos o menos), y para otros simplemente consiste en una disminución estadística significativa. Pero, en cualquier caso, parece que a la HDRS le falta sensibilidad para apreciar cambios en los pacientes más graves; en estos casos, escalas como la Escala de Evaluación de la Depresión de Montgomery-Asberg (Montgomery-Asberg Depression Ratings Scale, MADRS) pudieran ser más adecuadas.

Uno de los problemas de la escala es que existe cierta confusión sobre su empleo. Para valorar los síntomas se requiere tener en cuenta toda la información posible sobre el paciente y el empleo de una entrevista clínica. Aunque ha habido varios esfuerzos para concretar estos procedimientos (Hamilton, 1980; Loudon y Tiplady, 1981; Miller y cols., 1985), no existen aún pautas claras para llevar a cabo la HDRS, con la excepción de las escuetas orientaciones sugeridas por su autor (Hamilton, 1967). Además, los intentos por transformar esta escala heteroaplicada en un listado de síntomas de más fácil aplicación (Conde y Franch, 1984; Miller y cols., 1985), han sido frontal y directamente rechazados por el autor de la escala (Hamilton, 1987; Hamilton y Shapiro, 1990). Así pues, la escala deja una cierta parte de la variancia sin control: como no se especifican con detalle los procedimientos de evaluación, cada médico decide aspectos tan importantes como lo que significa intensidad moderada, ligera o grave (Thompson, 1989b).

Por último, una objeción a la escala se refiere a su contenido. Hay ítems, como la falta de *insight*, que además de su insuficiente fiabilidad posiblemente no deberían contemplarse como un síntoma más de la depresión. Es dudoso que esto tenga que ver en realidad con la gravedad del cuadro.

Escala de Evaluación de la Depresión de Montgomery-Asberg

La MADRS fue específicamente diseñada para ser sensible al cambio y a los efectos del tratamiento (Montgomery y Asberg, 1979). La MADRS procede de la Escala de Evaluación Psicopatológica Comprehensiva (Comprehensive Psychopathological Rating Scale, CPRS) de Asberg y cols. (1978).

En una escala corta, con claras instrucciones sobre su uso y sobre cómo asignar grados de gravedad (Snaith, 1987), y existe una versión en castellano de la MADRS (Conde y Franch, 1984). La MADRS consiste en 10 ítems que se evalúan de 0 a 6 (el 0 indica síntoma inexistente). Hay cuatro síntomas relacionados con los estados de ánimo (tristeza aparente, tristeza expresada, pensamientos pesimistas y pensamientos suicidas); dos con ansiedad (tensión interna, dificultades

de concentración); tres con síntomas somáticos (sueño disminuido, apetito disminuido y laxitud) y uno con anhedonia (incapacidad para sentir). La escala se debe utilizar tras una entrevista clínica. Una puntuación de 9 o inferior indica que no existe depresión (Levine y cols., 1987).

Fiabilidad. Aunque la consistencia interna de esta escala es aceptable (en torno a 0,70), es más baja que la hallada en la HDRS de 17 ítems y en la Escala de Melancolía de Bech-Rafaelsen. Una posible explicación de esto es que emplea un rango excesivamente amplio (7 grados) para evaluar la gravedad de cada ítem, lo cual implica obtener un alto acuerdo inter-examinadores (Maier y cols., 1988).

Validez. El rendimiento global de la MADRS es comparable a la HDRS (Kearns y cols., 1982) y, de hecho, su correlación con ésta es de 0,95. A pesar de su brevedad, como hay que emplear una entrevista *ad hoc*, posiblemente en la práctica se consume tanto tiempo como con la HDRS (Hamilton, 1987). Dado que no evalúa a fondo aspectos vegetativos e incide más en síntomas anímicos, es útil cuando se trata de pacientes con enfermedades físicas concomitantes cuyos síntomas somáticos pueden confundirse con los síntomas somáticos propiamente depresivos (Lloyd, 1990; Kearns y cols., 1982). De hecho, su correlación con la subescala de depresión de la Escala Hospitalaria de Ansiedad y Depresión (Hospital Anxiety and Depression, HAD), un instrumento específicamente diseñado para evaluar depresión y ansiedad en enfermos físicos hospitalizados, es de 0,81.

Utilidad. Una de las ventajas de la MADRS es que tiene más poder estadístico que la HDRS y, por lo tanto, se necesitan menos pacientes para obtener un mismo grado de sensibilidad (Thompson, 1989b).

Escala de Melancolía de Bech-Rafaelsen

Es una modificación efectuada sobre sólidas bases matemáticas de la de Hamilton, con el fin de conseguir mejorar su validez puesto que la HDRS ha sido criticada por su relativa falta de homogeneidad de constructo (Bech y Rafaelsen, 1980). Se trata de una escala heteroaplicada basada en los 11 ítems de la HDRS con mejores correlaciones con un criterio externo de gravedad otorgada por evaluadores expertos. Como resultado de este procedimiento, se suprimieron de la HDRS los ítems más vegetativos, con la excepción del sueño, aunque reducido éste a un único ítem. Lo más interesante del resultado de depuración descrito es la evaluación mediante 4 ítems del entumecimiento mo-

tor, verbal, intelectual y emocional que caracterizan, entre otros síntomas, la denominada depresión melancólica:

1. Estado de ánimo deprimido.
2. Sentimientos de culpabilidad.
3. Trabajo e intereses.
4. Trastornos de sueño.
5. Cansancio y dolores.
6. Ansiedad psíquica.
7. Impulsos suicidas.
8. Enlentecimiento emocional.
9. Enlentecimiento intelectual.
10. Actividad verbal disminuida.
11. Actividad motora disminuida.

Cada ítem se valora en una escala de 5 puntos, desde 0 (ausente) a 4 (grave). La correlación de los ítems con la puntuación total varía de 0,52 a 0,88.

Fiabilidad y validez. El trabajo psicométrico con la escala es muy completo (Bech, 1987; Maier y Philipp, 1985). La fiabilidad interexaminadores varía de 0,79 a 0,93 (Bech y cols., 1983) y su correlación con la escala de Hamilton es excelente: 0,97 (Bech y cols., 1979; Rafaelsen y cols., 1980). Por otro lado, como se ha demostrado en un completo estudio comparativo (Maier y cols., 1988), la consistencia interna de la escala es mejor que la de la HDRS y la MADRS.

Utilidad. La Escala de Melancolía de Bech-Rafaelsen parece una escala apropiada para la evaluación de pacientes melancólicos o con enlentecimiento o, lo que es prácticamente lo mismo, de pacientes graves (Bech, 1987; Vázquez y Sanz, 1991a). Así mismo, el trabajo de Maier y cols. (1988) ha demostrado que está especialmente indicada cuando se pretenden evaluar cambios terapéuticos, en cuyo caso su rendimiento es mejor que el de la HDRS y la MARS-D.

Al igual que la HDRS se supone que su uso está limitado a personas con cierto conocimiento clínico. No obstante, se han publicado unas pautas para facilitar su aplicación (Bech y cols., 1986). Junto a esta escala de melancolía, Bech diseñó una escala con ítems opuestos a los citados para evaluar manía, que posteriormente se revisa.

Escala de Evaluación del Retardo de Salpêtrière

La Salpêtrière Retardation Rating Scale (SRRS) de Widlöcher (Widlöcher, 1983a y b) es una escala de 15 ítems específicamente diseñada para evaluar enlentecimiento psicomotor. Cada ítem se evalúa de 0 a 4 puntos en términos de la gravedad presente. La pun-

tuación total varía entre 0 y 60. En realidad, incluye dos componentes (Brébion y cols., 1995): los 6 primeros ítems de la escala están referidos al comportamiento motor observable y a la conducta verbal, mientras que los 8 restantes abarcan aspectos psicológicos y subjetivos del enlentecimiento (p. ej., fatiga, falta de interés por hacer las cosas habituales). Finalmente, un último ítem hace referencia a la apreciación global sobre el enlentecimiento general del paciente. Aunque los ítems de la subescala del comportamiento observable, parece la más consistente y válida (Brébion y cols., 1995), el listado completo de áreas evaluadas es el siguiente:

1. Lentitud del paso.
2. Lentitud y movimientos extraños de extremidades y tronco.
3. Lentitud y movimientos extraños de cabeza y cuello.
4. Flujo lento del habla.
5. Intensidad y modulación de voz.
6. Brevidad en las respuestas.
7. Variedad de temas en el lenguaje espontáneo.
8. Riqueza de asociaciones ante temas propuestos por el entrevistador.
9. Experiencia subjetiva de rumiaciones.
10. Fatiga.
11. Interés en actividades habituales.
12. Percepción del paso del tiempo.
13. Déficit de memoria.
14. Concentración.
15. Valoración global por parte del entrevistador.

Fiabilidad. La subescala del comportamiento observable ha mostrado una buena consistencia interna: las correlaciones entre los 6 ítems oscilan entre 0,55 y 0,84 (Granier y cols., 1983). La escala completa tiene, en su conjunto, una buena fiabilidad interexaminadores: en torno a 0,90 (Leclercq, 1983).

Validez. La SRRS muestra correlaciones elevadas con la puntuación general en escalas de depresión como la HDRS (0,70). Existe por lo tanto, una polémica sobre si el enlentecimiento motor se trata sólo de un componente más de un factor general de depresión. Sin embargo, algunos análisis factoriales de la SRRS sugieren que los componentes más referidos al enlentecimiento motor de la escala (los primeros ítems) parecen formar un factor aparte (Widlöcher y Ghoslan, 1989). En efecto, la factorización de la escala ha demostrado que los ítems de la subescala del comportamiento observable son más específicos de un factor de enlentecimiento, mientras que los 8 ítems restantes están más vinculados a la gravedad general de la depresión tal y como se valora con la HDRS y otras esca-

las (Granier y cols., 1983). La correlación entre las dos subescalas de la SRSS es sólo moderada ($r = 53$) (Bré-bion y cols., 1995).

En cuanto a la validez discriminante, la escala de enlentecimiento parece diferenciar adecuadamente entre pacientes con depresión y pacientes con diversas enfermedades motoras como, por ejemplo, enfermos parkinsonianos (Widlöcher y Chozlan, 1989).

Utilidad. Dado que el enlentecimiento psicomotor es uno de los mejores predictores de respuesta al tratamiento con fármacos antidepressivos, esta escala puede ser de utilidad. No obstante, aún queda bastante trabajo para elucidar qué componentes concretos de éste pueden ser diferencialmente más válidos.

Instrumentos autoaplicados de depresión

Antecedentes históricos

Si la posguerra fue denominada la «era de la ansiedad» (May, 1950), posiblemente estamos viviendo en la «era de la depresión» como demuestran las tendencias seculares en ascenso de este cuadro. Los datos epidemiológicos indican que la depresión es un trastorno con una tendencia ascendente (Vázquez y Sanz, 1995): las cohortes más jóvenes (nacidos después de 1960) presentan tasas que llegan a duplicar o triplicar las de generaciones anteriores cuando tenían su misma edad (Fombonne, 1995). Sorprendentemente, la medida de la depresión fue relativamente ignorada hasta mediados los años sesenta (Berndt, 1993).

Uno de los primeros instrumentos de autoevaluación fue el propuesto por Jasper (1930) quien, queriendo establecer paralelismos con las teorías de inteligencia por entonces pujantes (p. ej., la teoría de Spearman sobre la existencia de un factor general «g» de inteligencia), creó una escala de depresión-alegría que, bajo esa dimensión general, englobaría dimensiones como optimismo-pesimismo o entusiasmo-apatía. Esta escala de 40 ítems introducía preguntas hoy en día un tanto atípicas (p. ej., apatía hacia instituciones sociopolíticas).

Tras algunos intentos posteriores (Chant y Myers, 1936; Guilford y Guilford, 1939), el gran revulsivo se produce tras la II Guerra Mundial y, en especial, con la aparición del MMPI, que incluía una subescala de depresión (MMPI-D, Hathaway y McKinley, 1942), que dominó el panorama durante varias décadas a pesar de su más que dudosa validez (véase más adelante) y su objetivo explícito de medir un rasgo de personalidad más que un estado de ánimo depresivo.

El antecedente de las escalas modernas se sitúa en 1961 con la creación de Beck de la escala que lleva su

nombre. Beck crea una escala con síntomas heterogéneos y muy descriptiva del grado sintomático. Aparte del prestigio incipiente de su autor en el ámbito de la nascente terapia cognitivo-conductual, una de las razones de su éxito fue que, hasta 1978, la escala tuvo una libre circulación, sin editor, lo que la convirtió en un instrumento sencillo, gratuito y de fácil acceso (Berndt, 1993). A esta escala la siguieron en 1965 la SDS de Zung (muy promocionada por laboratorios farmacéuticos) y, en esa misma fecha, aparecieron las escalas de adjetivos de Lubin (1965), ya con un formato completamente diferente a los autoinformes tradicionales. Finalmente, en 1969 aparece la propuesta de las escalas analógicas de Aitken, en las que el sujeto marca en una línea con dos polos opuestos dónde se situaría su estado de ánimo. En los últimos 25 años no ha habido desarrollos realmente innovadores en cuanto a los instrumentos, si bien desde el plano teórico de los afectos se han desarrollado formulaciones teóricas que podrían iluminar nuevos modos de evaluación (Watson y Kendall, 1989).

Inventario de Depresión de Beck

El BDI (Beck y cols., 1961; Beck, 1978) es el instrumento de autoinforme más citado en la bibliografía. Parte de su éxito se debe a que su autor, Aaron T. Beck es la figura más prominente de los modelos cognitivos de la depresión y la terapia cognitiva. El BDI consta de 21 ítems que evalúan un amplio espectro de síntomas depresivos: cognitivos, afectivos, conductuales y fisiológicos.

Aunque se diseñó con la idea de que no estuviese ligado a ninguna teoría en particular, el BDI se caracteriza por cubrir más síntomas cognitivos (desesperanza, pérdida de autoestima, falta de concentración, etc.) que otros instrumentos (tabla 21-3). Existen dos versiones; una de 1961 y otra de 1978, que deja intactos sólo 4 ítems de la versión inicial. En la versión de 1961 (Beck y cols., 1961) el instrumento albergaba entre cuatro y seis opciones de respuesta por ítem (tal y como recogió la versión española de Conde y cols., 1976), mientras que la versión posterior de 1978 (dada a conocer por Beck y cols. en 1983) alberga sistemáticamente cuatro alternativas de respuesta en cada uno de los 21 ítems.

Cada ítem del BDI está compuesto por cuatro alternativas, ordenadas de menor a mayor gravedad (de 0 a 3 puntos), y la tarea del sujeto es elegir aquella frase que se aproxime más a cómo se ha sentido durante la última semana. La puntuación total del BDI consiste simplemente en la suma de los puntos obtenidos en los 21 ítems (con un rango de 0 a 63). En la tabla 21-4 se señalan los puntos de corte sugeridos por Beck y cols. (1988).

Beck y Beck (1972) crearon una versión breve del BDI eligiendo los 13 ítems que se correlacionaban en mayor medida con la puntuación total (correlación media = 0,96) y con evaluaciones clínicas del nivel de gravedad (correlación media = 0,61). No obstante, en una serie de recientes directrices sobre el uso del BDI se ha desaconsejado el uso de estas versiones abreviadas pues su precisión clasificatoria, como instrumentos de cribado, es bastante pobre (Kendall y cols., 1987) y, además, la estructura factorial de la escala en su versión reducida difiere de la solución factorial de la escala normal (Gould, 1972).

El BDI cubre bastante bien el tipo de síntomas descritos en el DSM-IV, pero incluye otros que no se tienen en cuenta en estos criterios diagnósticos: irritabilidad, llanto, hipocondría y autoimagen corporal (véase un análisis comparativo en Vázquez y Sanz, 1991a). Por otro lado, la filosofía del BDI es la de recoger síntomas por defecto pero no por exceso»(p. ej., aumento de peso, agitación psicomotora, hipersomnio), de tal modo que la idea subyacente a este inventario es que la depresión se caracteriza por una inhibición general ignorando así otros posibles síntomas de naturaleza contraria que, por ejemplo, si se contemplan en un diagnóstico categorial tipo DSM-IV.

Fiabilidad. Los estudios psicométricos con el BDI son casi exhaustivos (Beck y cols., 1988) y es, con la HDRS, el instrumento más utilizado. La fiabilidad dos mitades es alta o muy alta (desde 0,58 a 0,93) y su consistencia interna muy alta tanto en pacientes (α de Cronbach media, 0,87) como en muestras de población general (α de Cronbach media, 0,81).

El examen de la fiabilidad test-retest es, como en cualquier escala que intente ser sensible al cambio y que pretenda evaluar estados de ánimo, menos relevante. No obstante, en términos generales, la fiabilidad test-retest se sitúa en un rango de 0,60-0,70 para periodos de entre 1 semana a 4 meses (Berndt, 1993). En resultados preliminares con muestras españolas universitarias indican una fiabilidad test-retest de 0,70 al cabo de 3 meses y una α de Cronbach de 0,82 (Vázquez y Sanz, 1991b).

Validez. En cuanto a su validez convergente, el BDI se ha comparado de forma exhaustiva con otras escalas (Robertson, 1989; Vizcarro y García, 1981; Thompson, 1989b; Vázquez, 1986; Shaw y cols., 1985; Beck y cols., 1988). Excepto en algunos estudios (Kearns y cols., 1982), en general la correlación media entre la puntuación total del BDI, la HDRS, la SDS de Zung, y las escalas analógicas visuales es moderadamente alta (en torno a 0,70). Por ejemplo, la correlación media con la HDRS es de 0,73 para pacientes psiquiátricos y 0,74 para sujetos de población general (Beck y

cols., 1988). La correlación es menor con listas de adjetivos depresivos, pero seguramente debido a la dudosa validez de estos últimos instrumentos (véase más adelante). Por otro lado, las correlaciones con otras muchas escalas de depresión son bastante altas (entre 0,50 y 0,80) quizá porque el BDI se suele utilizar como criterio de evaluación en la construcción de otras escalas por lo que, tautológicamente, se consiguen altas correlaciones (Shaver y Brennan, 1991).

El BDI es muy utilizado para evaluar los cambios terapéuticos (Dobson, 1989; Freeman y Munro, 1990) y, en este sentido, sus resultados son semejantes a la HDRS (Hedlund y Vieweg, 1979a). En efecto, la mejoría clínica juzgada por expertos correlacionada alrededor de 0,80 con los cambios en las puntuaciones en el BDI (Beck y cols., 1975). Bailey y Coppen (1976) demostraron que en algunos pacientes la HDRS refleja mejor los cambios terapéuticos que el BDI, mientras que en otros, sin que se sepa aún por qué, sucede lo contrario: el BDI es más sensible que la HDRS. Esto sugiere la conveniencia, cuando ello sea posible, de incluir ambos instrumentos en un mismo protocolo (Edwards y cols., 1984; Lambert y cols., 1986).

En nuestro país se dispone de la adaptación efectuada por Conde y cols. (1976), en la actualidad algo desfasada pues emplearon la primera versión del BDI y se introdujeron algunas modificaciones *sui generis* como, por ejemplo, la eliminación de 2 ítems completos (el que hace referencia a sentimientos de culpabilidad y el de autoimagen) y aleatorizar las alternativas de respuesta para evitar sesgos (en el original, las alternativas están siempre ordenadas de menor a mayor gravedad). En la actualidad se está efectuando una adaptación al castellano de la nueva versión de 1978 con más de 2.000 casos (1.600 de población general y 400 pacientes psiquiátricos ambulatorios) (Vázquez y Sanz, en preparación). De las decenas de análisis factoriales efectuados con el BDI (Beck y cols., 1988) parece desprenderse que hay un factor general de depresión que se puede descomponer en tres más intercorrelacionados: a) actitudes negativas hacia uno mismo o suicidio; b) molestias somáticas, y c) rendimiento disminuido. En su conjunto, la estructura factorial del BDI confirma que en su composición el área cognitiva tiene un peso mayor que la somática (Loks y cols., 1989).

En nuestro país, Guerrero y cols. (1986) han demostrado, utilizando análisis discriminante, que se puede emplear una subescala del BDI basada en los ítems de culpabilidad (5 ítems), sensación de castigo (6 ítems), autoacusación (8 ítems), aislamiento social (12 ítems) e indecisión (13 ítems) como un método sensible y específico de diagnóstico de la melancolía, una vez otorgados ciertos pesos a las puntuaciones en cada uno de esos 5 ítems.

Utilidad. Al igual que lo que sucede con la HDRS, se ha utilizado y se utiliza con poblaciones muy dispares, no sólo con pacientes depresivos, cuando se pretende evaluar la presencia de síntomas depresivos. Así, por ejemplo, existe una versión disponible para su empleo en niños (Kovacs y Beck, 1977).

Uno de los problemas del BDI es que, a pesar de su sencillez, para pacientes muy graves o con enlentecimiento puede ser difícil de entender, engorrosa o inaplicable (Kearns y cols., 1982). Un grave error común es pensar que el BDI diagnostica la depresión. El BDI, como la mayoría de los instrumentos de evaluación de síntomas, simplemente ofrece una aproximación cuantitativa, nunca diagnóstica, de los síntomas depresivos (Vázquez, 1986; Kendall y cols., 1987; Ruiz y Bermúdez, 1989): una alta puntuación en el BDI no indica necesariamente la presencia de un trastorno depresivo tal y como se define en clasificaciones como el DSM-IV. Por ejemplo, muchos pacientes pueden tener puntuaciones altas en el BDI, pero pocos cumplen criterios estrictos DSM-IV de trastorno del estado de ánimo (Smith y cols., 1985). Evidentemente, esto también sucede con la HDRS y el resto de las escalas cuyo objetivo no es diagnosticar sino evaluar la gravedad de síntomas: en un estudio con 80 pacientes con una puntuación mayor a 25 en la HDRS, sólo el 71 % satisfacía los RDC, como se sabe muy similares a los criterios DSM-III (APA, 1980).

Escala de Depresión del Centro de Estudios Epidemiológicos

Esta escala (Center for Epidemiological Studies Depression Scale, CES-D) es una de las más utilizadas en estudios epidemiológicos en la población general. Muchos de sus ítems se centran en el estado anímico y, al contrario que el BDI, la HDRS, y otros instrumentos similares, se diseñó para su uso preferente en muestras no clínicas (Radloff, 1977). Consta de 20 ítems, que Radloff identificó como los más característicos de la depresión en estudios clínicos y factoriales publicados. Se pide a los sujetos que señalen la frecuencia con la que experimentaron ese síntoma (p. ej., «me sentí triste») durante la semana pasada, en una escala de 4 puntos. La escala oscila desde 0 puntos [«raramente (menos de un día) o en ningún momento»] hasta 3 [«la mayor parte del tiempo (5-7 días) o todo el tiempo»]. Para minimizar sesgos de respuesta, hay 4 ítems formulados de modo inverso (p. ej., «estuve feliz»). La puntuación oscila entre 0 y 60 y no se ofrecen puntos de corte específicos de diferentes niveles de gravedad.

La CES-D se ha empleado en estudios comunitarios a gran escala desde la década de los setenta (Shaver y Brennan, 1991) y ha sido validada también con

algunas muestras, relativamente pequeñas, de pacientes depresivos (Radloff, 1977).

Fiabilidad. La consistencia interna de la escala es alta. La correlación dos mitades es de 0,85 para pacientes y de 0,77 en sujetos normales, mientras que los coeficientes α de Cronbach se sitúan en torno a 0,90 pacientes y 0,80 en normales (Radloff, 1977). En cuanto a la fiabilidad test-retest, los datos oscilan entre 0,67 para 4 semanas hasta 0,32 para un período de 12 meses.

Validez. En muestras clínicas la escala tiene índices de correlación moderados (alrededor de 0,60-0,70) con apreciaciones efectuadas por entrevistadores mediante la HDRS (Radloff, 1977). Las correlaciones con el BDI y la SDS son altas: 0,81 y 0,90, respectivamente (Weissman y cols., 1975).

Con el punto de corte habitual en la CES-D (una puntuación total de 15), la escala presenta una baja sensibilidad (detecta sólo alrededor del 60 % de los casos de depresión mayor), es decir, comete muchos falsos negativos, aunque su especificidad es adecuada (identifica correctamente un 90 % de las personas sin depresión mayor) (Berndt, 1993; Wetzler y Van Praag, 1989).

Diversos análisis factoriales exploratorios (Iartzog y cols., 1990) han confirmado que la CES-D contiene cuatro diferentes factores: afecto depresivo, síntomas somáticos, bienestar y relaciones interpersonales.

Utilidad. La CES-D es una escala sencilla, directa y muy fácil de comprender por parte de los sujetos. Una de las limitaciones más importantes de la escala es su limitada validez de contenido: el rango de síntomas que cubre es muy limitada al centrarse en aspectos anímicos y prestar muy poca o nula atención a síntomas vegetativos y comportamentales. En este sentido, aunque su correlación con el BDI o la HDRS es alta, su uso con muestras clínicas está seriamente comprometido. Sin embargo, su uso podría ser aconsejable si se intenta valorar el estado de ánimo en muestras de población general. Otra peculiaridad añadida es que evalúa frecuencia y no intensidad de síntomas. Por último, conviene indicar que existe una versión similar para niños (la CES-DC, Fautsch y cols., 1986).

Inventario Multifásico de Personalidad de Minnesota

El MMPI (Hathaway y MacKinley, 1942) cuenta con 566 ítems, de los cuales 60 componen específicamente una subescala de depresión (MMPI-Depression). El formato de respuesta es verdadero o falso y existen diversas versiones informatizadas. El MMPI es

uno de los instrumentos más utilizados en psicología clínica y la bibliografía existente es muy numerosa. Por desgracia, dado su enfoque «estático» (intenta medir la personalidad del individuo) su utilidad clínica para conocer problemas específicos del sujeto es mínima (Snaith y cols., 1971).

El MMPI no evalúa propiamente molestias o quejas del paciente sino más bien características generales del comportamiento que, según sus autores, son diferenciadoras de las personas con depresión de las que no lo son. Esta aproximación «empírica» es muy discutible. En efecto, en el caso de la MMPI-D se trataba de encontrar ítems que diferenciasen pacientes deprimidos de pacientes no deprimidos, lo cual plantea graves problemas de validez de criterio (Helmes y Reddon, 1993): un paciente con depresión puede diferenciarse en muchas cosas de un paciente no deprimido y, sin embargo, esos elementos de diferencia pueden no mostrar ninguna relevancia clínica, epidemiológica, causal o terapéutica.

Fiabilidad. La consistencia interna de la escala, dada su heterogeneidad intrínseca, es muy baja (Berndt, 1993). Pero incluso la subescala de depresión aisladamente tiene una estructura factorial compleja (lo que denota que no mide un constructo único) y con poco poder explicativo de la variancia total. Los intentos de crear una versión reducida a partir del conjunto de los 60 ítems depresivos (Cantor, 1960; Dempsey, 1964) no han resuelto los problemas de la falta de validez discriminante del MMPI-D (en concreto, parece medir más neuroticismo que depresión).

Validez. La escala del MMPI en su totalidad tiene una validez bastante dudosa (Helmes y Reddon, 1993) por lo que su uso no parece aconsejable. En términos de validez concurrente, la subescala de depresión muestra correlaciones inaceptablemente bajas, en su conjunto, con escalas como el BDI (Beck y cols., 1988).

En ocasiones se utilizan determinados perfiles de puntuaciones del MMPI como un modo de detección de casos (Winters y cols., 1985); sin embargo, estos perfiles no son específicos de la depresión sino que son también frecuentes en otros cuadros psicopatológicos. Por ejemplo, un estudio sobre la capacidad del MMPI-D para detectar casos de depresión mayor, el punto de corte típico (T mayor de 70) fue capaz de identificar correctamente el 92% de depresión mayor, pero produjo tasas inaceptablemente altas de falsos positivos del 55% (Post y cols., 1985); por otro lado, el índice de depresión de Mezzich (Mezzich y cols., 1974), basado en el MMPI completo, produjo tasas de falsos positivos del 73% en este mismo estudio. Así pues, mientras que la sensibilidad del MMPI puede ser adecuada, su coste en términos de falsos posi-

vos es excesivamente elevado debido a su inespecificidad.

Por otro lado, una puntuación elevada o en la subescala MMPI-D (o en la subescala de manía, MMPI-Ma) no indican necesariamente que la depresión o la manía sean el problema relevante dado que las subescalas del MMPI tienen una correlación elevada entre sí. Así pues, una puntuación elevada pudiera deberse a puntuaciones elevadas en otras escalas (hipocondría).

Otro problema añadido es que la validación de la escala se efectuó en la década de los cincuenta con grupos diagnosticados de diversos cuadros (depresión); la propia falta de fiabilidad de estos diagnósticos efectuados antes de los años ochenta (Vázquez, 1990a) cuestiona así mismo la validez intrínseca de la escala. Aunque el MMPI-2 parece haber mejorado algo el proceso de construcción, presenta aún serias deficiencias como han demostrado Helmes y Reddon (1993) en más acertadas y devastadoras críticas.

Utilidad. En general, no ofrece ventajas sobre otros instrumentos, pero presenta muchos inconvenientes conceptuales. En suma, a pesar de su elevado uso, el MMPI o el MMPI-2 son instrumentos poco recomendables tanto por sus limitaciones empíricas como teóricas.

Escala de Evaluación de Carroll

La Carroll Rating Scale (CRS) (Carroll y cols., 1981) fue diseñada con el fin de establecer un procedimiento de recogida de información que fuese paralelo, pero autoaplicado, al de la HDRS (Hamilton, 1960). Es una escala que se centra de modo casi exclusivo en síntomas somáticos y conductuales. Consta de 52 ítems formulados de modo sencillo que se refieren a síntomas percibidos en los últimos días. Se le pide al sujeto responder a cada ítem con una alternativa de verdadero o falso. Consta de 40 ítems directos (p. ej., «desde que comencé mi enfermedad he perdido todo mi interés por el sexo») y 12 en sentido inverso (p. ej., «me siento tan lleno de energía como antes»), para evitar sesgos de respuesta. La CRS cubre los 17 ítems originales de la HDRS y para ello emplea, en cada uno de los apartados, varios ítems. A diferencia de la HDRS (que en algunos casos se evalúan de 0 a 2 puntos y en otros de 0 a 4 puntos), todos los ítems de la CRS se valoran entre 0 y 1. El rango de puntuación total es de 0 a 52 y los autores sugieren emplear tentativamente un punto de corte de 10 puntos para diferenciar depresión de no depresión.

Fiabilidad. Los datos de Carroll y cols. (1981) derivados de diversas muestras de pacientes (depresivos,

psiquiátricos de control) y de personas de la población general, proporcionan unos índices de 0,87 en correlación de ítems pares e impares y una correlación ítem-puntuación total con un rango quizá excesivamente alto: entre 0,05 y 0,78, lo que, como sucedía con la HDRS, sugiere una cierta falta de consistencia. No hay información test-retest.

Validez. Dado que la CRS se diseñó para contar con un instrumento paralelo a la HDRS, hay muchos datos sobre la validez convergente de la CRS con la escala de Hamilton (Carroll y cols., 1981; Nasr y cols., 1984). Carroll y cols. (1981) publicaron una matriz de correlaciones de cada uno de los ítems de la HDRS. Pero, a pesar de ser escalas aparentemente paralelas, las correlaciones medias entre los ítems de ambas escalas fue sólo moderada, oscilando entre un decepcionante - 0,06 hasta 0,73 (mediana, 60). Esto indica, de nuevo, que la escala es heterogénea en cuanto a lo que realmente está midiendo. La correlación entre las puntuaciones totales de ambas escalas fue alta ($r = 80$), lo que es similar a lo hallado en otros estudios (Feinberg y cols., 1981; Nasr y cols., 1984). Así mismo, empleando una muestra de pacientes psiquiátricos de diferentes diagnósticos, la correlación entre la puntuación global de la CRS y el BDI o escalas analógicas visuales es alta: 0,86 (Carroll y cols., 1981) y 0,87 (Feinberg y cols., 1981), respectivamente.

Utilidad. No es una escala muy utilizada y no hay una gran abundancia de datos psicométricos. Puede ser una alternativa adecuada de la HDRS cuando se estime adecuado, si bien hay que tener en cuenta que parece sobredimensionar la gravedad de los pacientes más graves (Shaver y Brennan, 1991). De hecho, los mejores índices de concordancia entre la CRS y la HDRS se observan con pacientes menos graves, lo cual incide de nuevo en la cuestión de si ambas escalas cubren los mismos constructos.

Escala Autoaplicada de Depresión de Zung

La Self-rating Depression Scale (SDS) fue diseñada por Zung (1965). Consiste en 20 frases relacionadas con la depresión, diez sobre aspectos negativos (p. ej., «me canso por cualquier cosa») y diez sobre positivos («por la mañana es cuando mejor me siento»), que el sujeto evalúa en una escala Likert de 4 puntos en términos de frecuencia desde 1 (raramente o nunca) hasta 4 (casi todo el tiempo, o siempre). Es una escala que ha sido adaptada al castellano por Conde y su equipo (Conde y Esteban, 1976; Conde y Franch, 1984). El objetivo del autor (Zung, 1986) fue evaluar de forma se-

para los distintos síntomas de la depresión a través de cuatro grandes categorías:

1. Síntomas psicoafectivos: 2 ítems (estado de ánimo y llanto).
2. Síntomas fisiológicos: 8 ítems (sueño, apetito, libido, estreñimiento, etc.).
3. Entrecimiento o agitación psicomotora: 2 ítems.
4. Síntomas psicológicos: 8 ítems (desesperanza, confusión, irritabilidad, ideación suicida, etc.).

La puntuación total se expresa en forma de porcentaje sobre la puntuación total posible (80 puntos), es decir, multiplicando la puntuación del sujeto por 80 y dividiéndola por 100. De este modo, el rango oscila entre un 20 y un 100 %. En la tabla 21-4 se señalan los puntos de corte habituales.

Fiabilidad. Hay poca información sobre la fiabilidad de esta escala (Shaw y cols., 1985). En términos de la consistencia interna de la escala, Zung proporcionó un índice de 0,73 en la fiabilidad dos mitades (correlación entre ítems pares e impares), datos algo inferiores a los obtenidos en la adaptación española (Conde y cols., 1976): 0,80. En cuanto a su consistencia interna, los índices α de Cronbach son altos, oscilando entre 0,79 (Knigh y cols., 1983) y 0,92 (Zung, 1986).

Validez. Hay bastante información sobre la validez convergente de la escala. La correlación entre la SDS y la HDRS oscila entre 0,50 y 0,79. También presenta una correlación moderadamente alta con el MMPI-D: entre 0,59 y 0,70 (Zung, 1967), lo cual sugiere la poca sensibilidad de la SDS al cambio, dado que el MMPI-D se presenta como una medida de rasgo de personalidad, no de estado de ánimo actual.

Hay ciertos problemas respecto a la validez general de la escala. En primer lugar, la SDS es una escala no lo bastante sensible al cambio terapéutico (Hedlund y Vieweg, 1979b; Hamilton y Shapiro, 1990). En segundo lugar, el obstáculo más problemático de la escala es su dudosa validez de contenido. Por ejemplo, no tiene apenas ítems psicológicos, de entrecimiento, culpa, etc., lo que hace poco recomendable su uso (Snaith, 1987).

Utilidad. Aunque la SDS es fácil de aplicar y corregir, su uso parece cada vez menor debido a los problemas señalados. Una de las limitaciones de la SDS es que evalúa frecuencia de síntomas y no intensidad. Así pues, un ítem de gran intensidad pero que ha sido poco frecuente, obtendría una puntuación baja. De hecho, diversos estudios (Biggs y cols., 1978) han demostrado que es una escala inadecuada para valorar síntomas depresivos cuando la intensidad de los síntomas

es alta. Otro problema con la escala es que se le pide al sujeto que evalúe la frecuencia de síntomas de modo indeterminado, es decir, sin ningún marco temporal específico, lo cual incrementa el error potencial de la escala.

A pesar de estos marcados inconvenientes, merece señalarse que con esta escala se efectuó el denominado *Estudio Sociológico. Libro Blanco de la Depresión en España* (1982), un trabajo financiado por laboratorios, con el que se intentó efectuar una estimación general de la prevalencia de la depresión en nuestro país. Según los datos recogidos con este instrumento, un 7% de la población general urbana mayor de 15 años ($n = 3.000$ entrevistados) tendría un nivel de depresión entre moderada y grave (tabla 21-4).

Escala Hospitalaria de Ansiedad y Depresión

La HAD es una escala autoaplicada dirigida a evaluar síntomas de depresión y de ansiedad durante la pasada semana en pacientes con problemas médicos, no psiquiátricos (Zigmon y Snaith, 1983). Consiste en 14 ítems (7 de depresión y 7 de ansiedad) que evalúan la frecuencia del síntoma a través de una escala Likert de 4 puntos (valorables de 0 a 3). Por ejemplo, para el síntoma «me siento tenso o molesto», el sujeto ha de elegir entre cuatro opciones: todos los días, muchas veces, a veces, nunca. Una puntuación de 0-7 en cualquiera de las dos escalas se considera normal, 8-10 es una puntuación de «caso probable» y más de 11 se considera como caso de ansiedad o de depresión (tabla 21-4).

Dado que el instrumento se diseñó específicamente para su uso con pacientes hospitalizados con trastornos físicos, excluye la evaluación de síntomas somáticos que tienen una alta probabilidad de deberse a la enfermedad física existente.

Fiabilidad. La consistencia de la escala en términos del índice α de Cronbach en pacientes psiquiátricos es de 0,81 para la subescala de ansiedad y de 0,82 para la subescala de depresión (Tejero y cols., 1986).

Validez. En cuanto a la validez convergente de la HAD, muestra unas correlaciones moderadas con medidas de ansiedad-estado (en concreto, el STAI de Spielberger) y con medidas de depresión como el BDI: 0,71 y 0,74, respectivamente (Tejero y cols., 1986).

Uno de los problemas con la HAD es que, a pesar de que parece medir dos constructos diferenciados (ansiedad y depresión), la correlación entre ambas subescalas es efectivamente nula en algunos estudios (-0,04) (Aylard y cols., 1987) pero alta en otros (Bramley y cols., 1987; Anderson, 1993) aunque, como han demostrado Caro e Ibáñez (1992), estas diferencias

pueden depender del tipo de enfermedad de los pacientes evaluados. Pero, abundando en esta confusión las soluciones factoriales de la HAD no son claras y distan de ofrecer un resultado satisfactorio (Anderson, 1993).

Utilidad. La escala presenta algunos problemas. En primer lugar, la unidad de medida de la escala no es constante puesto que en algunos ítems se evalúa la intensidad del síntoma y en otros, como en el ejemplo citado, la frecuencia. Además, la evaluación de la depresión está bastante limitada: se centra básicamente en el área de la anhedonia (pérdida de placer), pues los autores consideran que esto es lo más distintivo en los pacientes con trastornos médicos (Snaith, 1987; Lloyd, 1990). Pero evidentemente, el síndrome depresivo no es sólo esto.

La ventaja de la HAD es que debido a su sencillez y facilidad de administración tiene utilidad como instrumento de cribado inicial (*screening*) en ambientes hospitalarios o para ofrecer una evaluación sencilla y general, aunque de relativa validez, de niveles de depresión y ansiedad (Tejero y cols., 1986). Además, la HAD puede que sea para los pacientes un instrumento más fácilmente comprensible que otros como el BDI (Robertson, 1989). No obstante, una evaluación más precisa de la depresión o de la ansiedad exigiría el empleo adicional de instrumentos diferentes a la HAD, que evalúen más en detalle esos constructos. De hecho, el empleo exclusivo de la HAD es muy probable que haga incurrir en muchos falsos positivos (Thompson, 1989c).

Escalas Analógicas Visuales

Estas escalas (Visual Analogue Scales, VAS) pueden ser utilizadas tanto por observadores como por los propios sujetos. Utilizadas desde los años veinte, fueron plenamente incorporadas a la psicología clínica por Zealley y Aitken (1969). La idea es muy simple: se trata de una línea con dos extremos (p. ej., desde «nada deprimido» hasta «más deprimido que nunca»), normalmente de 100 mm de longitud, en la que el sujeto ha de marcar el lugar en que se «sitúa» su estado de ánimo. La puntuación es la distancia desde el principio de la línea hasta donde se ha situado la marca. Aunque una VAS se podría emplear para cada ítem de una escala, en el caso de la depresión se suele emplear sólo para evaluar el estado de ánimo. A pesar de su aparente simpleza se trata de un método muy eficaz para evaluar diferentes estados de ánimo (depresión, ansiedad e irritabilidad), en especial cuando se trata de efectuar medidas muy repetidas (Aitken, 1969; Crawford-Little y McPhail, 1973). Por ejemplo, si

deseamos monitorizar diariamente el estado de ánimo de una persona y comprobar si existen fluctuaciones, estas sencillas escalas pueden ser el instrumento de elección si se desea disponer de un «termómetro» sensible del estado de ánimo (Beck y cols., 1983; Burns, 1990).

Fiabilidad. La fiabilidad test-retest a las 24 horas con pacientes es próxima a 0,70 (Folstein y Luria, 1973).

Validez. Evidentemente estas escalas son muy inespecíficas (es decir, no están ligadas a ningún cuadro diagnóstico concreto). No obstante, es sorprendente su buen comportamiento respecto a escalas más complejas. Por ejemplo, se ha hallado que la correlación de las VAS con la puntuación total del BDI es de aproximadamente 0,70 en universitarios españoles (Vázquez y Ring, 1993), una cifra semejante e incluso algo inferior a la hallada en otros estudios (Thompson, 1989b). Su correlación con la puntuación total de la escala de depresión de Hamilton es de 0,78 (Zealley y Aitken, 1969).

Utilidad. Las moderadamente altas correlaciones de las VAS con otros instrumentos no deben hacer nunca perder de vista sus limitaciones como procedimientos de evaluación. Son instrumentos muy inespecíficos y con un valor diagnóstico nulo. No obstante, pueden desempeñar un interesante papel por su economía, sencillez y complemento de información (McCormack y cols., 1988; Bech y cols., 1976). Por ejemplo, las VAS se utilizan de modo estándar si desea monitorizar con frecuencia el estado de ánimo diario, como suele hacerse en protocolos de terapia cognitivo-conductual.

Listas de adjetivos

Desde la década de los sesenta, se han venido desarrollando una serie de listas de adjetivos para medir diversos estados emocionales transitorios (Ávila y Giménez, 1991; Fernández-Ballesteros, 1992b). Uno de los instrumentos más conocidos es la Lista de Adjetivos Depresivos de Lubin (Depression Adjective Checklist, DACL) consistente en siete listas de unos 30 adjetivos (positivos y negativos) relacionados con la depresión.

Hay varias versiones. Una muy utilizada en psiquiatría es la de Zersen (Zersen y cols., 1974, 1985) consistente en 28 pares de adjetivos, teniendo que elegir el sujeto, el adjetivo de cada par que mejor refleje su estado actual (p. ej., alegre-triste).

Fiabilidad y validez. Aunque tienen una alta consistencia interna, en torno a 0,90 (Tanaka-Matsumi y

Kameoka, 1986), su validez es más bien cuestionable: sus correlaciones con instrumentos como los anteriormente mencionados no suelen superar el rango de 0,50-0,60 (Byerly y Carlson, 1982). De todos los instrumentos citados en este capítulo, las listas de adjetivos son las que manifiestan una menor validez concurrente con otras escalas de depresión o con juicios clínicos de gravedad: las correlaciones oscilan entre 0,25 y 0,58 (Shaw y cols., 1985). De hecho, las propuestas más modernas plantean que estos instrumentos valoran, simplemente, una dimensión anímica general de negatividad/positividad (Gotlib y Meyer, 1986).

Utilidad. Más que evaluar síndromes depresivos o de cualquier otra psicopatología, evalúan estados de ánimo de un modo sencillo y, por lo tanto, son muy inespecíficos (Vázquez y Lozoya, 1994). Con estas escalas no se le pide al sujeto que evalúe de algún modo sus síntomas sino que se le pide elegir aquellos adjetivos de una lista que se adecuen más a su estado de ánimo actual.

Estas escalas, al igual que las VAS, son bastante sensibles a pequeñas fluctuaciones en el estado de ánimo y aquí es donde pueden tener su mayor utilidad.

INSTRUMENTOS DE EVALUACIÓN DE LA MANÍA

El creciente empleo del litio desde los años setenta para el control de los trastornos bipolares ha favorecido el interés en desarrollar escalas que midan cambios en los síntomas maníacos tras la medicación. Además de la solución obvia de efectuar un listado de síntomas basado en clasificaciones como el DSM-IV, se pueden utilizar algunos instrumentos específicos para la evaluación de la manía aunque, por desgracia, son bastante escasos y normalmente heteroaplicados. Las escalas autoaplicadas no son muy útiles, ya que un paciente en fase maníaca suele estar agitado, hiperactivo, y con una escasa conciencia (*insight*) de su problema. En cualquier caso, los instrumentos existentes se basan en pocos pacientes y el trabajo de validación de las mismas no ha sido todo lo intenso que cabía esperar.

Escala de Estado Maníaco

La Escala del Estado Maníaco (Mania State Rating Scale, MSRS) es un instrumento de 26 ítems creado por Beigel y cols. (1971). Para cada ítem se evalúa por separado frecuencia y gravedad del síntoma mediante dos escalas Likert (de 1 a 5) multiplicándose las puntuaciones en las dos escalas, de modo que la puntuación

total para cada ítem oscila entre 0 y 25. La MS se diseñó como una escala de observación hospitalaria, requiere cierto entrenamiento, y el paciente es evaluado tras observar su conducta durante un día (unas 8 horas). La escala incluye dos elementos depresivos (ítems 1 y 14) puesto que el estado de ánimo deprimido no es infrecuente en estos pacientes. De hecho, la correlación entre estos ítems y la puntuación global en la escala es bastante elevada (Murphy y Beigel, 1974), lo que confirma que la manía no es lo opuesto a la depresión. Con el fin de mostrar el tipo de ítems que se suele incluir en las escalas de manía, se exponen los contenidos en la escala MS:

1. Parece deprimido.
2. Está hablando.
3. Se mueve de un sitio a otro.
4. Efectúa amenazas.
5. Tiene poco juicio.
6. Se viste inadecuadamente.
7. Parece feliz o dichoso*.
8. Busca la compañía de otros*.
9. Está distraible*.
10. Tiene ideas de grandiosidad.
11. Está irritable.
12. Está combativo o destructivo.
13. Está delirante.
14. Verbaliza sentimientos depresivos.
15. Está activo.
16. Está discutiendo.
17. Habla de temas sexuales.
18. Está enfadado.
19. Descuida su aspecto personal.
20. Tiene un control de impulsos disminuido*.
21. Verbaliza sentimientos de bienestar.
22. Está suspicaz.
23. Hace planes irreales.
24. Demanda el contacto con otros*.
25. Está preocupado sexualmente.
26. Salta de un tema a otro*.

Fiabilidad y validez. La escala tiene índices elevados de fiabilidad interexaminador; por ejemplo, la fiabilidad interexaminador del estudio original de la MS osciló entre 0,86 y 0,99. Así mismo, tiene buenos índices de validez concurrente con criterios de expertos (la puntuación total de la MS con las puntuaciones globales de los psiquiatras fue de 0,96 en el estudio original), y de sensibilidad al cambio.

Utilidad. Probablemente es una escala demasiado larga, poco clara en cuanto a sus normas de utilización, y con ítems bastante inespecíficos (Thompson, 1989b). En este sentido, Bech y cols. (1975) han demostrado que sólo 6 ítems (los indicados con un asterisco

en la lista anterior) pueden ser suficientes para obtener una puntuación sensible a diferentes niveles de intensidad.

Escala de Evaluación del Estado Maníaco-Modificada

La MS fue originalmente diseñada para personal de enfermería. Existe una versión posterior para psicólogos y psiquiatras: la Escala de Evaluación del Estado Maníaco-Modificado (MMS) de Blackburn y cols. (1977). Incluye 5 nuevos ítems (problemas de sueño, religiosidad, alucinaciones, desinhibición y labilidad del estado de ánimo) ignorados en la MS y se han condensado tres pares de los ítems originales (ira/combatividad; hablar sobre sexo/preocupación con el sexo, y aspecto deprimido/verbalización deprimido). La MMS evalúa sólo la gravedad de los síntomas.

Escala de Evaluación de la Manía de Bech-Rafaelsen

La Mania Rating Scale es posiblemente el instrumento más utilizado en la actualidad para la evaluación de la manía (Bech y cols., 1978). Se trata de un instrumento heteroaplicado de sólo 11 ítems, con definiciones precisas para evaluar la gravedad de cada uno de ellos, desde 0 (no presente) hasta 4 (grave/extremo).

Fiabilidad y validez. Es una escala bastante homogénea: correlaciones ítem-total entre 0,72 y 0,94 (con la excepción del sueño: 0,48), y una muy buena fiabilidad interexaminador (superior a 0,90). Es una escala que parece sensible al cambio y con una aceptable correlación con indicadores de gravedad global otorgados por clínicos expertos.

Utilidad. Dado que el trabajo psicométrico con la escala ha sido relativamente escaso, debería emplearse con ciertas cautelas.

VARIABLES COGNITIVAS RELACIONADAS

Además del estado general del paciente, existen instrumentos mucho más específicos, destinados a evaluar diversas áreas de funcionamiento. Quizá los instrumentos adicionales más empleados son los que evalúan variables cognitivas aunque, obviamente, la evaluación de los factores estresantes, apoyo social, etc., ayudan a completar un protocolo de evaluación

TABLA 21-5. Características generales de algunos de los instrumentos más utilizados para evaluar productos cognitivos en la depresión

Escala	Referencia	Comentarios
<i>Pensamientos automáticos</i> Cuestionarios de Pensamientos Automáticos (Automatic Thoughts Questionnaire: ATQ-30 y ATQ-R-40)	Hollon y Kendall (1980) y Kendall y cols. (1989)	30 pensamientos negativos automáticos. Evaluación en 5 puntos de la frecuencia durante la última semana Muy utilizada en práctica clínica e investigación El ATQ-R es idéntico pero incluye 10 frases positivas
Cuestionario de Pensamientos Molestos (Distressing Thoughts Questionnaire)	Clark (1986, 1992)	6 pensamientos ansiosos y 6 negativos (basados en el ATQ). En escalas Likert de 9 puntos, el sujeto ha de evaluar para cada uno: frecuencia, tristeza, preocupación, intrusividad y desaprobación
Inventario de Cogniciones de Crandell (Crandell Cognitions Inventory, CCI)	Crandell y Chambless (1986)	34 frases depresivas (con cuatro temas: indefensión, desesperanza, inferioridad y desapego) y 11 positivas. Se evalúa su frecuencia en una escala de 5 puntos
<i>Atribuciones causales</i> Cuestionario de Estilo Atribucional (Attributional Style Questionnaire, ASQ)	Peterson y cols. (1982)	12 situaciones hipotéticas (6 de logro y 6 de dependencia) para evaluar las atribuciones causales del paciente en tres dimensiones: internalidad, globalidad y estabilidad. Más utilizada en investigación que en práctica clínica
Cuestionario de Estilo Atribucional Expandido (Attributional Style Questionnaire)	Peterson y Villanova (1988)	24 situaciones hipotéticas negativas. Formato idéntico al ASQ
<i>Expectativas negativas</i> Escala de Desesperanza (Hopelessness Scale, HS)	Beck y cols. (1974)	21 ítems de verdadero/falso respecto a actitudes del cliente sobre el futuro. Muy utilizada. Buen predictor de suicidio
<i>Distorsiones cognitivas</i> Escala de Actitudes Disfuncionales (Dysfunctional Attitudes Scale, DAS)	Weissman y Beck (1978)	Grado de acuerdo (de 0 a 6) con 40 pensamientos generales absolutistas basados en Beck y Ellis. Muy utilizada en clínica e investigación
Cuestionario de Cogniciones (The Cognitions Questionnaire)	Fennell y Campbell (1984)	8 situaciones complejas de la vida. El sujeto responde a una serie de preguntas sobre cada una. No muy utilizado
Test de Estilos Cognitivos (Cognitive Styles Test)	Blackburn y cols. (1986)	30 situaciones de la vida. Mide errores de pensamiento negativo en el cliente
Test de Respuesta Cognitiva (Cognitive Response Test)	Watkins y Rush (1983)	38 frases incompletas. El sujeto las completa y se cuentan las «respuestas negativas irracionales» emitidas
Cuestionario de Sesgo Cognitivo (Cognitive Bias Questionnaire, CBQ)	Krantz y Hammen (1979)	6 situaciones problemáticas para jóvenes universitarios (3 sociales y 3 de logro)
Inventario de la Tríada Cognitiva (Cognitive Triad Inventory, CTI)	Beckham y cols. (1986)	Grado de acuerdo con 30 afirmaciones basadas con la tríada cognitiva de Beck: visión negativa de uno mismo, el mundo y el futuro
Cuestionario de Creencias (Belief Scale, BS)	Malouff y Schutte (1986)	20 afirmaciones que recogen los 10 tipos de creencias irracionales propuestas por Ellis
Test de Creencias Irracionales (Irrational Beliefs Test, IBT)	Jones (1969)	100 ideas irracionales basadas en Ellis. El sujeto señala el grado de acuerdo con cada una de ellas (de 1 a 5)

De Yáñez, 1995.

completo. En este apartado se efectúa un breve repaso de algunos de los que consideramos más útiles (véase una revisión más a fondo en, por ejemplo, Clark, 1988; Blackburn, 1988; Hammen y Krantz, 1985; Freeman y Munro 1990) (tabla 21-5).

Si pretender efectuar una revisión conceptual de los constructos que pretenden medir estos instrumentos (p. ej., pensamientos automáticos, desesperanza, ideas irracionales, etc.), es conveniente subrayar dos aspectos:

1. La inmensa mayoría de estos instrumentos evalúan productos cognitivos verbales, pero no evalúan ni operaciones mentales ni estructuras cognitivas (Sanz y Vázquez, 1991a).

2. Aunque para investigar los productos cognitivos que tienen las personas deprimidas existen diversos procedimientos de evaluación como las técnicas de muestreo del pensamiento, de pensar en voz alta, de listado de pensamientos, etc. (Kendall y Hollon, 1981), la estrategia más utilizada suele ser el empleo de cuestionarios de frases o autoafirmaciones (Clark, 1988).

Pensamientos automáticos

Cuestionario de Pensamientos Automáticos

Para evaluar «pensamientos automáticos» no hay un instrumento claramente mejor que otro, aunque el Cuestionario de Pensamientos Automáticos (Automatic Thoughts Questionnaire, ATQ-30) es probablemente el más utilizado. Contiene las 30 frases más comunes que una muestra grande de estudiantes universitarios ($n = 800$ estudiantes) indicó que aparecían de modo «automático» cuando se sentían deprimidos (Hollon y Kendall, 1980) por ejemplo «no valgo nada». Cada cognición es evaluada en términos de la frecuencia de aparición durante la semana pasada en una escala de 1 a 5, y la puntuación total es simplemente la suma de los 30 ítems. Es un instrumento de muy sencilla aplicación y con correlaciones en torno a 0,60 con la gravedad de la depresión evaluada con el BDI.

No obstante, es cuestionable si lo que mide son realmente pensamientos automáticos, en el sentido estricto de la psicología cognitiva, o más bien ideas de carácter rumiativo (Hartlage, Alloy, Vázquez y Dyckman, 1993). Aunque autores como Coyne y Gotlib (1983) han planteado que es difícil que alguien pueda estimar la frecuencia de pensamientos automáticos de una semana, lo cierto es que el ATQ diferencia bien entre deprimidos y no deprimidos, y es muy utilizado actualmente en la investigación básica y de eficacia de

tratamientos. Existe una versión más moderna, el ATQ-R (Kendall y cols., 1989), resultante de añadir 10 ítems que reflejan pensamientos automáticos positivos (Comeche y cols., 1995).

Ideas irracionales

Escala de Actitudes Disfuncionales

La Escala de Actitudes Disfuncionales (Dysfunctional Attitudes Scale, DAS) consiste en una escala de 40 ítems en la que se pide al sujeto que señale en qué medida está de acuerdo (en una escala de 7 puntos) con cada una de las frases que se presentan (Vázquez y Ring, 1993; Sanz y Vázquez, 1993). Cada ítem recoge aquellas ideas próximas al concepto de Ellis de «idea irracional» y de «pensamientos disfuncionales» de Beck que se supone son, según las teorías cognitivas de la depresión (Sanz y Vázquez, 1995), un factor de vulnerabilidad estable. Estos componentes de vulnerabilidad serían pensamientos o creencias falsas, absolutistas y psicológicamente dañinos (p. ej., «si fracaso en mi trabajo, soy un fracaso como persona»).

Es una escala con una buena consistencia interna (coeficientes α próximos a 0,90) fiabilidad test-retest (alrededor de 0,80 para un período de 8 semanas). Su correlación con el BDI oscila alrededor de 0,40. Sólo el 50-55 % de los pacientes que experimentan un episodio depresivo muestran durante el mismo una puntuación anormalmente elevada en la DAS (Norman y cols., 1983; Hamilton y Abramson, 1983), lo que pone de manifiesto que los elementos cognitivos evaluados por la DAS no están necesariamente presentes en todas las depresiones (Sanz y Vázquez, 1995). En cuanto a la especificidad de las creencias disfuncionales, los resultados no son concluyentes. Mientras que algunos estudios han encontrado que los pacientes depresivos manifiestan un mayor número de actitudes disfuncionales que otros pacientes psiquiátricos no deprimidos (Dobson y Shaw, 1986; Hamilton y Abramson, 1983), otros trabajos, por el contrario, no han conseguido encontrar diferencias significativas entre esos grupos (Hollon y cols., 1986; Silverman, Silverman y Eardley, 1984).

La DAS es una medida muy empleada en protocolos rigurosos de investigación sobre la eficacia diferencial de psicofármacos y psicoterapias en la depresión, por ejemplo, el del National Institute of Mental Health (Sotsky y cols., 1991). Una de sus propiedades es que es sensible al cambio terapéutico (Vázquez y Ring, 1993; Sanz y Vázquez, 1993), aunque esto plantea algunos problemas teóricos dado que, según la teoría de Beck, los pensamientos disfuncionales serían una característica relativamente estable de las personas, independientemente de su estado de ánimo. Así pues,

parece cuestionable la utilidad de la DAS como instrumento para evaluar actitudes depresogénicas estables. Existe una amplia revisión de la DAS y de su factorización con muestras españolas en Sanz y Vázquez (1993 y 1994).

Cuestionario de Sesgo Cognitivo

El Cuestionario de Sesgo Cognitivo (Cognitive Bias Questionnaire, CBQ) fue diseñado para comprobar la hipótesis de Beck de que los depresivos muestran sesgos negativos al interpretar información referente a uno mismo, el mundo y el futuro (Krantz y Hammen, 1979). Se presentan 6 historias cortas con una serie de preguntas finales sobre cómo interpretar cada una de ellas y se le ofrecen al sujeto, para cada pregunta, cuatro opciones de respuesta: una sesgada-depresiva, una no sesgada-depresiva, una sesgada-no depresiva y una no sesgada-no depresiva. El objetivo de la prueba es confirmar si existe un sesgo interpretativo de la realidad y, en caso de que así sea, si es de tipo depresivo o no.

Test de Creencias Irracionales

El Test de Creencias Irracionales (Irrational Beliefs Test, IBT) (Jones, 1969) es un test basado en la teoría cognitiva de Ellis, quien propone que la asunción de ideas irracionales, absolutistas, es la causa básica de los desajustes emocionales. El IBT consiste en 100 frases de carácter irracional en las que el sujeto ha de señalar su grado de creencia en ellas en una escala de 7 puntos. Contiene 10 subescalas diferentes derivadas de un análisis factorial, que se supone que representan diferentes grupos de ideas irracionales. El principal trabajo con esta escala se ha efectuado con estudiantes, no con pacientes. Por otro lado, dado que su contenido es semejante a la DAS de Weissman, en general parece más recomendable utilizar esta última o bien la sencilla Escala de Creencias de Malouff y Schutte (1986), reproducida en Comeche y cols. (1995) (tabla 21-5).

Expectativas negativas

Escala de Desesperanza

La desesperanza es una variable de importancia crucial y, de hecho, tiene un papel clave causal en algunos modelos actuales de depresión (Sanz y Vázquez, 1995). El modo más habitual de cuantificarla es mediante esta escala de muy sencilla aplicación. La Escala de Desesperanza (Hopelessness Scale, HS) (Beck, 1974) consta de 20 afirmaciones que expresan predic-

ciones generales sobre el propio futuro del sujeto, en las que se le pide que señale si cree verdadero o falso cada ítem (p. ej., «no tiene sentido intentar en serio conseguir algo que deseo porque probablemente no lo conseguiré»). El rango de puntuación de la escala es de 0 a 20. Es muy frecuente administrar la HS antes, durante, y después del tratamiento.

La escala HS tiene buenas cifras de fiabilidad y la validez queda demostrada, entre otras cosas, por ser la mejor variable predictora existente del suicidio. De hecho, la desesperanza se correlaciona mejor con el suicidio que el nivel de depresión. Beck y cols. (1989) hallaron que una puntuación mayor de 9 predecía el 91 % de los casos de suicidio en una muestra de 165 pacientes hospitalizados. Los estudios empleando la HS han demostrado que las personas deprimidas obtienen una puntuación significativamente más alta que los sujetos no deprimidos (Wilkinson y Blackburn, 1981; Hamilton y Abramson, 1983; Blackburn, Jones y Lewin, 1986; Dohr, Rush y Bernstein, 1989) o que los pacientes psiquiátricos no deprimidos (Wilkinson y Blackburn, 1981; Hamilton y Abramson, 1983).

Atribuciones causales

Cuestionario de Atribuciones Causales

El Cuestionario de Atribuciones Causales (Attributional Style Questionnaire, ASQ) (Peterson y cols., 1982) es la prueba más utilizada para evaluar cómo explica el sujeto lo que le sucede o, en otras palabras, el estilo atribucional. Se trata de 12 situaciones hipotéticas, seis positivas y seis negativas, cada una de las cuales contiene tres temas de logro y tres interpersonales. Para cada situación, el sujeto ha de pensar cuál sería la causa más probable de ese hecho si le ocurriese a él y entonces ha de valorar esta causa en varias dimensiones empleando un formato Likert de 7 puntos: internalidad (hasta qué punto esa causa es debida a uno mismo o a otras personas o circunstancias), globalidad (hasta qué punto esa causa influye en otras áreas de la vida de uno), y estabilidad (en qué medida esa causa volvería a estar presente en el futuro ante una situación semejante). Con este cuestionario se pretende comprobar el estilo atribucional de los sujetos puesto que según la teoría de la indefensión aprendida un «estilo atribucional negativo» (p. ej., hacer atribuciones causales internas, estables y globales para sucesos negativos) hace al sujeto especialmente vulnerable a la depresión (Peterson y Seligman, 1984; Vázquez y Sanz, 1991a).

Las cifras de consistencia interna de cada una de las tres dimensiones es realmente baja (α en torno a

0,50), pero estas cifras mejoran cuando se combinan las tres dimensiones para ofrecer un único índice de internalidad + globalidad + estabilidad para resultados positivos ($\alpha = 0,75$) y para resultados negativos ($\alpha = 0,70$). No obstante, esta combinación puede que no sea psicométrica y conceptualmente muy correcta (Hammen y Krantz, 1985).

Decenas de estudios han analizado la relación entre depresión y atribuciones causales medidas con la ASQ. En un metaanálisis de Sweeney y cols. (1986) demostraron que los resultados de los 104 trabajos publicados hasta esa fecha avalaban efectivamente la existencia en las personas deprimidas de un estilo atribucional interno-global-estable para sucesos negativos, aunque la correlación media no supera 0,30. Sin embargo, la existencia de un estilo depresivo externo-específico-inestable para sucesos positivos (p. ej., «he sacado sobresaliente sólo porque he tenido mucha suerte en el examen») no parece ser una característica de las personas deprimidas. Es decir, para sucesos positivos, parece que las atribuciones de las personas deprimidas no difieren de las que efectúan los no deprimidos. Ahora bien, la evidencia empírica sobre el papel causal de las atribuciones en la depresión es todavía bastante endeble (Sanz y Vázquez, 1991a). En nuestra experiencia, la ASQ parece útil para identificar determinados sujetos vulnerables a futuras depresiones (Vázquez Avia y cols., en preparación). Una nueva versión ampliada de la ASQ (Peterson y Vilanova, 1988) está reproducida en Comече y cols. (1995).

VARIABLES CONDUCTUALES RELACIONADAS

Las medidas conductuales de la depresión son un tema a menudo olvidado, incluso en libros específicos dedicados a la evaluación de la depresión (Sartorius y Ban, 1986). El énfasis clínico en los aspectos afectivos del cuadro y, por otro lado, el desarrollo de teorías cognitivas aparentemente con más poder predictivo, ha hecho olvidar los aspectos más conductuales del cuadro (Jacobson, 1981; Rehm, 1987). Quizás esto se deba a que los modelos estrictamente conductuales de la depresión están en franco retroceso dentro del panorama actual (Sanz y Vázquez, 1995). Desde el punto de vista diagnóstico, aunque los elementos estrictamente conductuales puede que no sean muy relevantes, desde el punto de vista de la planificación del tratamiento son esenciales (Beck y cols., 1983; Bas y Andrés, 1993). No obstante, en general falta por investigar tanto su posible utilidad diagnóstica como su valor para medir la eficacia de los tratamientos.

Actividad general

Greden y Carroll (1981) han indicado que sería deseable prestar más atención a la evaluación de los aspectos psicomotores de la depresión pues esto puede ayudar a mejorar las clasificaciones al uso y la monitorización del tratamiento. Variables como el gesto, la postura, los movimientos de cabeza, etc., han recibido alguna atención de los investigadores de la depresión pero no hay datos sistemáticos relevantes.

En general, con la excepción del sueño, hay una gran carencia de investigación en el área de la actividad conductual (p. ej., enlentecimiento motor, sexualidad, ingesta y apetito, etc.). Como señala Jacobson (1981), apenas existen estudios observacionales de estas variables en el ambiente natural de los pacientes, para comprobar si las quejas de éstos están correlacionadas con déficit apreciables por observadores. No obstante, parece que un incremento en la inmovilidad (incluso durante el sueño) parece ser una de las razones básicas de la mayor inactividad durante la depresión (Widlocher y Gozlan, 1989).

Los pacientes deprimidos y los no deprimidos parece que no se diferencian tanto por el nivel de actividad general que efectúan sino por el número de actividades agradables y desagradables que llevan a cabo (Lewinsohn y Amenson, 1978). En este sentido, dentro de las teorías conductuales de la depresión, quizá los instrumentos más interesantes fueron los desarrollados inicialmente por Lewinsohn para evaluar la frecuencia y disfrute de actividades reforzantes. Estas escalas pueden tener utilidad en el diseño de las terapias psicológicas con los pacientes y en la monitorización del tratamiento. Hay dos escalas básicas:

Inventario de Sucesos Agradables

El Inventario de Sucesos Agradables (Pleasant Events Schedule) (MacPhillamy y Lewinsohn, 1971) es una heterogénea lista de 320 actividades relativamente comunes, sociales o solitarias, que suelen ser agradables para muchas personas: ver viejos amigos, coser, estar con los hijos, ir al cine, etc. Respecto a la puntuación de la escala, para cada actividad, el sujeto ha de señalar con qué frecuencia la ha realizado durante los últimos 30 días (0, no ha sucedido; 1, ha sucedido poco; entre 1 y 6 veces; 2, ha sucedido a menudo; 7 o más veces) y en qué medida ha disfrutado con ella (0, no fue agradable; 1, fue algo agradable; 2, fue muy agradable). Se puede además obtener un índice de «placer obtenido» multiplicando la puntuación en frecuencia por la de disfrute. Aunque existen algunas normas para personas norteamericanas (MacPhillamy y Lewinsohn, 1982), no existe ningún estudio de tipi-

ficación española (véase una traducción en Bas y Andrés, 1994), si bien esto no impide su uso clínico individual con pacientes. Aunque existe una versión reducida de la escala con las 49 actividades que correlacionan más con el estado de ánimo (Lewinsohn y cols., 1979), la versión más utilizada es la de 320 ítems.

La fiabilidad test-retest de la escala es aceptable: aproximadamente 0,80 para un mes, y 0,60 para los 3 meses (MacPhillamy y Lewinsohn, 1982). En cuanto a los estudios de validez, son bastante escasos y realizados con pocos pacientes. Respecto a su validez de constructo, tiene una correlación aceptable con escalas de estado de ánimo y diferencia bien entre deprimidos y no deprimidos (estos últimos se implican en más actividades positivas reforzantes). No obstante, la existencia de diferencias entre deprimidos y no deprimidos no debe llevarnos a concluir que una baja tasa de reforzadores sea la causa de la depresión (Vázquez y Sanz, 1991a). Por el contrario, parece que el hecho de que los pacientes obtengan menos reforzadores está causado por elementos cognitivos tales como expectativas disminuidas, sensación de indefensión, etc. (Hammen y Krantz, 1985). En todo caso, éste es un instrumento de interés terapéutico (Lewinsohn, 1976; Beck y cols., 1983; Burns, 1990; Bas y Andrés, 1994).

Inventario de Sucesos Desagradables

Lewinsohn propuso que los depresivos no sólo experimentan y/o disfrutan menos experiencias positivas sino que además son más sensibles ante sucesos aversivos. Ésta es la razón por la que se creó el Inventario de Sucesos Desagradables (Lewinsohn y Talkington, 1979). Consta también de 320 actividades negativas y se puntúa de modo semejante al anterior. Tiene una buena consistencia interna (α , 0,96 para la puntuación total) y una buena fiabilidad test-retest (0,52 al mes y 0,62 a los 3 meses). Aunque es capaz de diferenciar entre normales y depresivos, las diferencias no están tanto en el número de experiencias aversivas que ambos dicen experimentar sino en cómo juzgan (un elemento cognitivo) su aversividad. Uno de los problemas conceptuales de la escala es que incluye tipos de estresores muy diferentes (agudos, crónicos, continuos y de la vida cotidiana), y esto hace que se pierda discriminabilidad dado que no todos estos factores estresantes tienen un papel equiparable en la etiología de la depresión (Vázquez, Crespo y Ring, en este mismo volumen; Brown, 1989). Por último, una crítica de estos instrumentos es que se tratan, en realidad, de autoinformes retrospectivos más que de instrumentos observacionales. Por lo tanto, hay que considerarlos sujetos a posibles distorsiones de memoria (Jacobson, 1981).

VARIABLES FISIOLÓGICAS RELACIONADAS

La depresión, como el resto de los trastornos mentales, no es una enfermedad en sentido estricto, al menos con los datos de que se dispone hasta ahora. No es de extrañar la ausencia de signos somáticos con valor diagnóstico en los sistemas de clasificación actuales. Sin embargo, existe una ingente cantidad de investigación biológica intentando buscar características somáticas diferenciadoras de los trastornos afectivos (Gastó y Vallejo, 1990). Se han empleado muchas medidas diferentes para intentar desentrañar diferencias sustanciales entre deprimidos y no deprimidos: diferencias en vascularización hemisférica cerebral, conductancia de la piel (García y cols., 1992), secreción de saliva, tensiones anómalas en músculos frontales (Teasdale y Rezin, 1978), etc. Los resultados muestran que todavía no existe ninguna variable fisiológica lo suficientemente sensible y específica como para que sea utilizable diagnósticamente (Blackburn, 1988). No obstante, existe una especial investigación sobre dos potenciales marcadores diagnósticos (Hamilton y Shapiro, 1990):

1. *Trastornos específicos del sueño.* Aunque varios parámetros pudieran estar afectados, en general, los resultados más robustos hacen referencia a trastornos en el sueño REM (fundamentalmente aparición más rápida de la primera fase REM y mayor tiempo en fase REM, en especial en el primer tercio de la noche). Estos trastornos no son generales: es más frecuente observarlos en determinados subgrupos de pacientes (los de más edad, los que presentan características psicóticas o los más endógenos). No obstante, los resultados aún no son nada definitivos y bastante contradictorios entre estudios que han seguido la misma metodología.

2. *Test de la supresión de dexametasona (DST).* El DST va dirigido a comprobar la actividad del sistema límbico, el hipotálamo, la hipófisis y las glándulas suprarrenales en la depresión. La inyección de dexametasona hace disminuir la actividad de este complejo sistema adrenocortical en personas normales. Sin embargo, en algunos deprimidos esta inyección no produce la disminución de dicha actividad, lo que sugiere que podría haber una hiperactividad del sistema en dichos sujetos. Pero el problema fundamental del DST es su inespecificidad. En efecto, aunque es bastante eficaz para diferenciar deprimidos de normales, tiene resultados pobres (da muchos falsos positivos) cuando se trata de personas con diabetes, hipotiroidismo, infecciones, alto estrés, manía, demencia, psicosis agudas, o alcoholismo, por citar sólo algunas condiciones

(Hamilton y Shapiro, 1990; Arana y Mossman, 1988; Braddock, 1986).

En definitiva, es posible que los parámetros del sueño y el DST puedan tener importancia en la investigación de la depresión (para monitorizar tratamientos o para cuantificar gravedad), pero aún es muy prematuro el intento de utilizarlos como marcadores diagnósticos.

CONCLUSIONES

Hasta aquí se ha efectuado un recorrido aparentemente transparente sobre algunos de los instrumentos más importantes para la evaluación de la depresión. Sin embargo, quizás hemos hurtado parte del debate de fondo: ¿qué miden realmente los instrumentos? Esta pregunta está íntimamente relacionada con el problema de la escasa distinción entre emociones (ansiedad, depresión, desmoralización, etc.) que ya se apuntó al principio.

Cuando se analizan factorialmente diversas escalas de psicopatología, el panorama resultante es más confuso de lo que cabría ingenuamente esperar. Por ejem-

plo, Gotlib (1984) efectuó un análisis factorial de segundo orden con 17 escalas psicopatológicas completadas por estudiantes universitarios y halló un factor general de malestar (distrés) en el que saturaban muchos ítems de instrumentos de valoración de depresión (como el BDI) y de valoración de la ansiedad (como el STAI). Este primer factor explicaba la mitad de la variancia. Este hallazgo de un factor común subyacente de malestar, en el que intervienen ítems de escalas que miden constructos o emociones supuestamente diferentes, se ha hallado igualmente en estudios con pacientes psiquiátricos (Luteijn y Bouman, 1988). Por consiguiente, resulta difícil aceptar *prima facie* que los instrumentos al uso miden de modo «puro» categorías ideales. Se necesita mucha investigación básica para poder salir de esta situación de freno al avance de la comprensión científica de la psicopatología de la depresión y trastornos relacionados.

Un ejemplo de los caminos a seguir puede ser el de la investigación sobre afectos negativos y positivos (Sanz y Vázquez, 1992). Por ejemplo, en el ámbito de la investigación clínica experimental parece cada vez más claro que si bien emociones como la ansiedad y la depresión comparten en efecto elementos de disforia negativa (y de ahí la desconcertante alta correlación entre escalas de ansiedad y depresión, focalizadas en

TABLA 21-6. Características de los principales instrumentos específicos para evaluar la depresión en ancianos

Instrumento	Número de ítems	Formato de administración	Tipo de respuesta	Utilidad	Ítems ejemplo	Comentarios
<i>Entrevistas</i>						
CARE (Gurland y cols., 1977)	314 ^a	Entrevista	Presencia/ausencia del ítem Heterovalorada	+++	¿Ha estado triste o deprimido durante el mes pasado?	Buenos índices psicométricos Motivante para el sujeto Poca especificidad para depresión
OARS (Duke University, 1978)	105 (18 de depresión)	Entrevista	Escala Likert Heterovalorada	++	¿Es usted feliz la mayor parte del tiempo?	Buenos índices psicométricos Subescala de depresión (ODS) Interpretación algo difícil
<i>Cuestionarios</i>						
GDS (Brink y cols., 1982; Yesavage y cols., 1983)	30 en la versión original, 20 en la española (González y Szurek, 1988, 1990)	Autoinforme en la original Entrevista estructurada en la española	Si/No en la original Grado de acuerdo (de 0 a 3) en la española	+++	¿Está usted satisfecho con su vida?	Buenos índices psicométricos No hay síntomas somáticos Sencilla y económica

^a Existe una versión abreviada (Gurland y cols., 1984).

Utilidad: + escasa; ++ mínima; +++ buena; ++++ muy buena.

CARE: Comprehensive Assessment and Referral Evaluation; GDS: Geriatric Depression Scale, OARS: Olders American Resources and Services Questionnaires.

De Vázquez y Lozoya, 1994.

el afecto negativo, quizá son dos emociones que en realidad difieren en la esfera del afecto positivo que muestran los sujetos: los deprimidos se diferencian claramente de los ansiosos por mostrar un afecto positivo bajo (Watson y Kendall, 1989).

Dejando de lado estos breves apuntes de elementos más teóricos quisiéramos terminar con una breve lista de conclusiones finales sobre el uso de instrumentos diagnósticos en la práctica:

1. Una evaluación diagnóstica completa de la depresión requiere tanto un diagnóstico categorial (p. ej., DSM-IV) como una cuantificación dimensional a través de escalas, inventarios, etc.

2. No existe un protocolo ideal de medida. De hecho, los instrumentos de evaluación dimensional de la depresión presentan intercorrelaciones bastante elevadas, en torno a 0,60, 0,70, a pesar de su aparente heterogeneidad en cuanto al tipo de síntomas que cubren, el marco temporal que utilizan, etc. (Lambert y cols., 1986; Schaefer y cols., 1985). Así pues, en general, la elección de una medida u otra va a depender más de preferencias y necesidades personales que de acentuadas diferencias psicométricas (tablas 21-2 y 21-3). En palabras de Hamilton (1976): «no puede existir una escala que sea mejor para todos los propósitos, en todas las circunstancias, y que sea utilizable por cualquier tipo de evaluador». Por lo tanto, una elección razonable podría incluir, si es posible, un instrumento de heterovaloración como la HDRS de Hamilton y una escala autoaplicada como el BDI de Beck. En todo caso, la elección depende de varios parámetros:

a) *Contexto de la evaluación.* Por ejemplo, en pacientes hospitalizados o con enfermedades físicas pueden ser adecuados instrumentos como la HAD, a pesar de las limitaciones que hemos señalado. Además, factores como una excesiva longitud del protocolo puede hacer aconsejable emplear instrumentos sencillos y fáciles (p. ej., VAS) aun no siendo psicométricamente idóneos.

b) *Edad del paciente.* Existen escalas específicas destinadas a evaluar la depresión en niños tanto en términos de escalas e inventarios (Polaino, 1988; del Barrio, 1992; Pedreira y cols., 1993) como de entrevistas estructuradas similares a las SADS o DIS de los adultos (Ezpeleta, 1995; Maganto, 1995). Así mismo, la evaluación de la depresión en ancianos puede exigir instrumentos diferentes a los empleados en adultos jóvenes (Vázquez y Lozoya, 1994) (tabla 21-6) aunque en este último caso es frecuente emplear también instrumentos normales como el BDI, la CES-D.

c) *Tipo de paciente.* Por ejemplo, es más adecuado emplear escalas de observación en pacientes maníacos. Así mismo, cuando se desea evaluar cambio terapéutico, hay que extremar el cuidado en la selección pues

no todos los instrumentos tienen la misma sensibilidad para diferentes aspectos del posible cambio (cognitivo, motor, vegetativo, etc.). Así mismo, como hemos analizado, las escalas difieren en el tipo de síntomas que evalúan, el marco temporal que utilizan, etc.

3. Es recomendable el empleo de otros instrumentos que evalúen áreas de interés, en especial aquellas más relacionadas teóricamente con la depresión (desesperanza, apoyo social, cogniciones, estilos de afrontamiento, reforzadores, *coping*, ansiedad, etc.) para efectuar una evaluación completa y con sentido clínico y terapéutico.

BIBLIOGRAFÍA

- Aitken RC. Measurement feeling using visual analogue scales. *Proceedings of the Royal Society of Medicine* 1969; 62: 989-993.
- Allerup P. Statistical analysis of MADRS-a rating scale. Copenhagen: The Danish Institute for Educational Research, 1986.
- American Psychiatric Association. DSM-IV. Manual diagnóstico y estadístico de los trastornos mentales. Barcelona: Masson, 1994.
- Arana GW, Mossman D. The dexamethasone suppression test: approaches to the use of a laboratory test in psychiatry. *Neurol Clin* 1988; 6: 21-39.
- Asberg M, Montgomery S, Perris C, Schalling D, Tuck J. A comprehensive psychopathological psychiatric rating scale (CPRS). *Acta Psychiatr Scand* 1978; 271 (Supl.): 5-27.
- Ávila A, Giménez A. Los adjetivos en tareas de evaluación psicológica: Propiedades y valor estimular. *Rev Psicol Gen Aplic* 1991; 44: 465-475.
- Aylard PR, Gooding J, McKenna PJ, Snaith RP. A validation study of three anxiety and depression self-assessment scales. *J Psychosom Res* 1987; 31: 261-268.
- Bailey J, Copen A. A comparison between the Hamilton Rating Scale and the Beck Inventory of Depression. *Br J Psychiatry* 1976; 128: 486-489.
- Bas F, Andrés V. Tratamiento cognitivo-conductual de la depresión: Un manual de tratamiento. Madrid: Fundación Universidad Empresa, 1993.
- Bech P. Rating scales for affective disorders: Their validity and consistency. *Acta Psychiatr Scand* 1981; 295 (Supl. 64): 1-101.
- Bech P. Construct validity of rating scales for melancholia. En: Hindmarch I, Stonier PD (eds.). *Human psychopharmacology: Measures and methods.* (Vol. 1). New York: Wiley, 1987.
- Bech P, Bolwig TG, Kramp P, Rafaelsen OJ. The Bech-Rafaelsen Mania Scale and the Hamilton Depression Rating Scale. *Acta Psychiatr Scand* 1979; 59: 420-430.
- Bech P, Clemmensen L. The diagnosis of depression: 20 years later. *Acta Psychiatr Scand* 1983; (Supl. 310): 9-30.
- Bech P, Kastrup M, Rafaelsen OJ. Minicompendium of rating scales for states of anxiety depression, mania, schizophrenia, with corresponding DSM-III syndromes. *Acta Psychiatr Scand* 1986; 73 (Supl. 326): 1-35.
- Bech P, Rafaelsen OJ, Kramp P, Bolwig TG. The Mania Rating Scale construction and interobserver agreement. *Neuropharmacol* 1978; 1: 430-431.

- Bech P, Rafaelsen OJ. The use of rating scales exemplified by comparison of the Hamilton and the Bech-Rafaelsen melancholia scale. *Acta Psychiatr Scand* 1980; 62 (Supl. 285): 128-132.
- Beck AT, Beck RW. Screening depressed patients in family practice. *Postgraduate Medicine* 1972; 52: 81-85.
- Beck AT, Brown G, Steer RA. Prediction of eventual suicide in psychiatric inpatients by clinical ratings of hopelessness. *J Consult Clin Psychol* 1989; 57: 309-310.
- Beck AT, Rush AJ, Shaw BF, Emery E. *Cognitive therapy of depression*. New York: Guilford; 1983. [Trad. esp. en Bilbao: Descle de Brower. Original de 1979].
- Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory. Twenty-five years of evaluation. *Clin Psychol Rev* 1988; 8: 77-100.
- Beck AT, Ward CH, Mendelson M, Mock JE, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961; 4: 561-585.
- Beck AT, Weissman A, Lester D, Trexler L. The measurement of pessimism: The hopelessness scale. *J Consult Clin Psychol* 1974; 42: 861-865.
- Beckham EE, Leber WR (eds.). *Handbook of depression. Treatment, assessment, and research*. Homewood: Dorsey, 1985.
- Beckham EE, Leber WR, Watkins J, Boyer JL, Cook JB. Development of an instrument to measure Beck's Cognitive Triad: The Cognitive Triad Inventory. *J Consult Clin Psychol* 1986; 54: 566-567.
- Beigel A, Murphy D, Bunney WE. The Manic State Rating Scale: scale construction, reliability, and validity. *Arch Gen Psychiatry* 1971; 25: 256-262.
- Berndt DJ. Cuestionarios y escalas. En: Wolman BB, Striker G (eds.). *Trastornos depresivos: Hechos, teorías y métodos de tratamiento*. Barcelona: Ancora, 1993; 265-285.
- Biggs JT, Wylie LT, Ziegler VE. Validity of the Zung Self Rating Depression Scale. *Br J Psychiatry* 1981; 32: 381-385.
- Blackburn IM. Psychological processes in depression. En: Miller E, Cooper PJ (eds.). *Adult abnormal psychology*. Edimburgo: Churchill Livingstone, 1988.
- Blackburn IM, Jones S, Lewin RJ. Cognitive style in depression. *Br J Clin Psychol* 1986; 25: 241-251.
- Blackburn IM, Loudon JB, Ashworth CM. A new scale for measuring mania. *Psychol Med* 1977; 7: 453-458.
- Bradock I. The dexamethasone suppression test: fact and artefact. *Br J Psychiatry* 1986; 148: 363-374.
- Bramley PN, Easton A, Morley S, Snaith RP. The differentiation of anxiety and depression by rating scales. *Acta Psychiatr Scand* 1987; 77: 133-138.
- Bravo M, Canino G, Biro H. El DIS en español: Su traducción y adaptación en Puerto Rico. *Acta Psiquiátrica Psicológica América Latina* 1987; 33: 27-42.
- Brébion C, Smith MJ, Allilaire JF. Psychometric characteristics of ideational retardation in depressives. *Br J Clin Psychol* 1995; 34: 371-382.
- Brewin CR, Andrews B, Gotlib IH. Psychopathology and early experience: A reappraisal of retrospective reports. *Psychol Bull* 1993; 113: 82-98.
- Brink TL, Rose TL, Huang V, Adeley MB, Leirer VO. Development and validation of a geriatric depression screening scale: A preliminary report. *J Psychiatr Res* 1983; 17: 37-49.
- Brink TL, Yesavage JA, Lum O, y cols. Screening tests for geriatric depression. *Clin Gerontol* 1982; 1: 37-43.
- Brown GW. Depression: A radical social perspective. En: Herbst K, Paykel E (eds.). *Depression: An integrative approach*. London: Heineman Medical Books, 1989.
- Burke JD, Regier DA. Epidemiología de los trastornos mentales. En: Talbott JA, Hales RE, Yudofsky SC (eds.). *Tratado de Psiquiatría*. Barcelona: Ancora, 1989. [Original: 1988.]
- Burns DD. *Sentirse bien*. Barcelona: Paidós, 1990. [Original de 1980.]
- Byerly FC, Carlson WA. Comparison among inpatients, outpatients, and normals on three self-report depression inventories. *J Clin Psychol* 1982; 38: 797-804.
- Cantor A. The efficacy of a short form of the MMPI to evaluate depression and morale loss. *J Consult Psychol* 1960; 24: 14.
- Caro J, Ibáñez E. La escala hospitalaria de ansiedad y depresión. *Bol Psicol* 1992; 36: 43-69.
- Carroll BJ, Feinberg M, Smouse PE, Rawson SG, Greden JE. The Carroll Rating Scale for Depression: I. Development, reliability and validation. *Br J Psychiatry* 1981; 138: 194-200.
- Clark DA. Cognitive-affective interaction: A test of the «specificity» and «generality» hypotheses. *Cognit Ther Res* 1986; 10: 607-623.
- Clark DA. The validity of measures of cognition: a review of the literature. *Cognit Ther Res* 1988; 12: 1-20.
- Clark DA. Depressive, anxious and intrusive thoughts in psychiatric inpatients and outpatients. *Behav Res Ther* 1992; 30: 93-102.
- Comeche MI, Diaz MI, Vallejo MA. *Cuestionarios, inventarios y escalas: ansiedad, depresión y habilidades sociales*. Madrid: Fundación Universidad Empresa, 1995.
- Compas B, Ey S, Grant K. Taxonomy, assessment, and diagnosis of depression during adolescence. *Psychol Bull* 1993; 114: 323-344.
- Conde V, Esteban T, Useros E. Revisión crítica de la adaptación castellana del cuestionario de Beck. *Rev Psicol Gen Aplic* 1976; 31: 469-497.
- Conde V, Esteban T, Useros E. Evaluación crítica de la SDS (Self-rating Depression Scale) de Zung: Experiencia propia y revisión de la literatura. *Rev Psicol Gen Aplic* 1976; 43: 842-857.
- Conde V, Franch JI. Escalas de evaluación comportamental para la clasificación de la sintomatología psicopatológica en los trastornos angustiosos y depresivos. Madrid: Upjohn, 1984.
- Coyne JC, Gotlib I. The role of cognition in depression: A critical appraisal. *Psychol Bull* 1983; 94: 472-505.
- Crandell DL, Cathy J, Chambless Dianne L. The validation of an inventory for measuring depressive thoughts: The Crandell Cognitions Inventory. *Behav Res Ther* 1986; 24: 403-411.
- Crawford-Little J, McPhail NI. Measures of depressive mood at monthly intervals. *Br J Psychiatry* 1973; 122: 447.
- Chant SN, Myers CR. An approach to the measurement of mental health. *Am J Orthopsychiatry* 1936; 6: 134-140.
- Dahlstrom WG, Brooks JD, Peterson CD. The Beck Depression Inventory: Item order and the impact of response sets. *J Personal Assess* 1990; 55: 224-233.
- Del Barrio V. Evaluación de las características psicopatológicas. En: Fernández-Ballesteros R (ed.). *Introducción a la evaluación psicológica I*. Madrid: Pirámide, 1992.
- Dempsey PL. A unidimensional scale for the MMPL. *J Consult Psychol* 1964; 28: 364-370.
- Dobson KS. A meta-analysis of the efficacy of cognitive therapy for depression. *J Consult Clin Psychol* 1989; 57: 414-420.
- Dobson KS, Shaw BF. Cognitive assessment with major depressive disorders. *Cognit Ther Res* 1986; 10: 13-30.
- Dohr KB, Rush AJ, Bernstein IH. Cognitive bases and depression. *J Abn Psychol* 1989; 98: 263-267.
- Duke University. Center for The Study of Aging And Human Development Multidimensional functional assessment: The

- OARS methodology (2^o ed.). Durham: Duke University, 1978.
- Edwards BC, Lambert MJ, Moran PW, McCully T, Smith KC, Elligson AG. A meta-analytic comparison of the Beck Depression Inventory and the Hamilton Rating Scale for Depression as measures of treatment outcome. *Br J Clin Psychol* 1984; 23: 93-99.
- Endicott J. A diagnostic interview: the Schedule for Affective Disorders and Schizophrenia. *Arch Gen Psychiatry* 1978; 135: 837-844.
- Endicott J, Cohen J, Nee J, Fleiss J, Sarantakos S. Hamilton Depression Rating Scale, extracted from regular and change versions of the Schedule for Affective Disorders and Schizophrenia. *Arch Gen Psychiatry* 1981; 35: 837-844.
- Endicott J, Cohen J, Nee J, Fleiss J, Sarantakos S. Hamilton Depression Rating Scale, extracted from regular and change versions of the Schedule for Affective Disorders and Schizophrenia. *Arch Gen Psychiatry* 1981; 138: 98-103.
- Endicott J, Spitzer R. Use of the Research Diagnostic Criteria and the Schedule for Affective Disorders and Schizophrenia to study affective disorders. *Arch Gen Psychiatry* 1979; 136: 52-56.
- Endler NS, Cox BJ, Parker J, Bagby RM. Self-reports of depression and state-trait anxiety: Evidence for differential assessment. *J Personality Soc Psychol* 1989; 63: 832-838.
- Estudio Sociológico. Libro Blanco de la Depresión en España. Madrid: Gabinete de Estudios Sociológicos Bernard Krief y Laboratorios Lederle, 1982.
- Ezpeleta L. Las entrevistas estructuradas en el diagnóstico psicopatológico infantil. En: Rodríguez Sacristán J (ed.). Psicopatología del niño y del adolescente (Vol. 1). Sevilla: Universidad de Sevilla, 1995; 305-330.
- Faulstich ME, Carey MP, Ruggiero L, Enyart P, Gresham F. Assessment of depression in childhood and adolescence: An evaluation of the Center for Epidemiological Studies depression scale for children (CES-DC). *Am J Psychiatry* 1986; 143: 1024-1027.
- Feinberg M, Carroll BJ, Spouse PE, Rawson SG. The Carroll Rating Scale for Depression: III. Comparison with other ratings instruments. *Br J Psychiatry* 1981; 138: 205-209.
- Fennell MJ, Campbell EA. The Cognitions Questionnaire: specific thinking errors in depression. *Br J Clin Psychol* 1984; 23: 81-92.
- Fernández-Ballesteros R (eds.). Introducción a la evaluación psicológica (2 vol.). Madrid: Pirámide, 1992a.
- Fernández-Ballesteros R. Técnicas subjetivas. En: Fernández-Ballesteros R (ed.). Introducción a la evaluación psicológica I. Madrid: Pirámide, 1992b.
- Fischer J, Corcoran K. Measures for clinical practice: a sourcebook. New York: Free Press, 1987.
- Folstein MF, Juria R. The validity and reliability of the Visual Analogue Mood Scale. *J Psychosom Res* 1973; 12: 51-57.
- Fombonne E. Depressive disorders: Time trends and possible explanatory mechanisms. En: Rutter M, Smith DJ (eds.). Psychosocial disorders in young people. Chichester: Wiley, 1995; 544-615.
- Freeman CP, Munro J. Process and outcome in psychotherapy. En: Peck DE, CM (ed.). Measuring human problems. New York: Wiley, 1990.
- Callagher-Thompson D, Futterman A, Hanley-Peterson P y cols. Endogenous depression in the elderly: Prevalence and agreement among measures. *J Consult Clin Psychol* 1992; 60: 300-303.
- García Merita M, Miquel M, Fuentes I. Psicofisiología clínica: Depresión y esquizofrenia. Valencia: Promolibro, 1992.
- Gastó C, Vallejo J. Biología de los trastornos afectivos. En: Vallejo J, Gastó C (eds.). Trastornos afectivos: Ansiedad y depresión. Barcelona: Salvat, 1990.
- Gold JR. Niveles de depresión. En: Wolman BB, Striker G (eds.). Trastornos depresivos: Hechos, teorías y métodos de tratamiento. Barcelona: Ancora, 1993; 210-236.
- González Felipe MA, Szurek Soler S. La Geriatric Depression Scale (GDS) como instrumento para la evaluación de la depresión en el anciano: Bases de la misma y modificaciones introducidas para la adaptación de la prueba a la realidad psicogeriatrica española. En: Diagnóstico y evaluación psicológica. Libro de comunicaciones del II Congreso Oficial de Psicólogos. Área 7. Valencia: Colegio Oficial de Psicólogos, 1990; 330-370.
- Gotlib IH. Depression and general psychopathology in university students. *J Abn Psychol* 1984; 93: 19-30.
- Gotlib IH, Cane DB. Self-report assessment of depression and anxiety. En: Kendall PC, Watson D (eds.). Anxiety and depression: Distinctive and overlapping features. San Diego: Academic, 1989; 131-169.
- Gotlib IH, Meyer JP. Factor analysis of the Multiple Affective Adjective Check List: A separation of positive and negative affect. *J Personality Soc Psychol* 1986; 50: 1161-1165.
- Gould J. A psychometric investigation of the standard and short form Beck Depression Inventory. *Psychol Rep* 1982; 51: 1167-1170.
- Granier F, Girard M, Cross D, Fernet J, Escande M. L'Echelle de Dépression de Widlöcher: analyse multifactorielle et comparaison avec l'Echelle de Hamilton chez 55 patients déprimés hospitalisés. *Ann Med Psychol* 1982; 140: 1134-1137.
- Greden JF, Carroll BS. Psychomotor function in affective disorders: an overview of new monitoring techniques. *Am J Psychiatry* 1981; 138: 1441-1448.
- Guerrero Hamilton EW, Abramson LY. Cognitive patterns and major depressive disorder: A longitudinal study in a hospital setting. *J Abn Psychol* 1983; 92: 173-184.
- Guilford JP, Guilford RB. Personality factors D, R, T and A. *J Abn Soc Psychol* 1939; 34: 23-36.
- Gurland BJ, Golen RR, Teresi JA, Challop J. The SHORT-CARE: an efficient instrument for the assessment of depression, dementia, and disability. *J Gerontol* 1984; 39: 166-169.
- Gurland BJ, Kuriansky J, Sharpe L, Simon R, Stiller P, Birkett P. The comprehensive assessment and referral evaluation (CARE) rationale, development and reliability: Part II. A factor analysis. *Int J Ag Hum Develop* 1977; 8: 9-42.
- Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatr* 1960; 23: 56-62.
- Hamilton M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol* 1967; 6: 278-296.
- Hamilton M. Comparative value of rating scales. *Br J Clin Pharmacol* 1976; 3 (Supl. 1): 58-60.
- Hamilton M. Rating depressive patients. *J Clin Psychiatry* 1980; 41: 21-24.
- Hamilton M. Assessment of Psychopathology. En: Hindmarch I, Stonier PD (eds.). Human psychopharmacology: Measures and methods. Vol. 1. New York: Wiley, 1987.
- Hamilton M, Shapiro CM. Depression. En: Peck DE, Shapiro CM (eds.). Measuring human problems. New York: Wiley, 1990.
- Hammen C, Krantz SE. Measures of psychological processes in depression. En: Beckham EE, Leber WR (eds.). Handbook

- of depression. Treatment, assessment and research. Home-wood: Dorsey, 1985.
- Hartlage S, Alloy L, Vázquez C, Dyckman L. Automatic and effortful processing in depression. *Psychol Bull* 1993; 2: 247-278.
- Hathaway SR, McKinley PE. Minnesota Multiphasic Personality Inventory. Psychological Corp., 1942.
- Hedlund JL, Vieweg BW. The Hamilton rating scale for depression: A comprehensive review. *J Operat Psychiatry* 1979a; 10: 149-165.
- Hedlund JL, Vieweg BW. The Zung Self Rating Depression Scale: A comprehensive review. *J Operat Psychiatry* 1979b; 10: 51-64.
- Helmes E, Reddon JR. A perspective on development in assessing psychopathology: A critical review of the MMPI and MMPI-2. *Psychol Bull* 1993; 113: 453-472.
- Helzer JE, Stotzman RK, Farmer A, Brockington IF, Works J. Comparison with a DIS/DSM-III based physician reevaluation in St. Louis. En: Eaton W, Kessler LG (eds.). *Epidemiological field methods in psychiatry: The NIMH Epidemiological Catchment Area Program*. New York: Academic Press, 1985.
- Hertzog C, Van Alstine J, Usala P, Hultsch D, Dixon R. Measurement properties of the Center for Epidemiological Studies Depression Scale (CES-D) in older populations. *Psychol Assess J Consult Clin Psychol* 1990; 2: 64-72.
- Hollon SD, Kendall PC. Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cogn Ther Res* 1980; 4: 383-395.
- Hollon SD, Kendall PC, Lumry A. Specificity of depressotypic cognitions in clinical depression. *J Abn Psychol* 1986; 95: 52-59.
- Ingram R, Wisnicki KS. Assessment of Positive Automatic Cognition. *J Consult Clin Psychol* 1988; 56: 898-902.
- Jacobson NS. The assessment to overt behavior in depression. En: Rehm LP (ed.). *Behavior therapy of depression: Present status and future directions*. New York: Academic, 1981.
- Jasper HH. The measurement of depression-elation and its relation to a measure of extroversion-introversion. *J Abn Soc Psychol* 1930; 25: 307-318.
- Jones RA. A factored measure of Elli's irrational belief system with personality and maladjustment correlates. *Dissert Abst Int* 1969; 29: 4379-4380.
- Jones RA. A factored measure of Elli's irrational belief system with personality and maladjustment correlates. *Dissert Abst Int* 1969; 29: 11-13.
- Kearns NP, Cruickshank CA, McGuigan SA, Siley SA, Shaw SP, Snaith RP. A comparison of depression rating scales. *Br J Psychiatry* 1982; 141: 45-49.
- Kendall PC, Hollon SD. Assessing self-referent speech: Methods in the measurement of self-statements. En: Kendall PC, Hollon SD (eds.). *Assessment strategies for cognitive-behavioral interventions*. New York: Academic Press, 1981.
- Kendall PC, Hollon SD, Beck AT, Ingram RE. Issues and recommendations regarding use of the Beck Depression Inventory. *Cogn Ther Res* 1987; 11: 289-299.
- Kendall PC, Howard BL, Hays RC. Self-referent speech and psychopathology: The balance positive and negative thinking. *Cogn Ther Res* 1989; 13: 583-598.
- Kobak KA, Reynolds WM, Rosenfeld R, Greist J. Development and validation of a computer-administered version of the Hamilton Depression Rating Scale. *Psychol Assess J Consult Clin Psychol* 1990; 2: 56-63.
- Krantz S, Hammen C. Assessment of cognitive bias in depression. *J Abn Soc Psychol* 1979; 88: 611-619.
- Lambert MJ, Hatch DR, Kingston MD, Edwards BC. Zung, Beck y Hamilton Rating Scales as measures of treatment outcome: A meta-analysis comparison. *J Consult Clin Psychol* 1986; 54: 54-59.
- Leber WR, Beckham EE, Danker-Brown P. Diagnostic criteria for depression. En: Beckham EE, Leber WR (eds.). *Handbook of depression. Treatment, assessment, and research*. Home-wood: Dorsey, 1985.
- Leclercubier Y. Présentation d'un instrument clinique. En: Widlöcher D (ed.). *La ralentissement dépressif*. Paris: PUF, 1983.
- Levine S, Deo R, Mahadevan K. A comparative trial of a new antidepressant, fluoxetine. *Br J Psychiatry* 1987; 150: 653-655.
- Lewinsohn PM, Amenson CS. Some relations between pleasant and unpleasant mood-related event depression. *J Abn Psychol* 1978; 87: 644-654.
- Lewinsohn PM, Talkington J. Studies in the measurement of unpleasant events and relations with depression. *Appl Psychol Meas* 1979; 3: 83-101.
- Lipton AA, Simon FS. Psychiatric diagnosis in a state hospital: Manhattan State revisited. *Hosp Comm Psychiatry* 1985; 36: 368-373.
- Loudon JB, Tiplady B. Assessment of depression and the non-psychiatrist. *Adv Biosci* 1981; 31: 101-110.
- Louks J, Hayne C, Smith J. Replicated factor structure of the Beck Depression Inventory. *J Nerv Ment Dis* 1989; 177: 473-479.
- Luteijn F, Bourman TK. The concepts of depression, anxiety and neuroticism in questionnaires. *Eur J Pers* 1988; 2: 113-120.
- Lloyd G. Adjustment to illness. En: Peck DF, Shapiro CM (eds.). *Measuring human problems*. New York: Wiley, 1990.
- MacPhillamy DJ, Lewinsohn PM. The Pleasant Events Schedule. Universidad de Oregon; 1971. Manuscrito no publicado.
- MacPhillamy DJ, Lewinsohn PM. The Pleasant Events Schedule: Studies on reliability, validity, and scale intercorrelation. *J Consult Clin Psychol* 1982; 50: 363-380.
- Maganto C. Instrumentos aplicados para la evaluación y el diagnóstico en psicopatología infantil. En: Rodríguez Sacristán J (ed.). *Psicopatología del niño y del adolescente*. Vol. 1. Sevilla: Universidad de Sevilla, 1995; 191-282.
- Maier W, Philipp M. Comparative analysis of observer depression scales. *Acta Psychiatr Scand* 1985; 72: 239-245.
- Maier W, Philipp M, Hauser I, Schlegler S, Raimund A, Wetzel H. Improving depression severity assessment: Reliability, internal validity and sensitivity to change of three observer depression scales. *J Psychiatry Res* 1988; 22: 3-12.
- Marsella AJ, Hirschfeld RMA, Katz MM (eds.). *The measurement of depression*. Chichester: Wiley, 1987.
- Matt J, Vázquez C, Campbell K. Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clin Psychol Rev* 1992; 2: 227-256.
- May R. *The meaning of anxiety*. New York: Ronald Press, 1950.
- McCormack HH, Horne D, Sheather S. Clinical application of visual analogue scales: A critical review. *Psychol Med* 1988; 18: 1007-1019.
- Mezzich JE, Damarin FL, Erickson JR. Comparative validity of strategies and indices for differential diagnosis of depressive states from other psychiatric conditions using the MMPI. *J Consult Clin Psychol* 1974; 42: 691-698.
- Miller IW, Bishop S, Norman W, Maddever H. The modified Hamilton Rating Scale for Depression: reliability and validity. *Psychiatr Res* 1985; 14: 131-142.

- Montgomery SA, Asberg M. A new depression rating scale designed to be sensitive to change. *Br J Psychiatry* 1979; 134: 382-389.
- Muñoz M, Vázquez C, Cruzado JA. Personas sin hogar en Madrid: Informe social y epidemiológico. Madrid: Imprenta de la Comunidad Autónoma de Madrid, 1995.
- Murphy DL, Beigel A. Depression, elation and lithium carbonate responses in manic patient subgroups. *Arch Gen Psychiatry* 1974; 31: 643-648.
- Nasr S, Altman EG, Rodin M, Jobe TH, Burg B. Correlation of the Hamilton and Carroll depression rating scales: A replication study among psychiatric outpatients. *J Clin Psychiatry* 1984; 45: 167-168.
- Norman WH, Miller IW, Klee SH. Assessment of cognitive distortion in a clinically depressed population. *Cogn Ther Res* 1983; 7: 133-140.
- Paykel ES, Norton KR. Self-report and clinical interview in the assessment of depression. En: Sartorius N, Ban TA (eds.). *Assessment of depression*. Berlin: Springer-Verlag y World Health Organization, 1986; 356-366.
- Peck DE, Shapiro CM (eds.). *Measuring human problems*. New York: Wiley, 1990.
- Pedreira JL, Sánchez B, Sardinero E, Martín L, Martín P. Aplicación del método epidemiológico para la detección de trastornos mentales en la infancia y adolescencia. En: González de Rivera JL, Rodríguez F, Sierra A (eds.). *El método epidemiológico en salud mental*. Barcelona: Masson-Salvat, 1993; 215-238.
- Peterson C, Seligman MEP. Causal explanations as a risk factor for depression: Theory and evidence. *Psychol Rev* 1984; 91: 347-374.
- Polaino A (ed.). *Las depresiones infantiles*. Madrid: Morata, 1988.
- Post RD y cols. The Mezzich MMPI regression formula revisited. *J Personal Assess* 1985; 49: 258-259.
- Radloff LS. The CES-D scale: A self-report scale for research in the general population. *Appl Psychol Meas* 1977; 1: 385-401.
- Rafaelsen OJ, Bech P, Bolwig T, Kramp P, Gjerris A. The Bech-Rafaelsen combined rating scale for mania and melancholia. *Psychiatria Fennica* 1980; (Supl.) 327-331.
- Rehm LP. The measurement of behavioral aspects of depression. En: Marsella AJ, Hirschfeld RM, Katz MM (eds.). *The measurement of depression*. New York: Guilford, 1987.
- Rehm LP, O'Hara MW. Item characteristics of the Hamilton Rating Scale for Depression. *J Psychiatr Res* 1985; 19: 31-42.
- Roa A (ed.). *Evaluación en Psicología Clínica y de la Salud*. Madrid: CEPE, 1995.
- Robertson MM. The assessment of mood and its components. En: Hindmarch I, Stonier PD (eds.). *Human psychopharmacology: Measures and methods*. (Vol. 2). New York: Wiley, 1989.
- Robins LN, Helzer JE. Diagnosis and clinical assessment: The current state of psychiatric diagnosis. *Ann Rev Psychol* 1986; 37: 409-432.
- Rodríguez-Pulido F, Montes de Oca D. Instrumentos para la detección de casos en psiquiatría: Cuestionarios y entrevistas. En: González de Rivera JL, Rodríguez F, Sierra A (eds.). *El método epidemiológico en salud mental*. Barcelona: Masson-Salvat, 1993; 125-144.
- Rosch E. Principles of categorization. En: Rosch E, Lloyd B (eds.). *Cognition and categorization*. Hillsdale: Erlbaum, 1978; 27-48.
- Rubio-Stipec M, Bravo M, Canino G. La Entrevista Diagnóstica Internacional Compuesta (CID): Un instrumento epidemiológico adecuado para ser administrado conjuntamente con otros sistemas diagnósticos en diferentes culturas. *Acta Psiquiatr Psicol Am Lat* 1991; 37: 191-204.
- Ruiz Caballero JA, Bermúdez J. Consideraciones en torno al Beck Depression Inventory como instrumento de identificación de sujetos deprimidos en muestras subclínicas. *Eval Psicol* 1989; 5: 255-272.
- Sanz J, Vázquez C. Trastornos depresivos (II): Productos, operaciones, proposiciones y estructuras cognitivas. En: Belloch A, Ibáñez E (eds.). *Manual de Psicopatología*. Vol. 2. Valencia: Promolibro, 1991a; 785-875.
- Sanz J, Vázquez C. Adaptación española de la escala de actitudes disfuncionales (DAS) de Weissman y Beck: Propiedades clínicas y psicométricas. *An Mod Cond* 1993; 19: 707-750.
- Sanz J, Vázquez C. Algunas consideraciones adicionales sobre la escala de actitudes disfuncionales (DAS) de Weissman y Beck. *An Mod Cond* 1994; 20: 669-673.
- Sanz J, Vázquez C. Trastornos del estado de ánimo: Teorías psicológicas. En: Belloch A, Sandín B, Ramos F (eds.). *Manual de Psicopatología*. Vol. 2. Madrid: McGraw Hill, 1995; 341-378.
- Sartorius N, Ban TA (eds.). *Assessment of depression*. Berlin: Springer-Verlag y World Health Organization, 1986.
- Schaefer A, Brown J, Watson C, Plemel D y cols. Comparison of the validities of the Beck, Zung and MMPI depression scales. *J Consult Clin Psychol* 1985; 53: 415-418.
- Schwab JJ, Bialow MR, Holzer CE. A comparison of two rating scales for depression. *J Clin Psychol* 1967; 23: 94-99.
- Seeman M. Alienation motifs and contemporary theorizing: The hidden continuity of classic themes. *Soc Psychol Quart* 1983; 46: 171-174.
- Seligman MEP, Abramson LY, Semmel A, Von Baeyer C. Depressive attributional style. *J Abn Psychol* 1979; 88: 242-247.
- Senra C, Polaino A. Concordance between clinical and self-report depression scales during the acute phase and after treatment. *J Affect Dis* 1993; 27: 13-20.
- Shaver PR, Brennan KA. Measures of depression and loneliness. En: Robinson JP, Shaver PR, Wrightsman LS (eds.). *Measures of personality and social psychology attitudes*. New York: Academic Press, 1991; 195-289.
- Shaw BF, Vallis TM, McCabe SB. The assessment of the severity and symptom patterns in depression. En: Beckham EE, Leber WR (eds.). *Handbook of depression*. Treatment, assessment and research. Homewood: Dorsey, 1985.
- Sholomskas DE. Métodos de entrevista. En: Wolman BB, Striker G (eds.). *Trastornos depresivos: Hechos, teorías y métodos de tratamiento*. Barcelona: Ancora, 1993; 239-256.
- Silva F. Evaluación conductual y criterios psicométricos. Madrid: Pirámide, 1989.
- Silverman JS, Silverman JA, Eardley DA. Do maladaptive attitudes cause depression? *Arch Gen Psychiatry* 1984; 41: 28-30.
- Smith MD, Hong BA, Robson AM. Diagnosis of depression in patients with end-stage renal disease. *Am J Med* 1985; 79: 160-166.
- Snaith RP. The concepts of mild depression. *Br J Psychiatry* 1987; 150: 387-393.
- Snaith RP, Ahmed SN, Metha S, Hamilton M. Assessment of the severity of primary depressive illness: Wakefield Self-Assessment Depression Inventory. *Psychol Med* 1971; 1: 143-149.
- Sositsky SM, Glass DR, Shea M, Pilkonis P y cols. Patient predictors of response to psychotherapy and pharmacotherapy: Findings in the NIMH Treatment of Depression Collaborative Research Program. *Am J Psychiatry* 1991; 148: 997-1008.

Parte III. Evaluación de síndromes y enfermedades

- Spitzer RL, Endicott J, Robins E. Research Diagnostic Criteria for a selected group of functional disorders. New York: Biometrics Research, New York State Psychiatric Institute, 1975.
- Spitzer RL, Skodol AE, Williams JB. Supervision intake diagnosis: A psychiatric «Rashomon». *Arch Gen Psychiatry* 1982; 39: 1299-1305.
- Spitzer RL, Williams JB, Gibbon M. Structured Clinical Interview for DSM-III-R. New York: Biometrics Research Department, New York Psychiatric Institute, 1987.
- Steer RA, Beck AT, Garrison B. Applications of the Beck Depression Inventory. En: Sartorius N, Ban TA (eds.). *Assessment of depression*. Berlin: Springer-Verlag y World Health Organization, 1986; 123-142.
- Sweeney PD, Anderson K, Bailey S. Attributional style in depression: A meta-analytic review. *Journal of Personality and Social Psychology* 1986; 50: 974-991.
- Tanaka-Matsumi J, Kameoka VA. Reliabilities and concurrent validities of popular self-report measures of depression, anxiety, and social desirability. *J Consult Clin Psychol* 1986; 54: 328-333.
- Teasdale JD, Rezin V. Effect of thought-stopping on thoughts, mood and corrugator EMG in depressed patients. *Behav Res Ther* 1978; 16: 97-102.
- Tejero A, Guimerá E, Farré JM, Peri JM. Uso clínico del HAD (Hospital Anxiety and Depression Scale) en población psiquiátrica: Un estudio de su sensibilidad, fiabilidad y validez. *Revista del Depto. de Psiquiatría de la Facultad de Medicina de Barcelona* 1986; 13: 233-238.
- Thompson C. Introduction. En: Thompson C (ed.). *The instruments of psychiatric research*. New York: Wiley, 1989a.
- Thompson C. Affective disorders. En: Thompson C (ed.). *The instruments of psychiatric research*. New York: Wiley, 1989b.
- Thompson C. Anxiety. En: Thompson C (ed.). *The instruments of psychiatric research*. New York: Wiley, 1989c.
- Vallejo J, Gastó C. Trastornos afectivos: ansiedad y depresión. Barcelona: Salvat, 1990.
- Vázquez C. Escalas evaluadoras de la depresión: limitaciones conceptuales y metodológicas. *Rev Psicol Gen Aplic* 1986; 41: 101-113.
- Vázquez C. Fundamentos teóricos y metodológicos de la clasificación en Psicopatología. En: Fuentenebro F, Vázquez C (eds.). *Psicología médica, Psicopatología, y Psiquiatría*. Vol. 2. Madrid: McGraw Hill, 1990a; 655-678.
- Vázquez C. Sistemas de clasificación en Psicopatología. En: Fuentenebro F, Vázquez C (eds.). *Psicología médica, Psicopatología y Psiquiatría*. Vol. 2. Madrid: McGraw Hill, 1990b; 679-708.
- Vázquez C. Trastornos del sueño. En: Fuentenebro F, Vázquez C (eds.). *Psicología médica, Psicopatología y Psiquiatría*. Vol. 2. Madrid: McGraw Hill, 1990c; 655-678.
- Vázquez C. Evaluación de trastornos depresivos y bipolares. En: Roa A (ed.). *Evaluación en Psicología Clínica y de la Salud*. Madrid: CEPE, 1995; 163-209.
- Vázquez C, Avia MD, Jiménez F, Saura F. Adaptación española del Cuestionario de Estilo Atribucional (ASQ). Facultad de Psicología: Universidad Complutense. En preparación.
- Vázquez C, Crespo M, Ring J. Estrategias de afrontamiento. En: Guimón J, Berrios B, Bulbena A (eds.). *Medición en Psiquiatría*. Barcelona: Salvat, 1996.
- Vázquez C, Lozoya G. Evaluación y diagnóstico diferencial de los trastornos depresivos en personas de edad. En: Buendía J (ed.). *Envejecimiento y psicología de la salud*. Madrid: Siglo XXI, 1994; 247-278.
- Vázquez C, Ring J. Altered cognitions in depression: Are dysfunctional attitudes stable? *Personality and Individual Differences* 1993; 15: 475-479.
- Vázquez C, Ring J. El CID (Coping Inventory of Depression): Un instrumento para evaluar estrategias de afrontamiento ante la depresión. *Psicología Conductual*. En prensa.
- Vázquez C, Sanz J. Adaptación española del Inventario de Depresión de Beck (1978): Propiedades psicométricas y clínicas en muestras clínicas y en población general. Madrid: Universidad Complutense, Facultad de Psicología. En preparación.
- Vázquez C, Sanz J. Trastornos depresivos (I): Datos clínicos y modelos teóricos. En: Belloch A, Ibáñez Y (eds.). *Manual de Psicopatología*. Vol. 2. Valencia: Promolibro, 1991a; 717-784.
- Vázquez C, Sanz J. Fiabilidad y validez factorial de la versión española del Inventario de Depresión de Beck. Comunicación presentada en el III Congreso de Evaluación Psicológica. Barcelona, 25-28 de septiembre, 1991b.
- Vázquez C, Sanz J. Trastornos depresivos: Historia y cuadros clínicos. En: Belloch A, Sandín B, Ramos F (eds.). *Manual de Psicopatología*. Vol. 2. Madrid: McGraw Hill, 1995; 299-340.
- Vázquez-Barquero JL. Los estudios comunitarios de salud mental: su diseño y ejecución. En: González de Rivera JL, Rodríguez F, Sierra A (eds.). *El método epidemiológico en salud mental*. Barcelona: Masson-Salvat, 1993; 19-35.
- Vizcarro C, García Marcos JA. Los auto-informes. En: Fernández Ballesteros R, Carrobbles JA (eds.). *Evaluación conductual. Metodología y aplicaciones*. Madrid: Pirámide, 1981.
- Vizcarro C, León O, Garciamartín J, García J, Llorente G, Romero J, Vargas JA. Aportaciones de una entrevista estructurada (SADS) al acuerdo interjueces en el diagnóstico por medio los Criterios Diagnósticos para la Investigación. *Evaluación Psicológica* 1986; 2: 99-114.
- Von Zerssen D. Personalidad y desórdenes afectivos. En: Paykel ES (ed.). *Psicopatología de los trastornos afectivos*. Madrid: Pirámide, 1985.
- Von Zerssen D, Strian F, Schwarz D. Evaluation of depressive states, especially in longitudinal studies. En: Pichot P, Olivier R (eds.). *Psychological measurement in psychopharmacology*. London: Karger, 1974.
- Watkins JT, Rush AJ. Cognitive Response Test. *Cognitive Therapy and Research* 1983; 7: 425-435.
- Watson D, Clark LA, Tellegen A. Development and validation of brief measures of Positive and Negative Affect: The PANAS Scales. *J Personal Soc Psychol* 1988; 54: 219-235.
- Watson D, Kendall PC. Understanding anxiety and depression: Their relation to negative and positive affective states. En: Kendall PC, Watson D (eds.). *Anxiety and depression: Distinctive and overlapping features*. San Diego: Academic Press, 1989; 3-26.
- Weissman AN, Beck AT. Development and validation of the Dysfunctional Attitude Scale: A preliminary investigation. Comunicación presentada en la reunión anual de la Association for the Advancement of the Behavior Therapy. Chicago, 1978.
- Weissman MM, Boyd JH. Affective disorders: Epidemiology. En: Kaplan HI, Sadock BJ (eds.). *Comprehensive Textbook of Psychiatry*. Vol. 1. Baltimore: Williams and Wilkins, 1985.
- Weissman MM, Prusoff B, Newberry P. Comparison of CES-D, Zung, Beck self-report depression scales. (Technical report ADM 42-47-83). Rockville, MD: National Institute of Mental Health, 1975.
- Wetzler S, Van Praag M. Assessment of depression. En: Wetzler S (ed.). *Measuring mental illness: Psychometric assessment for clinicians*. Washington: American Psychiatric Press, 1989.

- Whisman MA, Strosahl K, Fruzzetti AE, Schmaling KB, Jacobson NS, Miller DM. A structured interview version of the Hamilton Rating Scale for Depression: Reliability and validity. *Psychol Assess J Consult Clin Psychol* 1989; 3: 238-241.
- Widlöcher D. Psychomotor retardation: Clinical theoretical and psychometric aspects. En: Akiskal HS (ed.). *The Psychiatric Clinics of North America*. Vol. 6. Philadelphia: W.B. Saunders, 1983a; 27-40.
- Widlöcher D. Le ralentissement dépressif. Paris: Presses Universitaires de France, 1983b.
- Widlöcher D, Gozlan A. The measurement of retardation in depression. En: Hindmarch I, Stonier PD (eds.). *Human psychopharmacology: Measures and methods*. Vol. 2. New York: Wiley, 1989; 1-22.
- Wilcoxon LA, Schrader SL, Nelson RE. Formulaciones conductuales de la depresión. En: Craighead WE, Kazdin AE, Mahoney MJ (eds.). *Modificación de Conducta: Principios, Técnicas y Aplicaciones*. Barcelona: Omega, 1984.
- Williams JB. A structured interview guide for the Hamilton Depression Rating Scale. *Arch Gen Psychiatry* 1988; 45: 742-747.
- Winters KC, Newmark CS, Lumry AE y cols. MMPI codetypes in identifying DSM-III schizophrenics, unipolars, and bipolars. *J Clin Psychol* 1985; 41: 382-386.
- Young JC, Biggs JT, Ziegler V, Meyer DA. A rating scale for mania: Reliability, validity, and sensitivity. *Br J Psychiatry* 1978; 133: 429-435.
- Zealley AK, Aitken RC. Measurement of mood. *Proc Royal Soc Med* 1969; 62: 993-996.
- Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983; 67: 361-370.
- Zubin J. Use of research instruments in psychopathological assessment: Some historical perspectives. En: Wetzler S (ed.). *Measuring mental illness: Psychometric assessment for clinicians*. Washington: American Psychiatric Press, 1989.
- Zung WWK. A self-rating depression scale. *Arch Gen Psychiatry* 1965; 12: 63-70.
- Zung WWK. Zung Self-Rating Depression Scale and Depression Status Inventory. En: Sartorius N, Ban TA (eds.). *Assessment of depression*. Berlin: Springer-Verlag y World Health Organization, 1986; 221-231.