

ADAPTACIÓN ESPAÑOLA DEL INVENTARIO PARA LA DEPRESIÓN DE BECK—II (BDI-II): 1. PROPIEDADES PSICOMÉTRICAS EN ESTUDIANTES UNIVERSITARIOS

Jesús Sanz
María Eugenia Navarro
Carmelo Vázquez

Departamento de Personalidad, Evaluación y Psicología Clínica
Universidad Complutense de Madrid

RESUMEN

Este estudio presenta los primeros pasos para obtener una adaptación española del Inventario para la Depresión de Beck—II (BDI-II; Beck, Steer y Brown, 1996), ofreciendo datos sobre sus propiedades psicométricas en una muestra de 590 estudiantes universitarios españoles. La consistencia interna del BDI-II fue elevada (coeficiente alfa de 0,89). La distribución de puntuaciones del BDI-II fue similar a la encontrada en estudios previos, con una media superior en 3 puntos a la que se suele obtener con su predecesor, el BDI-IA, lo que justifica incrementar las puntuaciones para definir las categorías de gravedad de la depresión. Los análisis factoriales indicaron que el BDI-II mide una dimensión general de depresión compuesta por dos factores altamente relacionados, uno cognitivo-afectivo y otro somático-motivacional. Los ítems del BDI-II cubren todos los criterios diagnósticos sintomáticos del DSM-IV para el trastorno depresivo mayor y el distímico, lo que apoya su validez de contenido. En términos de consistencia interna, validez factorial y validez de contenido, el BDI-II parece mejor instrumento que el BDI-IA para evaluar sintomatología depresiva en estudiantes

universitarios. En una submuestra de 165 estudiantes, y tomando como criterio la entrevista estructurada por ordenador Quick DIS-III-R, el BDI-II demostró una validez aceptable para discriminar entre personas con y sin episodio depresivo mayor. No se encontraron diferencias entre sexos ni entre grupos de edad en la puntuación total del inventario, por lo que se ofrecen baremos generales del BDI-II para estudiantes universitarios.

Palabras clave: BDI-II, DEPRESIÓN, CUESTIONARIO, PROPIEDADES PSICOMÉTRICAS, ADAPTACIÓN ESPAÑOLA.

SUMMARY

This study shows the first steps to develop a Spanish adaptation of the Beck Depression Inventory—II (BDI-II; Beck, Steer y Brown, 1996) by providing data on its psychometric properties with a sample of 590 Spanish university students. Internal consistency estimate for the BDI-II was high (coefficient alpha of 0.89). The distribution of BDI-II scores was similar to that found in previous studies, yielding a mean that was 3 points higher than that found with its predecessor, the BDI-IA. This difference supports the decision of increasing the cutoff scores to define categories of depression severity. Factor analyses suggested that the BDI-II measures a general dimension of depression composed of two highly related factors: a cognitive-affective factor and a somatic-motivational one. All the symptomatic diagnostic criteria proposed by the DSM-IV for a major depressive disorder or a dysthymic disorder are accounted for the items of the BDI-II, supporting its content validity. In terms of internal consistency, factorial validity and content validity, the BDI-II appears to be a stronger instrument than the BDI-IA to assess depressive symptomatology in university students. In a subsample of 165 students, and taking the Quick DIS-III-R (a computerized, structured diagnostic interview) as criterion, the BDI-II showed acceptable validity to discriminate between persons with and without a major depressive episode. No sex or age differences on the BDI-II scores were found for this sample and, therefore, general norms scores for university students were provided.

Key words: BDI-II, DEPRESSION, QUESTIONNAIRE, PSYCHOMETRIC PROPERTIES, SPANISH ADAPTATION.

INTRODUCCIÓN

El Inventario para la Depresión de Beck (*Beck Depression Inventory*, BDI), tanto en su versión original de 1961 (BDI-I; Beck, Ward, Mendelson, Mock y Erbaugh, 1961) como en su versión revisada de 1978 (BDI-IA; Beck, Rush, Shaw y Emery, 1979; Beck y Steer, 1993), ha sido ampliamente utilizado como medida de sintomatología depresiva en pacientes con trastornos psicológicos y en población normal, hasta constituirse en el instrumento autoaplicado para evaluar la depresión más utilizado en la clínica (Beck, Steer y Garbin, 1988; Piotrowski, 1996) y en investigación (véase la Tabla 1).

En España, disponemos desde mediados de los años 70 de una adaptación española de la versión original de 1961 (Conde, Esteban y Useros, 1976) y, más recientemente, de una adaptación de la versión revisada de 1978 (Sanz y Vázquez, 1998; Vázquez y Sanz, 1997, 1999) y, con los años, el BDI también ha llegado a ser en nuestro país uno de los tests psicológicos más utilizados en la práctica clínica y en investigación. Por ejemplo, los resultados de una reciente encuesta enviada a miembros del Colegio Oficial de Psicólogos indican que el BDI se encuentra entre los seis tests más utilizados en España (Muñiz y Fernández-Hermida, 2000), y el tercero sin contar los instrumentos que evalúan inteligencia. Es más, el BDI ocuparía esa misma tercera posición entre los tests de personalidad y de psicopatología si se considera el número de investigaciones publicadas en español que han empleado dichos tests y que aparecen recogidas en *Psicodoc2002*, una de las principales bases de datos de literatura científica psicológica en español (véase la Tabla 2).

Recientemente se ha publicado una nueva versión del BDI denominada *Beck Depression Inventory—II* (BDI-II; Beck, Steer y Brown, 1996) que presenta modificaciones sustanciales con respecto a sus predecesores. Estas modificaciones van encaminadas a conseguir que el BDI-II cubra todos los criterios diagnósticos sintomáticos de los trastornos depresivos propuestos por el DSM-IV (APA, 1994). El BDI se desarrolló en los años 60 con la intención de reflejar el consenso clínico sobre los síntomas que entonces se consideraban definitorios de la depresión. En los 40 años que han pasado desde

Tabla 1.- Principales escalas para la evaluación de la depresión y número de estudios recogidos en las bases de datos PsycINFO y Psycodoc2002 que las han utilizado

| Test | <i>PsycINFO</i> (1992-2001) ^a | <i>Psycodoc2002</i> ^b |
|---|---|----------------------------------|
| Inventario para la Depresión de Beck (BDI) | 2769 | 90 |
| Escala de Depresión de Hamilton | 1396 | 30 |
| Escala Autoaplicada de la Depresión de Zung (SDS) | 504 | 26 |
| Escala del Centro de Estudios Epidemiológicos para la Depresión (CES-D) | 379 | 8 |
| Escala de Depresión del MMPI (MMPI-D) | 70 | 12 |

Nota. ^a Para su búsqueda se utilizaron las siguientes expresiones en cualquiera de los campos de *PsycINFO*: BDI = Beck Depression Inventory; Escala de Hamilton = Hamilton Rating Scale for Depression; Escala de Zung = Self Rating Depression Scale o Zung Depression Scale; CES-D = CES-D o Center for Epidemiologic Studies and Depression; MMPI-D = MMPI-D o MMPI-2 Depression o MMPI Depression o Depression scale of the MMPI/MMPI-2.
^b Para su búsqueda se utilizaron las siguientes expresiones en los campos "Resumen" o "Descriptores" de *Psycodoc2002*: BDI = BDI o (inventario y depresión y Beck); Escala de Hamilton = Hamilton y depresión; SDS = SDS o (escala y depresión y Zung); CES-D = CES-D o (centro de estudios epidemiológicos y depresión); MMPI-D = (MMPI y depresión) o (MMPI-2 y depresión).

Tabla 2.- Tests psicopatológicos y de personalidad más utilizados en España

| Test | Práctica Profesional (Muñiz y Fernández-Hermida, 2000) ^a | Investigación (<i>Psicodoc2002</i>) ^b |
|----------------|--|---|
| 16PF ó 16PF-5 | 1ª (26,53%) | 6ª (29) |
| MMPI ó MMPI-II | 2ª (15,06%) | 4ª (76) |
| BDI | 3ª (9,14%) | 3ª (90) |
| STAI | 4ª (7,87%) | 5ª (47) |
| Rorschach | 5ª (6,46%) | 1ª (204) |
| ISRA | 6ª (3,80%) | 7ª (34) |
| EPQ ó EPQ-R | s.d. | 2ª (91) |

Nota. s.d. = sin datos.

^a Entre paréntesis el porcentaje de psicólogos que informaron usar el test en su práctica profesional de la muestra de 3.455 miembros del Colegio Oficial de Psicólogos (12,4% del total de miembros) que coatestaron a una encuesta sobre los tests y su uso.

^b Entre paréntesis el número de estudios recogidos en la base de datos *Psicodoc2002* que emplearon el test; para su búsqueda se utilizaron las siguientes expresiones en los campos "Resumen" o "Descriptores" de *Psicodoc2002*: 16PF = 16PF o 16PF-5 o (cuestionario y personalidad y Cattell); MMPI = MMPI o MMPI_II o (inventario y personalidad y Minnesota); BDI = BDI o (inventario y depresión y Beck); STAI = STAI o (inventario y ansiedad y estado-rasgo) o STAI-R o STAI-E; Rorschach = test de Rorschach; ISRA = ISRA o inventario de situaciones y respuestas de ansiedad; EPQ = EPQ o EPQ-A o EPQ-R o (cuestionario y personalidad y Eysenck).

su construcción, han aparecido algunos desarrollos importantes en la forma de entender la depresión, desarrollos que quedan patentes en las definiciones de episodio depresivo mayor y distimia ofrecidas por el DSM-III (APA, 1980) y sus posteriores ediciones: el DSM-III-R (APA, 1987) y el DSM-IV (APA, 1994). A pesar de estos desarrollos, el BDI-I y, posteriormente, el BDI-IA han seguido reflejando de manera razonable las principales características de la depresión. Sin embargo, diversos estudios han señalado su falta de adecuación a varios criterios diagnósticos sintomáticos del DSM-III y DSM-IV (Moran y Lambert, 1983; Vázquez y Sanz, 1997; Vredenburg, Krames y Flett, 1985). Así, respecto a la definición de episodio depresivo mayor más consensuada hoy en día, la ofrecida por el DSM-IV (APA, 1994), el BDI-IA no cubre el criterio sintomático de enlentecimiento o agitación psicomotores, y sólo refleja de manera parcial (recogiendo sólo los aspectos deficitarios pero no de exceso) otros dos: problemas de sueño y problemas de apetito/peso (Vázquez y Sanz, 1997). Por otro lado, el BDI-IA cubre síntomas (p. ej., hipocondría, desesperanza) que no son criterios diagnósticos DSM para el episodio depresivo mayor (Burt y IsHak, 2002). Finalmente, el BDI-IA no permite evaluar la presencia de sintomatología depresiva durante el período mínimo necesario para el diagnóstico de episodio depresivo mayor (dos semanas en el DSM-IV frente al marco temporal de una semana utilizado por el BDI-IA).

Para solventar estas discrepancias, en la construcción del BDI-II el marco temporal se extendió a dos semanas, y cuatro ítems del BDI-IA (pérdida de peso, cambio en la imagen corporal, preocupación somática y dificultades en el trabajo) fueron omitidos y reemplazados por otros tantos nuevos (agitación, sentimientos de inutilidad, dificultad de concentración y pérdida de energía). Además, se modificaron los ítems de pérdida de apetito e insomnio para que éstos pudieran evaluar tanto los decrementos como los incrementos en apetito y sueño, y asimismo se introdujeron modificaciones en algunas de las afirmaciones usadas para la valoración de los otros síntomas. Por ejemplo, la redacción del ítem sobre desinterés social del BDI-IA se cambió para que incluyera no sólo desinterés por las relaciones interpersonales, sino también por

personas y actividades, de manera que el ítem del BDI-II evaluara de forma más adecuada el criterio sintomático de pérdida de interés en todas o casi todas las actividades propuesto por el DSM-IV. En definitiva, en la elaboración del BDI-II tan sólo tres ítems del BDI-IA (sentimientos de castigo, pensamientos de suicidio y pérdida de interés en el sexo) no fueron modificados y, por tanto, el BDI-II supone una importante actualización del popular BDI-IA.

Los estudios psicométricos realizados hasta la fecha con el BDI-II indican que esta nueva versión presenta índices de fiabilidad y validez tan buenos como su predecesor, el BDI-IA y, en algunos casos, claramente superiores como, por ejemplo, respecto a la validez factorial (Beck, Steer, Ball y Ranieri, 1996; Beck, Steer y Brown, 1996; Dozois, Dobson y Ahnberg, 1998; Steer, Clark, Beck y Ranieri, 1999; Whisman, Perez y Ramel, 2000).

El objetivo principal del presente estudio es desarrollar una versión española del BDI-II y ofrecer información sobre sus propiedades psicométricas de fiabilidad, validez de contenido, validez de criterio, validez factorial y normas en una muestra de estudiantes universitarios. El estudio forma parte de un proyecto más amplio que pretende adaptar el BDI-II en población española y que implica el análisis de las propiedades psicométricas de la versión en castellano del BDI-II que aquí se presenta en muestras de la población general (Sanz, Perdigón y Vázquez, 2003) y de pacientes con trastornos psicológicos (Sanz, García-Vera y Vázquez, 2003).

Por otro lado, el análisis psicométrico del BDI-II en muestras de estudiantes universitarios permite aportar un instrumento actualizado para la detección de probables casos de depresión en investigaciones con estudiantes así como, más en concreto, para la detección de análogos subclínicos para los estudios de validación de los modelos psicológicos de la depresión. Gran parte del aval empírico que apoya tales modelos proviene de estudios realizados con estudiantes universitarios subclínicamente deprimidos, los cuales son así clasificados en función de la puntuación obtenida en una determinada escala de depresión que, en la mayoría de los estudios, es el BDI-I o el BDI-IA (Vredenburg, Flett y Krames,

1993; Tennen, Hall y Affleck, 1995; Tennen, Eberhardt y Affleck, 1999). Esta estrategia metodológica es muy habitual tanto en la investigación internacional como en la nacional (Sanz y Vázquez, 1998).

Por último, y no menos importante, la adaptación de un instrumento para la evaluación de la depresión en muestras de estudiantes universitarios merece atención por sí misma en cuanto hace referencia a un grupo de edad en el que la depresión crece con una extraordinaria rapidez, como demuestran las denominadas "tendencias seculares" del trastorno (Fombonne, 1994; Seligman, 1990; Vázquez y Siles, 1999). Por ejemplo, un estudio epidemiológico realizado recientemente con estudiantes universitarios españoles empleando la entrevista diagnóstica estructurada DIS-III-R encontró que la prevalencia anual del episodio depresivo mayor era de 1,9% (Bethencourt, Peñate, González, Fumero, Fernández-Valdés y Díaz, 1998). Teniendo en cuenta que la población de estudiantes universitarios españoles menores de 30 años es de aproximadamente 1.417.470 personas (Consejo de Coordinación Universitaria, 2002), ese porcentaje de prevalencia supone que en España, al año, 26.932 estudiantes universitarios sufren episodios depresivos mayores, cifra que por sí misma justificaría la mejora de los instrumentos para su evaluación. De hecho, conviene recordar que, como han demostrado Lewinsohn, Hoberman y Rosenbaum (1988) o Horwath, Johnson, Klerman, y Weissman (1994), las personas que manifiestan niveles altos de sintomatología depresiva tienen un riesgo elevado de sufrir en el futuro un trastorno depresivo mayor, especialmente si además están sometidos a un elevado estrés (como suele ser el caso de los estudiantes universitarios). La importancia de una evaluación temprana de la sintomatología depresiva queda reflejada en las siguientes palabras de Horwath et al. (1994): "si los síntomas depresivos pudieran identificarse y tratarse antes de que se desarrolle por primera vez la depresión mayor, se podrían prevenir en potencia muchos casos nuevos de depresión mayor" [...] "las medidas diseñadas para identificar este grupo de personas antes del inicio de la depresión mayor pueden representar un paso importante en su prevención primaria" (p. 333).

MÉTODO

Participantes

La muestra inicial de participantes estaba formada por 613 estudiantes universitarios que completaron el BDI-II. No se analizaron los cuestionarios en los que faltaban datos demográficos básicos (edad o sexo) o la respuesta a más de un ítem del BDI-II. La muestra final ($N = 590$) quedó conformada por 131 varones y 459 mujeres que cursaban estudios universitarios en Madrid y que procedían de la Residencia Universitaria Sagrado Corazón y de diversas Facultades de la Universidad Complutense de Madrid y de la Universidad San Pablo-CEU. El rango de edades de la muestra total final oscilaba entre 18 y 58 años, con una media de 21,9 años ($DT = 3,8$). En la Tabla 3 se presenta información adicional sobre la edad de los participantes y su distribución por estudios universitarios.

Puesto que este estudio forma parte de un proyecto más amplio que pretende adaptar el BDI-II en todo tipo de población adulta, no sólo estudiantes universitarios, para comprobar la comprensibilidad y aplicabilidad de la traducción española del BDI-II se seleccionó un grupo de 10 varones y 10 mujeres de la población general que no habían cursado o no estaban cursando estudios universitarios, y cuya edad era más heterogénea (10% mayores de 65 años; 25% entre 45 y 65 años; 45% entre 25 y 44 años, y 20% entre 20 y 24 años).

Instrumentos

Inventario para la Depresión de Beck—Segunda Edición (Beck, Steer y Brown, 1996). El BDI-II es un instrumento de autoinforme de 21 ítems diseñado para evaluar la gravedad de la sintomatología depresiva en adultos y adolescentes con una edad mínima de 13 años. En cada uno de los ítems la persona tiene que elegir, entre un conjunto de cuatro alternativas ordenadas de menor a mayor gravedad, la frase que mejor describe su estado durante las últimas dos semanas incluyendo el día en que completa el instrumento. En cuanto a su corrección, cada ítem se valora de 0 a 3 puntos en

Tabla 3.- Características demográficas de la muestra de estudiantes universitarios y medias y desviaciones típicas (DT) de las puntuaciones del BDI-II en función de dichas características

| Variables Demográficas | Porcentaje | <i>n</i> | Media | DT |
|--------------------------------|--------------|------------|------------|------------|
| Sexo | | | | |
| Hombres | 22,2 % | 131 | 9 | 7,5 |
| Mujeres | 77,8 % | 459 | 9,3 | 7,5 |
| Edad | | | | |
| 18-24 años | 91,8 % | 542 | 9,3 | 7,5 |
| 25-58 años | 8,4 % | 48 | 8,8 | 7,8 |
| Estudios Universitarios | | | | |
| Psicología | 63,4 % | 374 | 8,5 | 7,3 |
| Derecho | 9,7 % | 57 | 10,2 | 7,4 |
| Periodismo | 8,3 % | 49 | 11,6 | 9,1 |
| Ingenierías | 4,9 % | 29 | 12,6 | 6,2 |
| Ciencias de la Salud | 5,3 % | 31 | 10,8 | 9,4 |
| Otras | 8,5 % | 50 | 7,9 | 5,6 |
| Muestra Total | 100 % | 590 | 9,2 | 7,5 |

función de la alternativa escogida y, tras sumar directamente la puntuación de cada ítem, se puede obtener una puntuación total que varía de 0 a 63. A veces se da la circunstancia de que la persona elige más de una alternativa en un ítem dado. En este caso se toma la puntuación de la frase elegida de mayor gravedad. Varios estudios psicométricos avalan la fiabilidad y validez del BDI-II en muy diversas muestras: pacientes psiquiátricos (Beck, Steer y Brown, 1996), pacientes con trastornos depresivos (Steer, Ball, Ranieri y Beck, 1999), adolescentes (Coelho, Martins y Barros, 2002), drogodependientes (Buckley, Parker y Heggie, 2001), ancianos (Jefferson, Powers y Pope, 2001), pacientes de atención primaria (Arnau, Meagher, Norris y Bramson, 2001) y estudiantes universitarios (Al Musawi, 2001; Beck, Steer y Brown, 1996; Dozois, Dobson y Ahnberg, 1998; O'Hara, Sprinkle y Ricci, 1998; Osman, Downs, Barrios, Kopper, Gutierrez y Chiro, 1997; Schulenberg y Yutrzenka, 2001; Steer y Clark, 1997; Whisman, Perez y Ramel, 2000).

Guión de Entrevista Diagnóstica Rápida III-R (Quick Diagnostic Interview Schedule III-R o Quick DIS-III-R; Marcus, Robins y Bucholz, 1991; Bucholz, Marion, Shayka, Marcus y Robins, 1996). El Quick DIS-III-R es una entrevista diagnóstica estructurada administrada por ordenador, bien de forma autoaplicada o heteroaplicada, que realiza diagnósticos psicopatológicos vitales y en el último año según criterios DSM-III-R (APA, 1987). La Quick DIS-III-R se basa en el *NIMH Diagnostic Interview Schedule Version III Revised (DIS-III-R; Robins, Helzer, Cottler y Goldring, 1989)* y, de hecho, todas las preguntas del Quick DIS-III-R provienen del DIS-III-R y la mayoría de ellas son idénticas. El DIS-III-R es una entrevista diagnóstica estructurada creada inicialmente para la realización de estudios epidemiológicos a gran escala, de forma que pudiera ser aplicada por personal sin formación clínica debidamente entrenado. El DIS-III-R, y su predecesor, el DIS, han sido utilizados en múltiples investigaciones, traducidos a varios idiomas y analizados en varios estudios en cuanto a sus propiedades psicométricas (Robins y Helzer, 1994). En España, por ejemplo, al menos dos estudios han empleado el DIS-III-R con fines epidemiológicos (Bethencourt, Peñate, Fumero y González, 1997; Bethencourt et al., 1998; López-García et al., 1996), y al menos otros dos han examinado su fiabilidad y validez con

resultados aceptables (Bethencourt, Peñate, González y Fumero, 1996; Pablo Alcázar et al., 1992). Tomando como criterio el DIS-III-R, el Quick DIS-III-R presenta una alta sensibilidad y especificidad para la mayoría de los diagnósticos (0,8 o más), y excelentes índices de acuerdo, la mayoría con kappas por encima de 0,70 (Bucholz et al., 1996). En concreto, respecto al episodio depresivo mayor se han encontrado índices de sensibilidad, especificidad y de acuerdo diagnóstico de 1, 0,67 y 0,76, respectivamente (Bucholz et al., 1996).

El Quick DIS-III-R está organizada en 25 módulos diagnósticos que son independientes entre sí, de los cuales, en esta investigación, sólo se administraron los 15 que permitían obtener diagnósticos de los siguientes trastornos: trastorno de angustia, trastorno de ansiedad generalizada, agorafobia, fobia social, fobia simple, trastorno por estrés postraumático, trastorno obsesivo-compulsivo, episodio maníaco, trastorno esquizofreniforme-esquizofrenia, anorexia, bulimia, abuso-dependencia de alcohol, y episodio depresivo mayor. Además, puesto que para los objetivos de este estudio se requería contar con información sobre la prevalencia puntual de los trastornos y, de manera especial, del episodio depresivo mayor, se modificaron los parámetros temporales de las preguntas del Quick DIS-III-R de forma que, por ejemplo, para el episodio depresivo mayor se indagaba por la presencia de sintomatología depresiva en las últimas dos semanas. Dado que el BDI-II pretende cubrir los criterios sintomáticos de los trastornos depresivos del DSM-IV, se podría cuestionar la conveniencia de utilizar en este estudio el Quick DIS-III-R en lugar de una entrevista diagnóstica que siguiera los criterios del DSM-IV. Aunque esto último hubiera sido deseable, no resulta imprescindible para los objetivos del estudio ya que los criterios sintomáticos del episodio depresivo mayor no han cambiado en el DSM-IV respecto a su predecesor el DSM-III-R (APA, 1987, 1994).

Procedimiento

El desarrollo de la traducción española del BDI-II se realizó en varias etapas, utilizando varios traductores, el procedimiento de traducción inversa (*back translation*) y un estudio piloto de comprensibilidad de sus instrucciones e ítems y de su facilidad de aplicación.

En primer lugar, cada uno de los autores de este trabajo realizó de manera independiente una traducción al español del BDI-II. Estas tres traducciones se compararon entre sí y a su vez con la versión española del BDI-IA, y tras discutir las discrepancias encontradas se obtuvo una traducción única al español.

En una segunda fase, la versión española fue traducida al inglés por un psicólogo clínico estadounidense bilingüe, cuya lengua materna es el inglés y su campo de especialización la depresión. Se comparó esta traducción inglesa con el original, se resolvieron las discrepancias, que tan sólo afectaban al 5,1% de las 84 frases de que consta el BDI-II, y se llegó a una traducción española consensuada cuya comprensibilidad y aplicabilidad fue evaluada posteriormente en un estudio piloto.

En este estudio piloto se aplicó de forma individual la traducción española del BDI-II a 20 personas de la población general que no habían cursado o no estaban cursando estudios universitarios, pidiéndoles que describieran las dificultades de comprensión o cumplimentación que encontraran tanto en las instrucciones como en cada uno de los ítems. A tenor de los comentarios obtenidos en este estudio piloto, se realizaron algunos cambios en la traducción española, siendo los más relevantes los introducidos en las instrucciones para mejorar su comprensibilidad y un ligero cambio realizado en el ítem 1. En este ítem, la primera alternativa "No me siento triste" fue sustituida por "No me siento triste habitualmente", ya que el 25% de la muestra tuvo dificultades para elegir la alternativa que mejor le describía porque consideraban que la primera traducción no resultaba creíble ya que no es posible no estar triste alguna vez.

Para su posterior análisis psicométrico, a la traducción española final del BDI-II se le añadieron los cuatro ítems del BDI-IA (pérdida de peso, cambio en la imagen corporal, preocupación somática y dificultades en el trabajo) que habían sido omitidos y reemplazados por otros nuevos, con un doble objetivo: (a) contar con posibles ítems alternativos en el caso de que algún ítem del BDI-II no obtuviera índices de fiabilidad y validez adecuados, y (b) comprobar si la decisión de omitir y reemplazar esos ítems era acertada en el caso de la población española.

En definitiva, la versión española del BDI-II que se administró a todos los participantes de este estudio contaba con 25 ítems, 21 del BDI-II original y cuatro del BDI-IA. La mayoría de los estudiantes de la muestra final ($n = 425$) completaron el BDI-II de forma individual entre los meses de enero y mayo de 2001. Los restantes 165 participantes, todos ellos estudiantes de la Facultad de Psicología de la Universidad Complutense de Madrid, fueron evaluados en grupos de 20-30 personas como parte de las prácticas de una de sus asignaturas entre los meses de marzo y mayo de 2001. Estos 165 estudiantes, tras completar el BDI-II, respondieron a la entrevista diagnóstica Quick DIS-III-R de forma autoaplicada a través de un ordenador.

RESULTADOS Y DISCUSIÓN

Distribución de las puntuaciones del BDI-II

Las puntuaciones totales, obtenidas tras sumar los 21 ítems originales del BDI-II, oscilaron entre 0 y 46, con una media igual a 9,2 y una desviación típica igual a 7,5. Estos resultados son consistentes con los que presenta la literatura científica sobre el BDI-II en muestras de estudiantes universitarios y que aparecen resumidos en la Tabla 4. En total, en los estudios previos recogidos en dicha tabla se ha administrado el BDI-II a 2623 estudiantes, y la media y desviación típica conjuntas (ponderadas por el número de participantes en cada estudio) son prácticamente iguales a las obtenidas en este estudio (9,3 frente a 9,2 y 7,4 frente a 7,5, respectivamente). Por otro lado, y como viene siendo habitual en los estudios sobre el BDI (Kendall, Hollon, Beck, Hammen e Ingram, 1987; Sanz y Vázquez, 1998; Vázquez y Sanz, 1997), la curva de distribución de las puntuaciones del BDI-II estaba desplazada hacia los valores más altos (índice de simetría = 1,4) y apuntada en los valores más bajos (curtosis = 2,5), de forma que si bien la mediana no era muy distinta de la media (7 frente a 9,2), el valor más frecuente en la muestra (moda = 5) era notablemente más bajo que dicha media.

Tabla 4.- Propiedades psicométricas del BDI-II en distintos estudios con muestras de estudiantes universitarios

| Estudio | Características de la muestra | | | | Puntuaciones del BDI-II | | | | Fiabilidad del BDI-II | | | |
|---|-------------------------------|------|-------------------|-------------------|---|------------------|-------------------|-------------------|-----------------------|--|------------------|---------------|
| | País | N | % de mujeres | Edad media | % por categorías de gravedad de depresión | | | | Alfa | Rango de correlaciones corregidas ítem-total | | |
| | | | | | M | DT | Mínima (0-13) | Leve (14-19) | | | Moderada (20-28) | Grave (29-63) |
| Boek, Steer y Brown (1996) | Canadá | 120 | 56 | 20 | 12,6 | 9,9 | 64 | 15 | 13 | 8 | 0,92 | 0,27 - 0,74 |
| Steer y Clark (1997) | Canadá | 160 | 67 | 19 | 11,9 | 8,1 | 56 | 29 | 13 | 2 | 0,89 | 0,37 - 0,62 |
| O'Hara et al. (1998) | EE. UU. | 152 | 52 | 20 | 8,8 | 8,1 | s.d. | s.d. | s.d. | s.d. | s.d. | s.d. |
| Dozoris et al. (1998) ^a | Canadá | 1022 | 67 | 21 | 9,1 | 7,6 | 77,2 | 13,4 | 6,9 | 2,4 | 0,91 | 0,41 - 0,62 |
| Whisman et al. (2000) | EE. UU. | 576 | 58 | 19 | 8,4 | 7,2 | 80,2 | 12,5 | 5,7 | 1,6 | 0,89 | 0,28 - 0,63 |
| Aasen (2001) | Noruega | 303 | 70 | 21 | 7,1 | 6 | 89 | 5,8 | 4,2 | 1 | 0,86 | 0,31 - 0,61 |
| Al-Mustawí (2001) | Bahrein | 200 | 63 | 23 | 13,4 | 6,7 | 53 | 32 | 14 | 1 | 0,84 | s.d. |
| Schulenberg y Yairzenka (2001) ^b | EE. UU. | 90 | 78,3 ^c | 22 ^c | 8,8 | 6,8 | s.d. | s.d. | s.d. | s.d. | 0,88 | s.d. |
| Estudios anteriores | - | 2623 | 64 ^d | 20,5 ^e | 9,3 ^d | 7,4 ^d | 75,3 ^d | 14,8 ^d | 7,6 ^e | 2,2 ^d | 0,88 | - |
| Este estudio (2002) | España | 590 | 78 | 22 | 9,2 | 7,5 | 77,8 | 12,7 | 6,6 | 2,9 | 0,89 | 0,36 - 0,60 |

Tabla 4.- (Continuación)

- Nota. ^a Los porcentajes por categorías de gravedad han sido proporcionados por Dozois (comunicación personal, 3 de noviembre de 2002).
- ^b En este estudio se aplicó dos versiones del BDI-II, una en formato escrito y otra en formato por ordenador, mediante un diseño que contrabalanceaba su orden de aplicación. Salvo que se indique lo contrario, sólo se presentan los datos de los dos grupos de estudiantes universitarios que completaron en primer lugar la versión escrita del BDI-II.
- ^c Los datos demográficos se refieren a los cuatro grupos de estudiantes universitarios que participaron en este estudio.
- ^d Valor medio ponderado por el número de participantes en cada estudio.
- ^e Coeficiente medio ponderado por el número de participantes en cada estudio según la fórmula propuesta por Rosenthal (1983, p. 8) basada en la transformación Z de Fisher de los coeficientes.
- s.d.: sin datos.

Tabla 5.- Frecuencia, media (M), desviación típica (DT) y correlaciones corregidas ítem-total (r_{tot}) de los ítems del BDI-II

| Ítems | Escala de Gravedad (en %) | | | M | DT | r_{tot} | |
|---------------------------------|---------------------------|------|-----|-----|------|------------------|------|
| | 0 | 1 | 2 | | | | 3 |
| 1. Tristeza | 80,5 | 16,4 | 2,7 | 0,3 | 0,23 | 0,5 | 0,57 |
| 2. Pesimismo | 59,2 | 37,3 | 3,1 | 0,5 | 0,45 | 0,6 | 0,44 |
| 3. Sentimientos de fracaso | 78 | 19,2 | 2,7 | 0,2 | 0,25 | 0,5 | 0,47 |
| 4. Pérdida de Placer | 72 | 24,1 | 3,7 | 0,2 | 0,32 | 0,6 | 0,5 |
| 5. Sentimientos de culpa | 62,7 | 34,9 | 1,9 | 0,5 | 0,4 | 0,6 | 0,51 |
| 6. Sentimientos de Castigo | 87,1 | 10,3 | 0,5 | 2 | 0,17 | 0,5 | 0,46 |
| 7. Insatisfacción con uno mismo | 70,2 | 22,2 | 4,7 | 2,9 | 0,4 | 0,7 | 0,59 |
| 8. Auto-críticas | 45,9 | 40,2 | 12 | 2,2 | 0,7 | 0,8 | 0,55 |
| 9. Pensamientos de suicidio | 89,5 | 10 | 0,5 | 0 | 0,11 | 0,3 | 0,4 |
| 10. Llanto | 79,2 | 13,1 | 5,8 | 2 | 0,31 | 0,7 | 0,5 |
| 11. Agitación | 47,6 | 41,5 | 6,3 | 4,6 | 0,68 | 0,8 | 0,45 |
| 12. Pérdida de interés | 62,7 | 33,2 | 3,4 | 0,7 | 0,42 | 0,6 | 0,5 |
| 13. Indecisión | 71,7 | 18 | 4,2 | 6,1 | 0,45 | 0,8 | 0,51 |
| 14. Inutilidad | 77,8 | 10,3 | 11 | 0,5 | 0,34 | 0,7 | 0,6 |
| 15. Pérdida de energía | 49,8 | 39,5 | 10 | 0,5 | 0,61 | 0,7 | 0,59 |

Tabla 5.- (Continuación)

| | | | | | | | |
|-----------------------------------|------|------|-----|-----|------|-----|------|
| 16. Cambios en el patrón de sueño | 36,6 | 51 | 9,5 | 2,9 | 0,79 | 0,7 | 0,4 |
| a. Incremento | 77,6 | 19 | 2,5 | 0,8 | | | |
| b. Decremento | 59 | 32 | 6,9 | 2 | | | |
| 17. Irritabilidad | 62,4 | 32,9 | 3,7 | 1 | 0,43 | 0,6 | 0,54 |
| 18. Cambios en el apetito | 50 | 41,2 | 4,6 | 4,2 | 0,63 | 0,8 | 0,38 |
| a. Incremento | 74,4 | 19,5 | 2,4 | 3,7 | | | |
| b. Decremento | 75,1 | 21,9 | 2,7 | 0,3 | | | |
| 19. Dificultad de concentración | 44,2 | 33,9 | 21 | 1 | 0,79 | 0,8 | 0,54 |
| 20. Cansancio o fatiga | 51,5 | 39 | 9 | 0,5 | 0,58 | 0,7 | 0,57 |
| 21. Pérdida de interés en el sexo | 87,8 | 9,5 | 2,2 | 0,5 | 0,15 | 0,5 | 0,36 |

Como puede verse en la Tabla 5, la puntuación media de todos los ítems fue 0,41, con un mínimo de 0,11 y un máximo de 0,78, y los siguientes ítems recibieron las mayores puntuaciones en frecuencia e intensidad: Cambios en el patrón de sueño, Dificultad de concentración y Auto-críticas. Por el contrario, los síntomas que se mostraban de forma más ligera y con menor frecuencia fueron Pensamientos de suicidio, Pérdida de interés en el sexo y Sentimientos de castigo. Estos resultados son muy semejantes a los encontrados en otras muestras de estudiantes universitarios. Por ejemplo, tanto en muestras canadienses (Beck, Steer y Brown, 1996; Dozois et al., 1998; Steer y Clark, 1997) como estadounidenses (Whisman et al., 2000), entre los cuatro ítems con mayor frecuencia e intensidad se encontraban Dificultad de concentración y Cambios en el patrón de sueño, siendo este último el que siempre mostraba la puntuación más alta de los 21 ítems del BDI-II, tal y como ocurría en la muestra española. Igualmente, en esos cuatro estudios también se encontró que los tres ítems con menor puntuación en intensidad y frecuencia fueron Pensamientos de suicidio, Pérdida de interés en el sexo y Sentimientos de castigo.

Diferencias demográficas y normas para estudiantes universitarios

Las medias y desviaciones típicas de la puntuación total de los 21 ítems originales del BDI-II para la muestra dividida por sexo, edad y tipo de estudios aparecen en la Tabla 3. Puesto que algunos de los grupos resultantes de combinar los niveles de las variables sexo (varón, mujer), edad (18 a 24 años, 25 a 58 años) y tipo de estudios (Psicología, no Psicología) estaban compuestos por muy pocos participantes comprometiendo, por tanto, la fiabilidad de los resultados de un ANOVA 2 x 2 x 2 sobre la puntuación total del BDI-II, se decidió realizar un ANCOVA 2 x 2 con la edad como covariable y las otras dos variables, sexo y tipo de estudios, como factores. Este análisis no reveló ningún efecto de la edad $_F(1, 585) = 1,95, n.s._$, ni ninguna diferencia estadísticamente significativa entre varones y mujeres $_F(1, 585) = 0,92, n.s._$ o entre estudiantes de psicología y estudiantes de otras carreras $_F(1, 585) = 3,00, n.s._$.

Tabla 6.- Puntuaciones centiles del BDI-II para la muestra total (N = 590) de estudiantes universitarios

| Centiles | BDI-II | Centiles | BDI-II | Centiles | BDI-II |
|----------|--------|----------|--------|----------|--------|
| 1 | 0 | 30 | 5 | 75 | 13 |
| 5 | 1 | 40 | 6 | 80 | 14 |
| 10 | 2 | 50 | 7 | 90 | 19 |
| 20 | 3 | 60 | 9 | 95 | 24 |
| 25 | 4 | 70 | 12 | 99 | 33 |

Tampoco se encontró ningún efecto significativo de la interacción sexo y tipo de estudios [$F(1, 585) = 0,64, n.s.$].

La ausencia de una diferencia estadísticamente significativa en el BDI-II entre mujeres y varones es consistente con los resultados obtenidos en la mayoría de los estudios previos con estudiantes universitarios (Aasen, 2001; Al-Musawi, 2001; Dozois et al., 1998; O'Hara et al., 1998; Schulenberg y Yutrzeuka, 2001; Steer y Clark, 1997), aunque los estudios iniciales de Beck, Steer y Brown (1996) y Osman et al. (1997) parecían sugerir que las mujeres puntuaban más alto en el BDI-II que los varones.

Para analizar con más detalle si el tipo de estudios cursados afectaba a las puntuaciones en el BDI-II, también se realizó un ANCOVA 2 x 5 con dos factores, sexo y tipo de estudios (Psicología, Derecho, Periodismo, Ingenierías y Ciencias de la Salud), y la edad como covariable. De nuevo, el tipo de estudios no mostró ningún efecto significativo ni por sí mismo [$F(4, 529) = 2,01, n.s.$], ni en interacción con el sexo [$F(4, 529) = 0,62, n.s.$].

Puesto que no se descubrieron diferencias significativas en la puntuación total del BDI-II debidas al sexo, la edad o el tipo de estudios, en la Tabla 6 se presentan puntuaciones normativas (en centiles) obtenidas a partir de la muestra total de estudiantes universitarios.

Rangos cuantitativos de depresión en el BDI-II

En la Tabla 7 se presentan los porcentajes de estudiantes universitarios que se corresponden con las diferentes categorías de gravedad de sintomatología depresiva (mínima, leve, moderada y grave) propuestas por Beck, Steer y Brown (1996) en el manual del BDI-II, así como las puntuaciones medias en el BDI-II para cada categoría de gravedad. Los puntos de corte que definen esas categorías de gravedad fueron obtenidos por Beck, Steer y Brown (1996) a partir de una muestra de pacientes con trastornos psicológicos mediante la utilización de las curvas ROC (*Receiver Operating Characteristic*) o curvas de rendimiento predictivo, pero, como los propios autores reconocen, "la decisión de usar puntos de corte

diferentes para el BDI-II debe basarse en las características particulares de la muestra y el propósito para que se usa el BDI-II" (p. 11, Beck, Steer y Brown, 1996). En este sentido, la mayoría de las investigaciones que han utilizado la estrategia metodológica de clasificar a estudiantes universitarios como subclínicamente deprimidos y no deprimidos mediante el BDI han utilizado los puntos de corte propuestos por Kendall, Hollon, Beck, Hammen e Ingram (1987) para el BDI-IA (Haaga y Solomon, 1993; Tennen, Eberhardt y Affleck, 1999; Tennen, Hall y Affleck, 1995). Dozois et al. (1998) han calculado empíricamente los valores del BDI-II que se corresponden con los puntos de corte del BDI-IA de Kendall et al. (1987) para así poder disponer de puntos de corte para el BDI-II que sean relevantes para los investigadores que utilizan esa estrategia metodológica, y que a su vez permitan la comparación directa de sus resultados con los estudios que hayan utilizado el BDI-IA. En la Tabla 7, se muestran los puntos de corte propuestos por Dozois et al. (1998) así como los porcentajes de estudiantes universitarios españoles que se corresponden con las categorías de gravedad definidas por dichos puntos, y las puntuaciones medias en el BDI-II que obtienen los estudiantes de cada categoría.

Independientemente de los puntos de corte elegidos, los de Beck, Steer y Brown (1996) o los de Dozois et al. (1998), la distribución de los estudiantes universitarios españoles en categorías de gravedad de depresión era similar a la obtenida en estudios anteriores con ese tipo de muestras. Los resultados de una prueba de chi cuadrado realizada sobre los porcentajes correspondientes a las categorías de gravedad de Beck, Steer y Brown (1996) que se recogen en la Tabla 4 indicaban que no había diferencias estadísticamente significativas entre los porcentajes hallados en el presente estudio y los porcentajes medios ponderados obtenidos de los datos de los estudios previos realizados con estudiantes universitarios [$\chi^2(3, N = 2971) = 3,33, n.s.$]. Igualmente, los resultados de otra prueba de chi cuadrado realizada para comparar los porcentajes correspondientes a las categorías de gravedad de Dozois et al. (1998) obtenidos en el presente estudio (véase la Tabla 7) y aquellos obtenidos en el propio estudio de Dozois et al. (1998; no deprimido = 73,7%; disfórico = 16,9%; disfórico o deprimido = 9,4%) no mostraron diferencias estadísticamente significativas.

Tabla 7.- Rangos cuantitativos de depresión del BDI: Porcentajes y valores medios

| Criterios de Depresión | | Porcentaje | n | Media | DT |
|----------------------------|-------|------------|-----|-------|------|
| Beck, Steer y Brown (1996) | | | | | |
| Mínima | 0-13 | 77,8 % | 459 | 5,99 | 3,71 |
| Leve | 14-19 | 12,7 % | 75 | 16,28 | 1,69 |
| Moderada | 20-28 | 6,6 % | 39 | 23,1 | 2,55 |
| Grave | 29-63 | 2,9 % | 17 | 33,59 | 5,01 |
| Dozois et al. (1998) | | | | | |
| No deprimido | 0-12 | 73,6 % | 434 | 5,59 | 3,4 |
| Distórico | 13-19 | 16,9 % | 100 | 15,46 | 2,04 |
| Distórico o deprimido † | 20-63 | 9,5 % | 56 | 26,28 | 5,95 |

Nota. DT = Desviación típica. † Dependiendo de si cumple o no cumple los criterios diagnósticos para un episodio depresivo mayor o un trastorno depresivo.

mido = 9,4%) indicaban que dichos porcentajes no diferían de forma estadísticamente significativa [$\chi^2(2, N = 1612) = 0,01, n.s.$].

Consistencia interna

El análisis de la consistencia interna de los 21 ítems originales del BDI-II arrojó un coeficiente alfa de 0,89, lo que indica una muy buena consistencia interna para el BDI-II y replica los coeficientes encontrados en la literatura con muestras semejantes (véase la Tabla 4). Las correlaciones entre las puntuaciones en cada uno de los ítems y la puntuación total corregida en el BDI-II (es decir, la puntuación total sin tener en cuenta el ítem en cuestión) se presentan en la Tabla 5. Los coeficientes de correlación hallados fueron todos estadísticamente significativos, oscilando entre 0,35 para el ítem de pérdida de interés en sexo y 0,60 para el ítem de inutilidad. La media de las correlaciones interítems fue 0,28 ($DT = 0,08$), con un mínimo de 0,09 y un máximo de 0,61.

Comparación entre el BDI-II y el BDI-IA

Comparando los resultados encontrados con el BDI-II con los obtenidos con la adaptación española del BDI-IA en muestras de universitarios (Sanz y Vázquez, 1998), llama la atención que la media de las puntuaciones totales del BDI-II es superior en más de tres puntos a la media del BDI-IA (9,2 frente a 5,4). Esta diferencia es consistente con los resultados hallados en estudios previos en todo tipo de muestras. Por ejemplo, respecto a los estudiantes universitarios, la media conjunta que se calculó a partir de los estudios del BDI-II recogidos en la Tabla 4 era de 9,3 ($N = 2623$), mientras que la media conjunta (ponderada por el número de participantes) de los ocho estudios con estudiantes universitarios recogidos por Kendall y Sheldrick (2000) en su revisión de datos normativos del BDI-IA fue de 7,6 ($N = 3030$). De hecho, el único estudio con estudiantes universitarios que ha administrado simultáneamente el BDI-II y el BDI-IA (Dozois et al., 1998) ha corroborado que la pun-

tuación media del primero es superior de forma estadísticamente significativa a la del segundo (9,1 frente a 7,4). En la misma línea, los estudios llevados a cabo por Beck y sus colaboradores (Beck, Steer y Brown, 1996; Beck, Steer, Ball y Ranieri, 1996) administrando simultáneamente el BDI-II y el BDI-IA a muestras de pacientes con trastornos psicológicos también corroboran que la puntuación media del BDI-II es 2-3 puntos mayor que la del BDI-IA, y que la diferencia es estadísticamente significativa.

Por otro lado, cabría plantearse si la eliminación en el BDI-II de cuatro ítems del BDI-IA y su reemplazo por cuatro nuevos estarían justificados en la muestra de estudiantes universitarios españoles. Beck, Steer y Brown (1996) argumentaban en el manual del BDI-II que los ítems originales del BDI-IA sobre cambio en la imagen corporal, pérdida de peso y preocupación somática habían sido eliminados porque tales ítems, que "típicamente se habían observado en pacientes hospitalizados a largo plazo... con el paso de los años habían llegado a ser gradualmente menos útiles para evaluar la gravedad de la depresión en pacientes externos o en pacientes hospitalizados a corto plazo" (p. 3). Aunque en principio este argumento podría extenderse a la sintomatología depresiva que experimentan los estudiantes universitarios, parece conveniente comprobar empíricamente la plausibilidad de tal afirmación. Para ello, se calculó el índice de homogeneidad de los ítems eliminados calculando su correlación con la puntuación total corregida de una versión del BDI-II en la que se habían sustituido los ítems nuevos del BDI-II por los ítems originales del BDI-IA que habían sido eliminados. Las correlaciones ítem-total obtenidas para los ítems eliminados fueron 0,45 (Cambio en la imagen corporal), 0,47 (Dificultades laborales), 0,38 (Preocupación somática) y 0,07 (Pérdida de peso), correlaciones que, en general, fueron inferiores a los que mostraban los nuevos ítems del BDI-II (Agitación, Inutilidad, Pérdida de energía y Dificultad de concentración) que oscilaban entre 0,45 y 0,60 (véase la Tabla 5). De hecho, la consistencia interna de la versión del BDI-II que reemplazaba los ítems nuevos por los ítems originales del BDI-IA eliminados era ligeramente inferior a la que mostraba el BDI-II original (0,87 frente a 0,89).

Finalmente, la consistencia interna del BDI-II es superior a la que muestra el BDI-IA en estudiantes universitarios españoles (0,89 frente a 0,83; Sanz y Vázquez, 1998), diferencia que, según el test de Alsawalmeh y Feldt (1992) para la igualdad de dos coeficientes alfa independientes, fue estadísticamente significativa [$T(727,5, 1683,6) = 1,55, p < 0,001$]. Esta superioridad también es consistente con la literatura previa. En su revisión de investigaciones realizadas con el BDI-I y el BDI-IA, Beck et al. (1988) localizaron ocho trabajos que informaban de coeficientes de consistencia interna obtenidos con un total de 2435 estudiantes universitarios. A partir de esos estudios se puede calcular que, para el BDI-I y el BDI-IA, la consistencia interna media (ponderada por el número de participantes en cada uno de ellos según la fórmula sugerida por Rosenthal, 1983, p. 8) es 0,85, mientras que el coeficiente alfa medio de los siete estudios sobre el BDI-II que se recogen en la Tabla 4 es 0,88.

Validez de contenido

El BDI-II surgió con la intención de evaluar los síntomas que se recogen en los criterios diagnósticos de los trastornos depresivos del DSM-IV (APA, 1994): el trastorno depresivo mayor y el trastorno distímico. Por tanto, cabría esperar que los nueve criterios sintomatológicos del episodio depresivo mayor y los siete criterios sintomatológicos del trastorno distímico (cinco de ellos comunes a los del episodio depresivo mayor) tendrían su reflejo en uno o varios de los ítems del BDI-II. Como puede observarse en la Tabla 8, este objetivo parece haberse cumplido plenamente. El BDI-II cubre los siete síntomas que definen el trastorno distímico y ocho de los nueve del episodio depresivo mayor y, aunque parcialmente, también refleja el síntoma restante (agitación o enlentecimiento psicomotores). Efectivamente, el BDI-II evalúa los cambios de la actividad psicomotora que suponen un aumento en dicha actividad (ítem de Agitación), pero no su enlentecimiento o retardo (p. ej., lenguaje, pensamiento y movimientos corporales enlentecidos; aumento de la latencia de respuesta; bajo volumen de la voz; menos inflexiones y cantidad o variedad de contenido en el lenguaje; mutismo).

Tabla 8.- Comparación entre los criterios diagnósticos sintomáticos del DSM-IV para los trastornos depresivos y los ítems del BDI-II y del BDI-IA

| DSM-IV (1994) * | BDI-IA (Beck et al., 1979) | BDI-II (Beck, Steer y Brown, 1996) |
|--|---|--|
| 1. Estado de ánimo deprimido (o irritable en niños y adolescentes) | (1) Tristeza (10) Llanto (11) Irritabilidad | (1) Tristeza (10) Llanto (17) Irritabilidad |
| 2. Disminución del placer o del interés | (4) Insatisfacción general (12) Desinterés social (21) Desinterés por el sexo | (4) Pérdida de placer (12) Pérdida de interés (21) Pérdida de interés en el sexo |
| 3. Aumento/disminución del peso o del apetito | (18) Disminución del apetito (19) Disminución del peso | (18) Cambios en el apetito |
| 4. Insomnio o hipersomnia | (16) Insomnio | (16) Cambios en el patrón de sueño |
| 5. Agitación o enlentecimiento psicomotor | (17) Fatiga | (11) Agitación |
| 6. Fatiga o pérdida de energía | (17) Fatiga | (15) Pérdida de energía (20) Cansancio o fatiga |

Tabla 8.- (Continuación)

| | | |
|---|---|---|
| 7. Sentimientos excesivos de inutilidad o culpa | (3) Sentimiento de fracaso (6) Sentimiento de castigo (5) Culpabilidad (8) Autoculpación | (3) Sentimientos de fracaso (6) Sentimientos de castigo (5) Sentimientos de culpa (8) Autoctricas (14) Inutilidad |
| 8. Disminución de la capacidad para pensar, concentrarse o tomar decisiones | (13) Indecisión | (13) Indecisión (19) Dificultad de concentración |
| 9. Ideas recurrentes de muerte o suicidio | (9) Pensamientos de suicidio | (9) Pensamientos de suicidio |
| 10. Baja autoestima | (7) Autodecepción (14) Autoimagen | (7) Insatisfacción con uno mismo |
| 11. Sentimientos de desesperanza | (2) Pesimismo | (2) Pesimismo |
| Síntomas no considerados en el DSM-IV | (20) Hipocondria | |

Nota. Entre paréntesis los números de los correspondientes ítems en el BDI. * Del 1 al 9 aparecen los criterios sintomáticos del episodio depresivo mayor, mientras que los del trastorno distímico corresponden a los números 1, 3, 4, 6, 8, 10 y 11.

Como sus predecesores, el BDI-II se distingue en cuanto al contenido por evaluar sobre todo características cognitivas: aproximadamente la mitad (47,6%) de sus ítems tienen un contenido cognitivo (ítems 2, 3, 5, 6, 7, 8, 9, 13, 14 y 19), mientras que sólo el 38% evalúa aspectos somáticos (ítems 16, 18 y 20) o motivacionales-conductuales (ítems 4, 11, 12, 15 y 21), frente al 50% de la Escala para la Valoración de la Depresión de Hamilton (1960) o de la Escala Autoaplicada de la Depresión de Zung (1965), por ejemplo (véase el análisis comparativo efectuado por Vázquez, 1995, y Vázquez y Jiménez, 2000). El 14,3% restante de los ítems del BDI-II evalúa aspectos anímicos (ítems 1, 10 y 17).

Validez factorial

Se realizó un análisis factorial de ejes principales sobre las correlaciones entre los 21 ítems del BDI-II. La extracción inicial arrojó cuatro factores con valores propios mayores que 1; sin embargo, a partir del primer factor, que explicaba un 31,7% de varianza, la varianza explicada por los restantes factores era muy pequeña y tendía a disminuir de manera asintótica (6,5%, 5,7%, 4,8%, 4,3%, 4,1%, 3,9% y 3,7%, para los restantes siete factores). De hecho, un análisis visual del gráfico de sedimentación (*scree test* de Cattell) sugería la existencia de un solo factor de depresión (los valores propios de los ocho primeros factores fueron, respectivamente, 6,6, 1,4, 1,2, 1,0, 0,91, 0,87, 0,81 y 0,77). Además, en la matriz factorial de un único factor (véase la Tabla 9) se observa cómo todos los ítems del BDI-II saturaban en ese único factor con valores por encima de 0,40, excepto los ítems de Cambios en el apetito y Pérdida de interés en el sexo (con saturaciones de 0,39 y 0,38, respectivamente).

Dado que en la mayoría de los estudios anteriores con el BDI-II, tanto en estudiantes universitarios (Beck, Steer y Brown, 1996; Dozois et al., 1998; Steer y Clark, 1997; Whisman, Pérez y Ramel, 2000) como en población clínica (Beck, Steer y Brown, 1996; Steer, Ball, Ranieri y Beck, 1999), se han encontrado soluciones con dos factores que correlacionaban entre sí de forma moderada o muy alta

Tabla 9.- Resultados del análisis factorial sobre las puntuaciones del BDI-II

| Ítems | Solución unifactorial * | | Solución bifactorial † | |
|---------------------------------|-------------------------|--|------------------------|----------|
| | Solución unifactorial * | | Factor 1 | Factor 2 |
| 1. Tristeza | 0,62 | | 0,48 | 0,19 |
| 2. Pesimismo | 0,48 | | 0,41 | 0,11 |
| 3. Sentimientos de fracaso | 0,5 | | 0,61 | -0,08 |
| 4. Pérdida de placer | 0,53 | | 0,14 | 0,45 |
| 5. Sentimientos de culpa | 0,54 | | 0,59 | -0,01 |
| 6. Sentimientos de castigo | 0,5 | | 0,5 | 0,02 |
| 7. Insatisfacción con uno mismo | 0,64 | | 0,69 | -0,01 |
| 8. Auto-críticas | 0,58 | | 0,5 | 0,11 |
| 9. Pensamientos de suicidio | 0,44 | | 0,47 | -0,01 |
| 10. Llanto | 0,52 | | 0,43 | 0,12 |
| 11. Agitación | 0,47 | | 0,32 | 0,18 |
| 12. Pérdida de interés | 0,54 | | 0,08 | 0,52 |
| 13. Indecisión | 0,54 | | 0,52 | 0,04 |
| 14. Inutilidad | 0,64 | | 0,75 | -0,07 |
| 15. Pérdida de energía | 0,62 | | -0,01 | 0,72 |

Tabla 9.- (Continuación)

| | | | |
|-----------------------------------|-------------|-------|-------------|
| 16. Cambios en el patrón de sueño | 0,42 | -0,06 | 0,54 |
| 17. Irritabilidad | 0,57 | 0,32 | 0,29 |
| 18. Cambios en el apetito | 0,39 | 0,03 | 0,39 |
| 19. Dificultad de concentración | 0,57 | 0,31 | 0,3 |
| 20. Cansancio o fatiga | 0,6 | -0,09 | 0,79 |
| 21. Pérdida de interés en el sexo | 0,38 | 0,04 | 0,38 |

Nota. Las saturaciones mayores de 0,40 se presentan en negrita. * Matriz factorial tras extraer un solo factor mediante el método de ejes principales. † Matriz de configuración tras rotar, mediante la rotación oblicua *promax*, los dos factores extraídos mediante el método de ejes principales.

(como excepción véanse las soluciones trifactoriales de Aasen, 2001; Al-Musawi, 2001; Osman et al., 1998), se realizó un segundo análisis factorial de ejes principales extrayendo dos factores que fueron rotados mediante un procedimiento oblicuo (*promax*). La solución factorial resultante (véase la Tabla 9) identificó un primer factor que explicaba un 31,7% de varianza y que venía definido fundamentalmente por los ítems de Inutilidad, Insatisfacción con uno mismo, Sentimientos de fracaso, Sentimientos de culpa, Indecisión, Autocríticas, Tristeza, Llanto y Pesimismo. Dado que estos 11 síntomas son de naturaleza esencialmente cognitiva y afectiva, se puede considerar que este factor representa la dimensión Cognitivo-Afectiva que previos estudios factoriales han identificado en estudiantes universitarios (Beck, Steer y Brown, 1996; Dozois et al., 1998; Steer y Clark, 1997; Whisman et al., 2000). El segundo factor, que explicaba un 6,6% de varianza, quedaba configurado principalmente por los ítems de Cansancio-fatiga, Pérdida de energía, Cambios en el patrón de sueño, Pérdida de interés y Pérdida de placer, los cuales parecen representar un factor somático-motivacional en línea con la dimensión Somática que también ha sido identificada con estudiantes universitarios en estudios anteriores (Beck, Steer y Brown, 1996; Dozois et al., 1998; Steer y Clark, 1997; Whisman et al., 2000). En línea también con estos estudios previos, los dos factores, el cognitivo-afectivo y el somático-motivacional, se mostraban altamente correlacionados entre sí ($r = 0,71$), lo que apoya aún más la idea de que el BDI-II mide una dimensión general de depresión que está compuesta de dos dimensiones sintomáticas altamente relacionadas, una cognitivo-afectiva y otra somático-motivacional. De hecho, un análisis comparativo de los ítems que definen las soluciones bifactoriales del BDI-II en estudios con estudiantes universitarios que han utilizado los mismos procedimientos de extracción y rotación factorial (ejes principales y rotación *promax*), sugiere que los dos factores encontrados en el presente estudio se corresponden razonablemente bien con los factores cognitivo-afectivo y somático encontrados en la literatura previa (véase la Tabla 10). Para cuantificar el grado de convergencia entre la solución bifactorial del presente estudio y las encontradas en estudios anteriores y que aparecen recogidas en la Tabla 10, se calcularon dos índices: el coeficiente

Tabla 10.- Soluciones bifactoriales del BDI-II obtenidas mediante análisis factorial exploratorio con estudiantes universitarios

| Estudio | Factor 1 ^a (F1) — Cognitivo-afectivo | Factor 2 ^a (F2) — Somático-motivacional | r entre F1 y F2 | Congruencia factorial con este estudio | | | |
|----------------------------|---|--|-----------------|--|-----------------------|------|------|
| | | | | Factor 1 ^a | Factor 2 ^a | r | C |
| Beck, Steer y Brown (1996) | 14, 1, 13, 10, 7, 8, 5, 3, 6, 9, 17, 12, 11 | 15, 20, 18, 19, 16 | 0,62 | 0,85 | 0,95 | 0,8 | 0,89 |
| Dozois et al. (1998) | 3, 14, 7, 8, 2, 13, 5, 9, 1, 6 | 20, 15, 16, 17, 11, 4, 12, 19, 21, 18 | 0,6 | 0,92 | 0,96 | 0,91 | 0,96 |
| Steer y Clark (1997) | 14, 7, 9, 8, 1, 4, 10, 3, 13, 21, 2, 12, 17 | 20, 16, 15, 11, 19, 18, 12 | 0,58 | 0,7 | 0,9 | 0,8 | 0,91 |
| Este estudio | 14, 7, 3, 5, 13, 8, 6, 1, 9, 10, 2 | 20, 15, 16, 12, 4, 18, 21 | 0,71 | | | | |

Nota. * Los ítems aparecen indicados mediante su número correspondiente del BDI-II y en orden decreciente respecto a la magnitud de su saturación factorial; en negrita los ítems con saturaciones significativas ($\geq ,35$) que coinciden en las soluciones de los cuatro estudios.
r = coeficiente de correlación de Pearson. C = índice de congruencia de Tucker.

de congruencia factorial C de Tucker y el coeficiente de correlación de Pearson. El rango de valores de C varía entre -1 y 1 , indicando el valor cero falta de acuerdo. Cureton y D'Agostino (1983) y Mulaik (1972) sugieren valores de C mayores que $0,80$ ó $0,90$ para afirmar que dos factores son congruentes. Para el coeficiente de correlación, Cliff (1966) ha propuesto una correlación mínima $0,75$ para afirmar que dos factores tienen una interpretación similar. Como puede verse en la Tabla 10, los valores de C que se obtuvieron superaron en todos los casos el estándar de $0,80$, mientras que los valores de r también superaron en todos los casos el estándar de $0,75$ con la excepción de la comparación con el factor cognitivo-afectivo identificado por Steer y Clark (1997). Esta elevada congruencia con las soluciones bifactoriales previas y, en general, el que el BDI-II original muestre un estructura factorial compuesta de dos factores, cognitivo-afectivo y somático-motivacional, que aparece de forma consistente en la mayoría de los estudios, contrasta con los resultados factoriales que se obtenían con el BDI-I y el BDI-IA, en los cuales la consistencia era la excepción más que la regla (véase la revisión de estudios factoriales de Beck et al., 1988, y, respecto a los estudiantes universitarios españoles, los estudios factoriales de Ibáñez, Peñate y González, 1997; Ibáñez, González y Peñate, 1997; Vázquez y Sanz, 1991).

Validez de criterio: discriminación entre estudiantes con y sin episodio depresivo mayor

Tomando como criterio de validez el diagnóstico llevado a cabo por medio de la entrevista computerizada Quick DIS-III-R, se pueden obtener, para las distintas puntuaciones del BDI-II, diversos índices de validez relacionados con su capacidad para diagnosticar un episodio depresivo mayor. Entre los índices más utilizados para este propósito se encuentran (Kessel y Zimmerman, 1993): sensibilidad (capacidad del BDI-II para identificar correctamente estudiantes deprimidos, es decir, porcentaje de estudiantes con episodio depresivo mayor que puntúan igual o por encima de un punto de corte establecido), especificidad (capacidad del BDI-II para identificar

correctamente estudiantes no deprimidos, es decir, porcentaje de estudiantes no deprimidos que puntúan por debajo de un determinado punto de corte en el BDI-II), valor predictivo positivo (probabilidad de que un estudiante universitario identificado por el BDI-II como deprimido padezca realmente un episodio depresivo mayor, es decir, porcentaje de estudiantes que tienen un episodio depresivo mayor y que puntúan igual o por encima de un punto de corte del BDI-II), valor predictivo negativo (probabilidad de que un estudiante identificado por el BDI-II como no deprimido efectivamente no tenga un episodio depresivo mayor, es decir, porcentaje de estudiantes que puntúan por debajo de un punto de corte determinado en el BDI-II que sufren un episodio depresivo mayor), eficiencia (porcentaje de estudiantes con y sin episodio depresivo mayor correctamente clasificados por el BDI-II) y kappa (grado de acuerdo entre el BDI-II y el diagnóstico ofrecido por el Quick DIS-III-R que no es explicable por azar).

En la Tabla 11 aparece la precisión diagnóstica del BDI-II en la muestra de estudiantes entrevistados para el rango de puntuaciones totales entre 10 y 35. Para interpretar estos resultados es imprescindible tener en cuenta el propósito para el que va a ser utilizado el test. De manera que si el BDI-II va a ser utilizado como instrumento de discriminación entre estudiantes universitarios que sufren trastornos depresivos y no depresivos (p. ej., en un servicio de asistencia psicológica universitario), se debería utilizar un punto de corte que ofrezca simultáneamente valores altos de sensibilidad y especificidad, mientras que si se pretende utilizar el BDI-II como instrumento de cribado para detectar personas con una potencial depresión en la población general de estudiantes universitarios, al igual que se ha hecho recientemente en población española general empleando el BDI-IA (Lastra et al., 2000), se podría emplear un punto de corte con valores más altos de sensibilidad. Con el objetivo de analizar estas cuestiones a lo largo de todas las puntuaciones del BDI-II se empleó un análisis de la curva ROC. En dicha curva se representa gráficamente la sensibilidad en la ordenada y la tasa de falsos positivos ($1 - \text{especificidad}$) en la abscisa para cada uno de los puntos de corte del instrumento. El caso ideal estaría representado por un punto de corte que arrojarase una sensibilidad del

Tabla 11.- Puntuaciones del BDI-II y su precisión diagnóstica para detectar personas con episodio depresivo mayor en la muestra de estudiantes universitarios

| BDI-II | Sensibilidad | Especificidad | VPP | VPN | Eficiencia | Kappa |
|--------|--------------|---------------|-------|-------|------------|-------|
| 10 | 92,94 | 76,83 | 27,15 | 99,15 | 78,20 | 0,33 |
| 11 | 92,94 | 80,11 | 30,27 | 99,19 | 81,20 | 0,38 |
| 12 | 92,94 | 84,15 | 35,27 | 99,23 | 84,90 | 0,44 |
| 13 | 85,88 | 88,09 | 40,11 | 98,53 | 87,90 | 0,49 |
| 14 | 78,82 | 90,05 | 42,41 | 97,86 | 89,10 | 0,50 |
| 15 | 78,82 | 91,37 | 45,89 | 97,89 | 90,30 | 0,53 |
| 16 | 78,82 | 92,02 | 47,86 | 97,91 | 90,90 | 0,55 |
| 17 | 71,76 | 94,00 | 52,59 | 97,29 | 92,11 | 0,56 |
| 18 | 71,76 | 94,75 | 55,96 | 97,31 | 92,80 | 0,59 |

Tabla 11.- (Continuación)

| | | | | | | |
|----|-------|-------|-------|-------|-------|------|
| 19 | 57,14 | 96,72 | 61,54 | 96,09 | 93,39 | 0,56 |
| 20 | 57,14 | 96,72 | 61,54 | 96,09 | 93,39 | 0,56 |
| 21 | 50 | 97,38 | 63,64 | 95,50 | 93,39 | 0,52 |
| 22 | 50 | 98,69 | 77,78 | 95,56 | 94,59 | 0,58 |
| 23 | 50 | 98,69 | 77,78 | 95,56 | 94,59 | 0,58 |
| 24 | 35,29 | 98,69 | 71,43 | 94,26 | 93,30 | 0,44 |
| 25 | 28,24 | 99,34 | 80,00 | 93,71 | 93,30 | 0,39 |
| 26 | 21,18 | 99,34 | 75,00 | 93,14 | 92,70 | 0,30 |
| 27 | 14,12 | 99,34 | 66,67 | 92,57 | 92,10 | 0,21 |
| 28 | 14,12 | 100 | 100 | 92,61 | 92,70 | 0,23 |

Tabla 11.- (Continuación)

| | | | | | | | |
|----|-------|-----|-----|-----|-------|-------|------|
| 29 | 14,12 | 100 | 100 | 100 | 92,61 | 92,70 | 0,23 |
| 30 | 14,12 | 100 | 100 | 100 | 92,61 | 92,70 | 0,23 |
| 31 | 7,06 | 100 | 100 | 100 | 92,05 | 92,10 | 0,12 |
| 32 | 7,06 | 100 | 100 | 100 | 92,05 | 92,10 | 0,12 |
| 33 | 7,06 | 100 | 100 | 100 | 92,05 | 92,10 | 0,12 |
| 34 | 7,06 | 100 | 100 | 100 | 92,05 | 92,10 | 0,12 |
| 35 | 7,06 | 100 | 100 | 100 | 92,05 | 92,1 | 0,12 |

Nota. VPP = Valor predictivo positivo; VPN = Valor predictivo negativo.

100% y una tasa de falsos positivos de 0% (especificidad del 100%); lo cual quedaría representado en el vértice superior izquierdo de la curva. De este modo, el mejor punto de corte para el instrumento en cuestión sería aquel que estuviese situado más cerca de dicho vértice. En el caso de que el poder discriminante del instrumento fuese nulo, los valores de la curva ROC formarían una línea recta en la diagonal principal.

En este estudio, la curva ROC identificó como mejor punto de corte para el BDI-II la puntuación de 12 (véase la Figura 1). Utilizando este punto de corte, el BDI-II fue capaz de identificar correctamente al 92,9% de los individuos con episodio depresivo mayor y al 84,1% de los individuos sin depresión, lo cual supone unos índices de eficiencia diagnóstica y de concordancia (coeficiente kappa) de 84,9% y 0,44, respectivamente. El grado de acuerdo encontrado entre el BDI-II y el diagnóstico del Quick DIS-III-R (kappa = 0,44), se considera aceptable en cuanto a su significación clínica según las directrices presentadas por Cicchetti (1994) para evaluar los valores de kappa. Por otra parte, el punto de corte del BDI-II que nos ofrece el mejor valor de concordancia con el diagnóstico es el de 18 (kappa = 0,59). A su vez, los puntos de corte que nos ofrecen los valores más altos de sensibilidad son los situados en el rango de 10 a 12, mientras que los que nos ofrecen los niveles más altos de especificidad son las puntuaciones iguales o superiores a 28. En general, las características operativas del BDI-II para el diagnóstico del episodio depresivo mayor en muestras de estudiantes son aceptables. Una medida útil del rendimiento diagnóstico de un instrumento es el área bajo la curva ROC (Swets, 1979), la cual varía desde 0,5 para un rendimiento al azar a lo largo de la diagonal hasta 1 para una curva ROC con una precisión pronosticadora perfecta. El área bajo la curva ROC para el BDI-II fue de ,91, con un intervalo de confianza asintótico al 95% de confianza entre 0,80 y 1,0, lo que supone una capacidad discriminadora muy buena (Swets, 1979).

Para determinar si un conjunto específico de ítems del BDI-II podría diferenciar los estudiantes con episodio depresivo mayor de los no deprimidos, se realizó un análisis discriminante por pasos utilizando como criterio de clasificación el diagnóstico puntual de episodio depresivo mayor realizado con el Quick DIS-III-R. Este

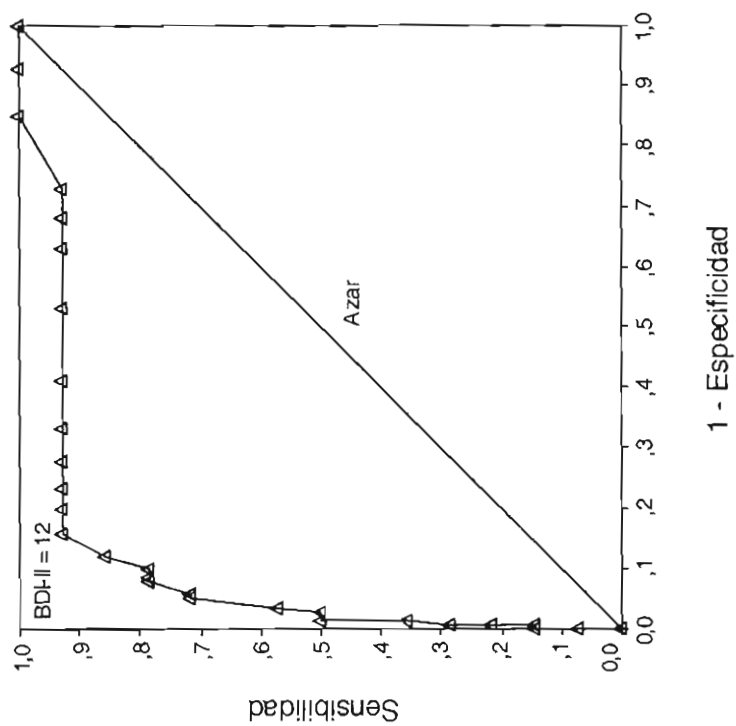


Figura 1.- Curva ROC para las puntuaciones del BDI-II

análisis reveló que tan sólo siete ítems podrían discriminar aceptablemente entre los estudiantes deprimidos y no deprimidos: Tristeza, Sentimiento de castigo, Pensamientos de suicidio, Agitación, Pérdida de energía, Cambios en el apetito y Pérdida de interés en el sexo (todos ellos con correlaciones mayores de 0,23 con la función discriminante, entre los que destaca la correlación de 0,62 del ítem de Tristeza). La función discriminante obtenida podía clasificar correctamente el 96,4 % de los sujetos como deprimidos o no deprimidos con un lambda de Wilks de 0,49 ($c^2(7, N = 165) = 114,1, p < 0,0001$). La función resultante era fundamentalmente acertada para clasificar a los estudiantes no deprimidos como tales (98% de estudiantes no deprimidos bien clasificados frente a un 78,6% de estudiantes con depresión bien clasificados).

CONCLUSIONES

1. En este trabajo, se ha desarrollado una *versión española del BDI-II* siguiendo un cuidadoso proceso de traducción y estudio de su aplicabilidad y comprensibilidad, y se han analizado sus propiedades psicométricas en una muestra de estudiantes universitarios españoles.

2. La *distribución de las puntuaciones* del BDI-II en esta muestra es similar a la encontrada en estudios anteriores con estudiantes universitarios de otros países, con una media que, tal y como también ocurre en la literatura previa, es superior en 2-3 puntos a la media que se suele obtener con su inmediato predecesor, el BDI-IA. Este hecho sustenta la decisión de incrementar en el BDI-II, y con respecto al BDI-IA, las puntuaciones criterio para definir las categorías de gravedad de la sintomatología depresiva.

3. Se replican transculturalmente los porcentajes de estudiantes universitarios que se corresponden a las distintas *categorías de gravedad* de la depresión medidos por el BDI-II.

4. Entre los estudiantes universitarios españoles no hay *diferencias entre sexos* en cuanto a la puntuación total en el BDI-II, dato que replica la ausencia de diferencias que también aparecía con la versión española del BDI-IA (Sanz y Vázquez, 1998). Este dato

puede explicarse en función de la mayor homogeneidad de este tipo de población en cuanto a roles, actividades, nivel de estudios, profesiones, etc. De hecho, variables como el nivel de estudios y la profesión afectan a las puntuaciones del BDI-IA cuando son analizadas en la población general de adultos españoles (Vázquez y Sanz, 1997). Aunque las diferencias en tasas de depresión entre sexos son muy consistentes en diferentes estudios y en diferentes culturas, la existencia de tasas más altas en las mujeres sólo es evidente justamente *después* de la adolescencia; de hecho, en la mayor parte de los trabajos con estudiantes universitarios aún no se detectan diferencias significativas en niveles de depresión entre hombres y mujeres (véase las excelentes revisiones de Nolen-Hoeksema, 1987, 1993). En concreto, la mayoría de las investigaciones realizadas con el BDI-II confirman esta ausencia de diferencias entre sexos en muestras de estudiantes universitarios (Aasen, 2001; Al-Musawi, 2001; Dozois et al., 1998; O'Hara et al., 1998; Schulenberg y Yutrzenka, 2001; Steer y Clark, 1997). De modo que esta ausencia de diferencias entre hombres y mujeres jóvenes universitarios es de gran importancia ya que parece apoyar el gran peso que tienen los factores socioculturales y de roles en la existencia de mayores tasas de depresión en las mujeres (Nolen-Hoeksema, 1994).

5. La fiabilidad en términos de *consistencia interna* de la versión española del BDI-II es buena y se asemeja a los niveles encontrados en otros ámbitos culturales.

6. Los resultados de los *análisis factoriales* realizados con la versión española del BDI-II indican que en muestras de estudiantes universitarios este instrumento parece medir una dimensión de depresión general que está compuesto de dos dimensiones sintomáticas altamente relacionadas, una cognitivo-afectiva y otra somático-motivacional, que replica los resultados encontrados en la literatura previa.

7. El BDI-II cubre aceptablemente el espectro de síntomas incluidos como criterios sintomatológicos de los *trastornos depresivos en el DSM-IV* (trastorno depresivo mayor y trastorno distímico), lo que hace lógica la elección de este instrumento para evaluar cuantitativamente la sintomatología depresiva. Sin embar-

go, ya que su contenido está muy centrado en los aspectos cognitivos de la depresión, parece recomendable emplear otros instrumentos simultáneamente para recabar información complementaria del individuo, especialmente en relación a los síntomas depresivos somáticos. Así, por ejemplo, se podría emplear la Escala para la Valoración de la Depresión de Hamilton o, si se prefiere un instrumento autoaplicado, no heteroaplicado como el anterior, se podría emplear la Escala Autoaplicada para la Depresión de Zung (ambas adaptadas al español por el grupo de investigación de Conde; véase Conde y Franch, 1984), si bien la eficacia de estos instrumentos es muy semejante, al menos en cuanto a su capacidad para detectar posibles “casos” de depresión (Williams, Noel, Cordis et al., 2002). El lector también puede consultar otras opciones de medida en Vázquez (1995) y Vázquez y Jiménez (2000).

8. Obviamente, el hecho de que el BDI-II cubra el espectro de síntomas que definen un trastorno depresivo mayor o un trastorno distímico no justifica el uso exclusivo del BDI como instrumento diagnóstico. Hacerlo conllevaría confundir los niveles de análisis: síntoma-síndrome-trastorno. El BDI-II sirve para *identificar síntomas depresivos y cuantificar su intensidad*, lo que obviamente es muy importante para evaluar, por ejemplo, el estado general del paciente o la efectividad de un proceso terapéutico. Sin embargo, el diagnóstico de un trastorno depresivo mayor o un trastorno distímico se efectúa teniendo en cuenta no sólo el tipo y número de síntomas presentes, sino también ciertos criterios de duración (p. ej., los síntomas aparecen la mayor parte del tiempo de la mayoría de los días durante al menos dos años), gravedad (p. ej., al menos aparecen cinco síntomas y éstos provocan un malestar clínicamente significativo), curso (p. ej., los síntomas representan un cambio respecto a la actividad previa), incapacidad (p. ej., los síntomas provocan deterioro social, escolar, o laboral), y ausencia de ciertas causas posibles o de ciertos diagnósticos concurrentes (p. ej., no se considera el efecto fisiológico directo de una enfermedad médica o de la ingestión de medicamentos o drogas, o no se explica mejor por la muerte reciente de un ser querido o por la presencia de una esquizofrenia).

9. No obstante, los análisis realizados para evaluar la capacidad del BDI-II para *diagnosticar un episodio depresivo* mayor sugieren que este instrumento puede ser de gran ayuda en este tipo de labores, al menos como instrumento de cribado (*screening*) para detectar casos de depresión en muestras de estudiantes universitarios, casos que deberían posteriormente ser confirmados mediante algún tipo de entrevista diagnóstica.

10. En términos de validez de contenido, fiabilidad de consistencia interna y validez factorial, la versión española del BDI-II parece *mejor que su predecesor, el BDI-IA*, como instrumento para evaluar la sintomatología depresiva en estudiantes universitarios, ya que (a) sus ítems cubren mejor los criterios sintomáticos de los trastornos depresivos del DSM-IV, (b) exhibe un coeficiente alfa más alto, y (c) presenta una estructura factorial que refleja en mejor medida una dimensión general de depresión, distingue de forma más clara dos factores interrelacionados de sintomatología cognitivo-afectiva y somático-motivacional, y es más consistente con la literatura previa.

11. En suma, la versión española del BDI-II parece tener unas *propiedades psicométricas aceptables* como instrumento de evaluación de la sintomatología depresiva en estudiantes universitarios españoles que recomiendan su uso en este tipo de población en ámbitos clínicos tanto de prevención (p. ej., para identificar estudiantes en riesgo que pudieran formar parte de programas preventivos) como de tratamiento (p. ej., valorar el estado de los estudiantes con trastornos y su progreso terapéutico), así como en ámbitos de investigación (p. ej., para identificar estudiantes deprimidos y no deprimidos en los estudios de validez de los modelos de depresión). Por otro lado, y en función de la experiencia acumulada con su predecesor, la versión española del BDI-IA (Sanz y Vázquez, 1998; Vázquez y Sanz, 1997, 1999), es posible aventurar que estas buenas propiedades psicométricas con estudiantes universitarios auguran que el BDI-II también mostrará un adecuado rendimiento como instrumento de evaluación de la sintomatología depresiva en la población española general y clínica, cuestiones éstas que abordaremos en próximos trabajos (Sanz, Perdígón y Vázquez, 2003; Sanz, García-Vera y Vázquez, 2003).

Notas de los autores

Este artículo se basa en parte en el trabajo de investigación de doctorado conducente al Diploma de Estudios Avanzados llevado a cabo por la segunda autora bajo la dirección del primero. Los autores quisieran agradecer al Dr. Jeff Ring su colaboración en la realización de la traducción inversa. La correspondencia en relación a este artículo puede dirigirse a Jesús Sanz, Departamento de Personalidad, Evaluación y Psicología Clínica, Universidad Complutense de Madrid, Campus de Somosaguas, 28223 Madrid, o, mediante correo electrónico, a jsanz@psi.ucm.es.

BIBLIOGRAFÍA

- Aasen, H.** (2001). *An empirical investigation of depression symptoms: norms, psychometric characteristics and factor structure of the Beck Depression Inventory-II*. Facultad de Psicología, Universidad de Bergen. Documento consultado en Internet en: <http://www.ub.uib.no/eipub/2001/h/308003/Hovedoppgave.pdf>.
- Al Musawi, N. M. M.** (2001). Psychometric properties of the Beck Depression Inventory-II with university students in Bahrain. *Journal of Personality Assessment*, 77(3), 568-579.
- Alsawalmeh, Y. M. y Feldt, L. S.** (1992). Test of the hypothesis that the intraclass reliability coefficient is the same for two measurement procedures. *Applied Psychological Measurement*, 16, 195-205.
- American Psychiatric Association.** (1980). *Diagnostic and statistical manual of mental disorders* (3^o ed.). Washington, DC: APA (Trad. esp. en Barcelona: Masson, 1980).
- American Psychiatric Association.** (1987). *Diagnostic and statistical manual of mental disorders* (3^o ed., revisada). Washington, DC: APA (Trad. esp. en Barcelona: Masson, 1989).
- American Psychiatric Association.** (1994). *Diagnostic and statistical manual of mental disorders* (4^o ed.). Washington, DC: APA (Trad. esp. en Barcelona: Masson, 1996).
- Arnau, R. C., Meagher, M. W., Norris, M. P. y Bramson, R.** (2001). Psychometric evaluation of the Beck Depression Inventory-II with primary care medical patients. *Health Psychology*, 20(2), 112-119.

- Beck, A.T., Rush, A.J., Shaw, B.F. y Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press (Trad. esp. en Bilbao: Desclée de Brouwer, 1983).
- Beck, A.T. y Steer, R.A. (1993). *Beck Depression Inventory. Manual*. San Antonio, TX: The Psychological Corporation.
- Beck, A. T., Steer, R. A., Ball, R. y Ranieri, W. F. (1996). Comparison of Beck Depression Inventories-IA and -II in psychiatric outpatients. *Journal of Personality Assessment*, 67(3), 588-597.
- Beck, A. T., Steer, R. A. y Brown, G. K. (1996). *BDI-II. Beck Depression Inventory-Second Edition. Manual*. San Antonio, TX: The Psychological Corporation.
- Beck, A.T., Steer, R.A. y Garbin, M.C. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8, 77-100.
- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J. y Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- Benthencourt, J.M., Peñate, W., Fumero, A. y González, M. (1997). Trastornos de ansiedad y estrés en los estudiantes de la Universidad de La Laguna. *Ansiedad y Estrés*, 3(1), 25-36.
- Benthencourt, J.M., Peñate, W., González, M. y Fumero, A. (1996). El programa de entrevista diagnóstica estructurada (D.I.S.): presentación de la versión española y primeros datos de validez a través de la fiabilidad intercalificadores. *Psicologemas*, 10, 47-154.
- Benthencourt, J.M., Peñate, W., González, M., Fumero, A., Fernández-Valdés, A. y Díaz, F. (1998). Prevalencia de trastornos psicopatológicos en los estudiantes de la Universidad de La Laguna. *Psicologemas*, 12(23), 4-29.
- Bucholz, K. K., Marion, S. L., Shayka, J. J., Marcus, S. C. y Robins, L.N. (1996). A short computer interview for obtaining psychiatric diagnoses. *Psychiatric Services*, 47(3), 293-297.
- Buckley, T. C., Parker, J. D. y Heggie, J. (2001). A psychometric evaluation of the BDI-II in treatment-seeking substance abusers. *Journal of Substance Abuse Treatment*, 20(3), 197-204.
- Burt, T. y IsHak, W.W. (2002). Outcome measures in mood disorders. En W.W. IsHak, T. Burt y L.I. Sederer (Eds.), *Outcome measurement in Psychiatry* (pp. 155-190). Washington, DC: American Psychiatric Publishing.
- Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290.

- Cliff, J. (1966). Orthogonal rotation to congruence. *Psychometrika*, 31, 33-42.
- Coelho, R., Martins, A. y Barros, H. (2002). Clinical profiles relating gender and depressive symptoms among adolescents ascertained by the Beck Depression Inventory II. *European Psychiatry*, 17(4), 222-226.
- Conde, V., Esteban, T. y Useros, E. (1976). Revisión crítica de la adaptación castellana del Cuestionario de Beck. *Revista de Psicología General y Aplicada*, 31, 469-497.
- Conde, V. y Franch, J.I. (1984). *Escalas de evaluación comportamental para la cuantificación de la sintomatología psicopatológica en los trastornos angustiosos y depresivos*. Madrid: Upjohn Farmaquímica.
- Consejo de Coordinación Universitaria (2002). Avance estadístico. Curso 2001-2002. Documento consultado en Internet: <http://www.mec.es/consejou/estadis/avan0102/index.html>.
- Cureton, E.E. y D'Agostino, R.B. (1983). *Factor analysis: An applied approach*. London: Lawrence Erlbaum Associates.
- Dozois, D. J. A., Dobson, K. S. y Ahnberg, J. L. (1998). A psychometric evaluation of the Beck Depression Inventory-II. *Psychological Assessment*, 10(2), 83-89.
- Fombonne, E. (1994). Increased rates of depression: Update of epidemiological findings and analytical problems. *Acta Psychiatrica Scandinavica*, 90(3), 145-156.
- Haaga, D. A. y Solomon, A. (1993). Impact of Kendall, Hollon, Beck, Hammen, and Ingram (1987) on treatment of the continuity issue in «depression» research. *Cognitive Therapy and Research*, 17(4), 313-324.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, 23, 56-62.
- Ibáñez, I., González, M. y Peñate, W. (1997). Comparación de modelos factoriales alternativos para la versión española del Inventario de Depresión de Beck. *Análisis y Modificación de Conducta*, 23, 283-303.
- Ibáñez, I., Peñate, W. y González, M. (1997). La estructura factorial del Inventario de Depresión de Beck. *Psicología Conductual*, 5(1), 71-91.
- Jefferson, A. L., Powers, D. V. y Pope, M. (2001). Beck Depression Inventory-II (BDI-II) and the Geriatric Depression Scale (GDS) in older women. *Clinical Gerontologist*, 22(3-4), 3-12.
- Kendall, P. C., Hollon, S. D., Beck, A. T., Hammen, C. L. e Ingram, R.E. (1987). Issues and recommendations regarding use of the Beck Depression Inventory. *Cognitive Therapy and Research*, 11(3), 289-299.
- Kendall, P.C. y Sheldrick, R.C. (2000). Normative data for normative comparisons. *Journal of Consulting and Clinical Psychology*, 68(5), 767-773.

- Kessel, J.B. y Zimmerman, M. (1993). Reporting errors in studies of the diagnostic performance of self administered questionnaires: extent of the problem, recommendations for standardized presentation of results, and implications for the peer review process. *Psychological Assessment*, 5, 395-399.
- Lasa, L., Ayuso-Mateos, J.L., Vázquez-Barquero, J.L. Díez-Manrique, F.J., y Dowrick, C.F. (2000). The use of the Beck Depression Inventory to screen for depression in the general population: a preliminary analysis. *Journal of Affective Disorders*, 57, 261-265.
- López-García, G., Cortegano, M.C., Martín, C., Murcia, L., López-Mora, I., García Montalvo, C. y López Soler, C. (1996). Estudio epidemiológico en una muestra clínica de mujeres a través del DIS. *Revista de Psicopatología y Psicología Clínica*, 1(3), 255-268.
- Marcus, S., Robins, L. N. y Bucholz, K. (1991). *Quick Diagnostic Interview Schedule III-R. Version 1.0*. St. Louis, MO: Washington University School of Medicine.
- Mulaik, S.A. (1972). *The foundations of factor analysis*. Nueva York: McGraw-Hill.
- Muñiz, J. y Fernández, J.R. (2000). La utilización de los tests en España. *Papeles del Psicólogo*, 76, 41-49.
- Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin*, 101, 259-282.
- Nolen-Hoeksema, S. (1994). The emergence of gender differences in depression during adolescence. *Psychological Bulletin*, 115, 424-443.
- O'Hara, M. M., Sprinkle, S. D. y Ricci, N. A. (1998). Beck Depression Inventory—II: College population study. *Psychological Reports*, 82(3, Pt 2), 1395-1401.
- Osman, A., Downs, W. R., Barrios, F. X., Kopper, B. A., Gutierrez, P. M. y Chiros, C. E. (1997). Factor structure and psychometric characteristics of the Beck Depression Inventory-II. *Journal of Psychopathology and Behavioral Assessment*, 19(4), 359-376.
- Pablo Alcázar, F. de, Albacete Belmonte, A., Meroño Méndez, A., Artiz Martínez, M, Berbell Torrecilla, F., García Peñalver, A., Giménez Giménez, L. , Martínez de Salazar Arboles, A., Mateu Sabater, A., Plaza Tobarra, R., Sepulcre Canto, R. y Sánchez Hurtado, N. (1992). Validez de la versión española del Diagnostic Interview Schedule. Tercera Edición Revisada (DIS-III-R). *Actas Luso-Españolas de Neurología, Psiquiatría y Ciencias Afines*, 20(6), 257-262.
- Piotrowsky, C. (1996). Use of the Beck Depression Inventory in clinical practice. *Psychological Reports*, 79, 873-874.

- Robins, L. y Helzer, J. E.** (1994). The half-life of a structured interview: The NIMH Diagnostic Interview Schedule (DIS). *International Journal of Methods in Psychiatric Research*, 4(2), 95-102.
- Robins, L., Helzer, J., Cottler, L. y Goldring, E.** (1989). *NIMH Diagnostic Interview Schedule, Version III Revised (DIS-III-R)*. St Louis, MO: Washington University.
- Rosenthal, R.** (1983). Assessing the statistical and social importance of the effects of psychotherapy. *Journal of Consulting and Clinical Psychology*, 51, 4-13.
- Sanz, J., García-Vera, M.P. y Vázquez, C.** (2003). *Adaptación española del Inventario para la Depresión de Beck—II (BDI-II): 3. Propiedades psicométricas en pacientes con trastornos psicológicos*. Manuscrito en preparación.
- Sanz, J., Perdigón, L.A. y Vázquez, C.** (2003). *Adaptación española del Inventario para la Depresión de Beck—II (BDI-II): 2. Propiedades psicométricas en población general*. Manuscrito en preparación.
- Sanz, J. y Vázquez, C.** (1998). Fiabilidad, validez y datos normativos del Inventario para la Depresión de Beck. *Psicothema*, 10(2), 303-318.
- Seligman, M.** (1990). Why is there so much depression today? The waxing of the individual and the waning of the commons. En R. Ingram (Ed.), *Contemporary psychological approaches to depression: Theory, research, and treatment* (pp. 1-7). Nueva York: Plenum Press.
- Schulenberg, S. E. y Yutzenka, B. A.** (2001). Equivalence of computerized and conventional versions of the Beck Depression Inventory-II (BDI-II). *Current Psychology: Developmental, Learning, Personality, Social*, 20(3), 216-230.
- Steer, R. A., Ball, R., Ranieri, W. F. y Beck, A. T.** (1999). Dimensions of the Beck Depression Inventory-II in clinically depressed outpatients. *Journal of Clinical Psychology*, 55(1), 117-128.
- Steer, R. A. y Clark, D. A.** (1997). Psychometric characteristics of the Beck Depression Inventory-II with college students. *Measurement and Evaluation in Counseling and Development*, 30(3), 128-136.
- Steer, R. A., Clark, D. A., Beck, A. T. y Ranieri, W. F.** (1999). Common and specific dimensions of self-reported anxiety and depression: The BDI-II versus the BDI-IA. *Behaviour Research and Therapy*, 37(2), 183-190.
- Swets, J.A.** (1979). ROC analysis applied to the evaluation of medical imaging techniques. *Investigative Radiology*, 14, 110-121.
- Tennen, H., Eberhardt, T.L. y Affleck, G.** (1999). Depression research methodologies at the social-clinical interface: Still hazy after all these years. *Journal of Social and Clinical Psychology*, 18(2), 121-159.

- Tennen, H., Hall, J.A. y Affleck, G. (1995). Depression research methodologies in the Journal of Personality and Social Psychology: A review and critique. *Journal of Personality and Social Psychology*, 68(5), 870-884.
- Vázquez, C. (1995). Evaluación de los trastornos depresivos y bipolares. En A. Roa (Ed.), *Evaluación en psicología clínica y de la salud* (pp. 163-209). Madrid: CEPE.
- Vázquez, C. y Jiménez, F. (2000). Depresión y manía. En A. Bulbena, G. Berrios y Fernández de Larrinoa, P. (Eds.), *Medición clínica en Psiquiatría y Psicología* (pp. 255-308). Barcelona: Masson.
- Vázquez, C. y Sanz, J. (1991). *Fiabilidad y validez factorial de la versión española del Inventario de Depresión de Beck*. Comunicación presentada en III Congreso de Evaluación Psicológica, Barcelona, 25-28 de septiembre.
- Vázquez, C. y Sanz, J. (1997). Fiabilidad y valores normativos de la versión española del Inventario para la Depresión de Beck de 1978. *Clínica y Salud*, 8, 403-422.
- Vázquez, C. y Sanz, J. (1999). Fiabilidad y validez de la versión española del Inventario para la Depresión de Beck de 1978 en pacientes con trastornos psicológicos. *Clínica y Salud*, 10(1), 59-81.
- Vázquez, C. y Siles, S. (1999). Efectos de la psicopatología materna y paterna en los hijos. En J. Buendía (Ed.), *Psicología clínica. Perspectivas actuales*. Madrid: Pirámide.
- Vredenburg, K., Flett, G.L. y Krames, L. (1993). Analogue versus clinical depression: A critical reappraisal. *Psychological Bulletin*, 113, 327-344.
- Whisman, M. A., Perez, J. E. y Ramel, W. (2000). Factor structure of the Beck Depression Inventory—Second Edition (BDI-II) in a student sample. *Journal of Clinical Psychology*, 56(4), 545-551.
- Williams, J.B., Noël, P.H., Cordes, J.A., Ramirez, G. y Pignone, M. (2002). Is this patient clinically depressed? *JAMA*, 287, 1160-1170
- Zung, W.W.K. (1965). A self-rating depression scale. *Archives of General Psychiatry*, 12, 63-70.