



Consejo General de  
Colegios Oficiales  
de Psicólogos

Miembro de la Federación Europea  
de Asociaciones de Psicólogos

Conde Peñalver, 45. 5º Izq.  
28006 Madrid - España  
Tel.: +34 91 444 90 20  
Fax: +34 91 309 56 15  
secop@cop.es  
www.cop.es

# EVALUACIÓN DEL INVENTARIO

# **BDI-II**

## RESUMEN DE LA VALORACIÓN DEL TEST

### Descripción general

Características	Descripción
<b>Nombre del test</b>	Inventario de Depresión de Beck-II (BDI-II).
<b>Autor</b>	Aaron T. Beck, Robert A. Steer y Gregory K. Brown.
<b>Autor de la adaptación española</b>	Jesús Sanz y Carmelo Vázquez, en colaboración con el Departamento de I+D de Pearson Clinical and Talent Assessment España: Frédérique Vallar, Elena de la Guía y Ana Hernández.
<b>Editor de la adaptación española</b>	Pearson Educación, S.A.
<b>Fecha de la última revisión del test en su adaptación española</b>	2011
<b>Constructo evaluado</b>	Depresión
<b>Área de aplicación</b>	Psicología clínica, forense y Neuropsicología
<b>Soporte</b>	Papel y lápiz y administración oral

### Valoración general

Características	Valoración	Puntuación
<b>Calidad de los Materiales y documentación</b>	Excelente	5
<b>Fundamentación teórica</b>	Excelente	5
<b>Adaptación Española</b>	Buena	4
<b>Análisis de los ítems</b>	Buena	4
<b>Validez de contenido</b>	Buena	4
<b>Validez de constructo</b>	Excelente	5
<b>Análisis del sesgo</b>	No se aporta	-
<b>Validez predictiva</b>	Excelente	5
<b>Fiabilidad: equivalencia</b>	No se aporta	-
<b>Fiabilidad: consistencia interna</b>	Excelente	5
<b>Fiabilidad: estabilidad</b>	No se aporta	-
<b>Baremos</b>	No se aporta	-

## Comentarios generales

El Inventario de Depresión de Beck-II (BDI-II) es un autoinforme de lápiz y papel compuesto por 21 ítems de tipo Likert. El inventario inicialmente propuesto por Beck y sus versiones posteriores han sido los instrumentos más utilizados para detectar y evaluar la gravedad de la depresión. De hecho, es el quinto test más utilizado por los psicólogos españoles (Muñiz y Fernández-Hermida, 2010). Sus ítems no se derivan de ninguna teoría concreta acerca del constructo medido, sino que describen los síntomas clínicos más frecuentes de los pacientes psiquiátricos con depresión. El BDI-II ha experimentado algunas modificaciones respecto a las versiones anteriores para representar mejor los criterios para el diagnóstico de los trastornos depresivos recogidos en el DSM-IV (*Manual diagnóstico y estadístico de los trastornos mentales*, cuarta edición, American Psychiatric Association, 1994) y CIE-10 (*Clasificación estadística internacional de enfermedades y problemas relacionados con la salud*, Organización Mundial de la Salud, 1993). La prueba ha de ser destinada preferentemente para un uso clínico, como un medio para evaluar la gravedad de la depresión en pacientes adultos y adolescentes con un diagnóstico psiquiátrico y con 13 años o más de edad.

Entre sus **puntos fuertes** destacaríamos los siguientes:

1. Se ha realizado un cuidado proceso de adaptación de la prueba. La traducción se ha llevado a cabo siguiendo en lo sustancial las recomendaciones para la traducción y adaptación de la International Test Commission.
2. El impreso que contiene las instrucciones y los ítems está muy bien editado y es de fácil comprensión y administración. El completo manual del test incluye los fundamentos de la prueba, los estudios psicométricos y una amplia bibliografía sobre la misma.
3. Los estudios psicométricos de la versión española del BDI-II se han llevado a cabo con varias muestras de tamaño moderado o alto, extraídas en su mayor parte de la Comunidad de Madrid: pacientes adultos con diversos trastornos psicológicos que seguían tratamientos psicológicos ambulatorios, adultos de la población general y estudiantes universitarios de diversas facultades. Los ítems presentan altas saturaciones en un factor o dimensión básica de depresión. Las puntuaciones presentan una alta fiabilidad, estimada mediante diseños de consistencia interna. La validez convergente con otras medidas de depresión es adecuada.
4. Las puntuaciones muestran una alta discriminación entre personas con un trastorno depresivo mayor y sin depresión, diagnosticadas mediante entrevistas estructuradas, tanto en muestras de pacientes como en muestras de estudiantes universitarios. En el Manual se incluyen tablas con índices estadísticos valiosos para seleccionar los puntos de corte más apropiados para diversos propósitos.

Entre los **puntos a mejorar**, destacaríamos los siguientes:

1. No se aportan en el Manual algunos datos que pueden ser de interés, como indicadores de la estabilidad de las puntuaciones (fiabilidad test-retest) de la versión española, estudios sobre el funcionamiento diferencial de los ítems, y baremos que puedan facilitar más la interpretación de las puntuaciones.
2. Se podría considerar incluir en el Manual información sobre las implicaciones teóricas y prácticas acerca de las relaciones del BDI-II con el sexo y la edad. En este sentido, podría ser de utilidad disponer de puntuaciones de corte específicas para la población española, por sexo y edad. Asimismo, vendrían bien estudios que evalúen los efectos de los sesgos de memoria y respuesta en el BDI-II (tal y como se ha hecho para versiones anteriores como, por ejemplo, en el BDI-IA).

En resumen, las buenas propiedades psicométricas de la prueba avalan su uso en la población española, tanto en ámbitos clínicos como de investigación, para la detección de la presencia y de la gravedad de la depresión en adultos y adolescentes de 13 años o más.

## ANÁLISIS DETALLADO DE LA PRUEBA

### 1. DESCRIPCIÓN GENERAL DEL TEST

#### 1.1. Nombre del test

Inventario de Depresión de Beck-II (BDI-II).

#### 1.2. Nombre del test en su versión original

Beck Depression Inventory-Second Edition (BDI-II).

#### 1.3. Autor

Aaron T. Beck, Robert A. Steer y Gregory K. Brown

#### 1.4. Autor de la adaptación española

Jesús Sanz y Carmelo Vázquez, en colaboración con el Departamento de I+D de Pearson Clinical and Talent Assessment España: Frédérique Vallar, Elena de la Guía y Ana Hernández.

#### 1.5. Editor del test en su versión original

NCS Pearson, Inc. (U.S.A.)

#### 1.6. Editor de la adaptación española

Pearson Educación, S.A

#### 1.7. Fecha de publicación del test original

1996.

#### 1.8. Fecha de la publicación del test en su adaptación española

2011

#### 1.9. Fecha de la última revisión del test en su adaptación española

2011

#### 1.10. Área general de la variable que pretende medir el test

- Escalas Clínicas

#### 1.11. Breve descripción de la variable que pretende medir el test

El BDI-II es un autoinforme que proporciona una medida de la presencia y de la gravedad de la depresión en adultos y adolescentes de 13 años o más. Se compone de 21 ítems indicativos de síntomas tales como tristeza, llanto, pérdida de placer, sentimientos de fracaso y de culpa, pensamientos o deseos de suicidio, pesimismo, etc. Estos síntomas se corresponden con los criterios para el diagnóstico de los trastornos depresivos recogidos en el DSM-IV (*Manual diagnóstico y estadístico de los trastornos mentales*, cuarta edición, American Psychiatric Association, 1994) y CIE-10 (*Clasificación Estadística Internacional de Enfermedades y Problemas relacionados con la Salud*, Organización Mundial de la Salud, 1993). En concreto, se incluyen todos los criterios propuestos en las dos referencias citadas para el diagnóstico de un episodio depresivo mayor y casi todos los propuestos para el diagnóstico de distimia (en concreto, todos los del DSM-IV y el 75% del CIE-10).

El BDI-II es fácil de utilizar. Se puede aplicar de forma individual o colectiva, con formato de papel y lápiz o de forma oral. En general, requiere entre 5 y 10 minutos para ser completado; aunque los pacientes con depresión grave o trastornos obsesivos a menudo pueden tardar más tiempo en cumplimentarlo. Las instrucciones para el BDI-II solicitan a las personas evaluadas que elijan las afirmaciones más características que cubren el marco temporal de las últimas dos semanas, incluido el día de hoy, para ser consistente con los criterios del DSM-IV para la depresión mayor. Cada ítem se responde en una escala de 4 puntos, de 0 a 3, excepto los ítems 16 (cambios en el patrón de sueño) y 18 (cambios en el apetito) que contienen 7 categorías. Si una persona ha elegido varias categorías de respuesta en un ítem, se toma la categoría a la que corresponde la puntuación más alta. Las puntuaciones mínima y máxima en el test son 0 y 63. Se han establecido puntos de corte que permiten clasificar a los evaluados en uno de los siguientes cuatro grupos: 0-13, mínima depresión; 14-19, depresión leve; 20-28, depresión moderada; y 29-63, depresión grave.

#### 1.12. Áreas de aplicación

- Psicología clínica
- Neuropsicología
- Psicología forense

#### 1.13. Formato de los ítems

- Tipo Likert

En las instrucciones de la versión original del BDI-II se solicita a la persona evaluada que *“si dentro del mismo grupo hay más de una afirmación que considere igualmente aplicable a su caso, elija el número más alto”* y, en consecuencia, se le pide también que se asegure *“de no haber elegido más de una respuesta para cada grupo, particularmente en los ítems 16 (cambios en el patrón de sueño) y 18 (cambios en el apetito)”*. No obstante, en el estudio piloto de la adaptación española, dos personas (10%) informaron no sentirse cómodas eligiendo una sola alternativa, ya que en algunos ítems había varias alternativas que las describían. Por tanto, y en consonancia con las instrucciones para la aplicación oral del cuestionario, para los estudios de investigación iniciales y análisis psicométricos de la adaptación española, se mantuvieron esas instrucciones, en las que se le pide a la persona evaluada que *“si dentro de un mismo grupo, hay más de una afirmación que considere aplicable a su caso, márkela también”*, dejando que la elección de la alternativa con el número más alto descansara en el evaluador, no en la persona evaluada. Sin embargo, para poder comparar los resultados que se obtengan en la adaptación española BDI-II con los obtenidos internacionalmente, se han mantenido las instrucciones de la versión original en inglés del BDI-II (en las que se sugiere elegir solo una categoría).

#### **1.14. Número de ítems**

21 ítems

#### **1.15. Soporte**

- Administración oral
- Papel y lápiz

#### **1.16. Cualificación requerida para el uso del test de acuerdo con la documentación aportada**

- Nivel C

Aunque el BDI-II puede ser fácilmente aplicado y corregido por técnicos sanitarios, las puntuaciones deben ser interpretadas únicamente por profesionales con la adecuada preparación y experiencia clínica.

#### **1.17. Descripción de las poblaciones a las que el test es aplicable**

- El BDI-II ha sido construido preferentemente para un uso clínico como un medio para evaluar la gravedad de la depresión en pacientes adultos y adolescentes con un diagnóstico psiquiátrico y con 13 años o más de edad.

#### **1.18. Existencia o no de diferentes formas del test y sus características**

- No existen formas diferentes del test BDI-II.

#### **1.19. Procedimiento de corrección**

- Manual

#### **1.20. Procedimiento de obtención de las puntuaciones directas**

- El formato de los ítems es de tipo Likert con cuatro categorías de respuesta ordenadas que se codifican de 0 hasta 3. La puntuación es la suma de las respuestas a los 21 ítems. Si una persona ha elegido varias opciones en un ítem, se toma la opción con la puntuación más alta. También se puntúan de 0 a 3 los ítems 16 y 18, que tienen 7 categorías. En estos ítems el Manual recomienda registrar la categoría seleccionada, por su valor diagnóstico.
- El rango de las puntuaciones va desde 0 a 63 puntos. Cuanto más alta sea la puntuación, mayor será la severidad de los síntomas depresivos. Se establecen cuatro grupos en función de la puntuación total: 0-13, mínima depresión; 14-19, depresión leve; 20-28, depresión moderada; y 29-63, depresión grave.

#### **1.21. Transformación de las puntuaciones**

- Característica no aplicable para este instrumento

#### **1.22. Escalas utilizadas**

#### **1.23. Posibilidad de obtener informes automatizados**

- No

**1.24. Oferta por parte del editor de un servicio para la corrección y/o elaboración informes**

- No

**1.25. Tiempo estimado para la aplicación del test**

- En aplicación individual: 5-10 minutos
- En aplicación colectiva: 5-10 minutos

**1.26. Documentación aportada por el editor**

- Manual

**1.27. Precio de un juego completo de la prueba**

- 72.80 euros.

**1.28. Precio y número de ejemplares del paquete de cuadernillos**

- No procede.

**1.29. Precio y número de ejemplares del paquete de hojas de respuesta**

- El precio de 50 HOJAS DE RESPUESTA es de 46,80 € (+ IVA). El precio del Manual de la prueba es de 41,60 € (+ IVA).

**1.30. Precio de la corrección y/o elaboración de informes por parte del editor**

No procede

**1.31. Bibliografía básica acerca del test aportada en la documentación**

- En el Manual se incluyen 111 referencias bibliográficas. Este apartado es exhaustivo y está actualizado.



## 2. VALORACIÓN DE LAS CARACTERÍSTICAS DEL TEST

### Aspectos generales

Contenido	Valoración	Puntuación
2.1. Calidad de los materiales del test	Excelente	5
2.2. Calidad de la documentación aportada	Excelente	5
2.3. Fundamentación teórica	Excelente	5
2.4. Adaptación del test	Buena	4
2.5. Calidad de las instrucciones	Excelente	5
2.6. Facilidad para comprender la tarea	Excelente	5
2.7. Facilidad para registrar las respuestas	Buena	4
2.8. Calidad de los ítems (aspectos formales)	Buena	4
2.9. Datos sobre el análisis de los ítems	Buena	4

### Validez de contenido

Contenido	Valoración	Puntuación
2.10. Calidad de la representación del contenido o dominio	Buena	4
2.11. Consultas a expertos	No se aporta	

## Validez de constructo

Contenido	Valoración	Puntuación
<b>2.12. Diseños empleados</b>	<ul style="list-style-type: none"> <li>▪ Correlaciones con otros tests</li> <li>▪ Diferencias entre grupos</li> <li>▪ Análisis factorial exploratorio</li> <li>▪ Discriminación diagnóstica</li> </ul>	
<b>2.13. Tamaño de las muestras</b>	Excelente	5
<b>2.14. Procedimiento de selección de las muestras</b>	Incidental	
<b>2.15. Correlaciones del test con otros tests similares</b>	Excelente	5
<b>2.16. Calidad de los test empleados como criterio o marcador</b>	Buena	4
<b>2.17. Datos sobre el sesgo de los ítems</b>	No se aporta	

### Comentarios sobre la validez de constructo

- Para investigar las características psicométricas de la adaptación española, se administró el BDI-II a las siguientes muestras:

a) Tres muestras de pacientes adultos con diversos trastornos psicológicos que seguían tratamientos psicológicos ambulatorios en centros públicos de la Comunidad de Madrid (número total de pacientes = 712; rango de edad: 18 – 79 años; el porcentaje de mujeres osciló entre 63.6 y 76.5). Como criterio de exclusión se utilizó la existencia de dificultades intelectuales o lingüísticas para completar la prueba.

b) Dos muestras de adultos de la población general con un perfil demográfico (sexo y edad) similar al de la población española (número total de participantes = 569). Los participantes de la primera muestra fueron captados por estudiantes de Psicología de la UCM entre sus familiares y amigos. Los integrantes de la segunda procedían de un estudio sobre depresión y ansiedad realizado en diferentes regiones españolas.

c) Dos muestras de estudiantes universitarios de diversas facultades de la Comunidad de Madrid (número total de estudiantes = 727; rango de edad: 18 – 86 años; el porcentaje de mujeres osciló entre 51.5 y 77.8).

## Validez predictiva

Contenido	Contenido	Puntuación
<b>2.18. Diseño de selección del criterio</b>	<ul style="list-style-type: none"> <li>▪ Concurrente</li> <li>▪ Retrospectivo</li> </ul>	
<b>2.19. Tamaño de las muestras</b>	Bueno	4
<b>2.20. Procedimiento de selección de las muestras</b>	Incidental	
<b>2.21. Correlaciones del test con los criterios</b>	En lugar de correlaciones, se aportan datos basados en diferencias de medias, $d$ de Cohen, índices de discriminación para distintos puntos de corte (sensibilidad, especificidad, eficiencia, kappa, curvas ROC, etc).	

## Comentarios sobre validez predictiva

En este apartado pueden ser incluidos los estudios sobre la utilidad o discriminación diagnóstica de las puntuaciones de la adaptación española del BDI-II entre:

1.- Muestras de pacientes externos clasificados en cuatro grupos diagnósticos (trastornos del estado de ánimo, trastornos de ansiedad, trastornos adaptativos y otros trastornos psicológicos).  
Número total de casos: 322.

2.- Muestra de pacientes con un trastorno depresivo mayor y muestra de pacientes sin depresión.  
Número total de casos: 171.

3.- Muestra de estudiantes universitarios con episodio depresivo mayor y muestra de estudiantes no deprimidos. Número total de casos: 165.

Como criterio se han empleado entrevistas diagnósticas estructuradas: SCID-I VC (estudios 1 y 2) y Quick DIS-III-R (estudio 3).

## Comentarios generales sobre validez

- Se ha utilizado una metodología adecuada para traducir y adaptar los ítems a la población española. El contenido de la prueba incluye todos los criterios propuestos por DSM-IV y CIE-10 para el diagnóstico de un episodio depresivo mayor y casi todos los propuestos (todos los del DSM-IV y el 75% de los del CIE-10) para el diagnóstico de distimia. No se aportan resultados de consultas a expertos.
- Los estudios empíricos de validación de las puntuaciones aportan evidencias procedentes del análisis de la estructura interna del instrumento, de su utilidad diagnóstica y de correlaciones con marcadores relevantes de los constructos de depresión y ansiedad (MCMI-II, LED, STAI y BAI).
- En lo que se refiere a la estructura interna, se realizaron análisis factoriales exploratorios en los que se replicaron los resultados de estudios factoriales en población clínica realizados por los autores de la prueba, obteniéndose dos factores: una *dimensión somática* (pérdida de energía, fatiga, pérdida de placer e interés, cambios en el patrón del sueño, etc) y un *factor cognitivo* (autocríticas, insatisfacción con uno mismo, inutilidad, pensamientos o deseos de suicidio, etc). Dada la alta asociación entre ambos factores, es razonable considerar que la prueba mide un factor dominante de gravedad de la depresión. De hecho, el primer factor en la extracción inicial daba cuenta de un porcentaje de varianza sensiblemente superior (mayor del 30%) al de los factores sucesivos (menores del 7%).
- Las puntuaciones de la prueba presentaron una adecuada validez convergente: elevadas correlaciones (superiores a 0.50) con las medidas de depresión del MCMI-II y del LED. La validez discriminante resultó más comprometida. Por un lado, se obtuvieron correlaciones moderadas con la medida de ansiedad del STAI. Por otro, sin embargo, se observaron altas correlaciones con la medida de ansiedad del BAI. Los autores argumentan que esta elevada asociación se explica por la presencia de síntomas compartidos en las definiciones clínicas de depresión y ansiedad. De hecho, los análisis factoriales conjuntos de los ítems del BAI y el BDI-II revelan la presencia de dos factores correlacionados pero distinguibles: uno de depresión y otro de ansiedad. No se aportan estudios sobre el funcionamiento diferencial de los ítems asociado al sexo, la patología (grupos normales y clínicos), etc.
- Los resultados expuestos indican que el BDI-II es un test apropiado para la medición de la presencia y gravedad de la sintomatología depresiva en la población española, tanto en ámbitos clínicos como de investigación. Sin embargo, estas buenas propiedades no justifican que el juicio diagnóstico sobre si una persona padece o no un trastorno depresivo recaiga solo en los resultados que el test ofrece. El BDI-II es un

test útil para identificar síntomas depresivos y cuantificar su intensidad, lo que obviamente es muy importante para evaluar, por ejemplo, la gravedad de un paciente o su progreso terapéutico. Sin embargo, el diagnóstico de un trastorno depresivo se debe efectuar teniendo en cuenta no solo el tipo y número de síntomas presentes, sino también ciertos criterios de duración, gravedad, curso, incapacidad y ausencia de ciertas causas posibles o de ciertos diagnósticos concurrentes.

### Fiabilidad

Contenido	Valoración	Puntuación
2.22. Datos aportados sobre fiabilidad	<ul style="list-style-type: none"> <li>Coeficientes de fiabilidad para diferentes grupos de sujetos</li> </ul>	

### Equivalencia formas paralelas

Contenido	Valoración	Puntuación
2.23. Tamaño de las muestras	No se aporta	-
2.24. Coeficientes de equivalencia	No se aporta	-

### Consistencia interna

Contenido	Valoración	Puntuación
2.25. Tamaño de las muestras	Excelente	5
2.26. Mediana de los coeficientes de consistencia	Excelente ( $r \geq 0.85$ )	5

### Estabilidad (test-retest)

Contenido	Valoración	Puntuación
2.27. Tamaño de las muestras	No se aporta	
2.28. Coeficientes de estabilidad	No se aporta	

- En un estudio realizado con el test original se obtuvo una correlación test-retest muy elevada (0,93). La muestra estaba formada por 26 pacientes y el intervalo entre aplicaciones fue de una semana.

## Comentarios generales sobre fiabilidad

Los datos sobre la fiabilidad de las puntuaciones se basan exclusivamente en la consistencia interna. Los coeficientes son elevados (superiores a 0.85) y similares a los obtenidos en los estudios internacionales con la versión original del BDI-II.

## Baremación

Contenido	Valoración	Puntuación
<b>2.29. Calidad de las normas</b>		
<b>2.30. Tamaño de las muestras</b>		
<b>2.31. Procedimiento de selección de las muestras</b>	No se aporta	

El Manual no contiene baremos. La principal estrategia para la interpretación de una puntuación consiste en su comparación con 3 puntos de corte y su consecuente clasificación en una de 4 categorías. Sanz y Vázquez (1998) y otras publicaciones de los adaptadores de la prueba, que figuran en la lista de referencias del Manual, ofrecen distintos baremos para las poblaciones general española y de estudiantes españoles universitarios.

## Comentarios generales sobre baremación

- Para facilitar la interpretación de las puntuaciones se ofrecen 3 puntos de corte, que permiten calificar el nivel de depresión como mínima, leve, moderada o grave. La decisión de usar estos (u otros) puntos de corte se debe basar en las características concretas de la muestra y el propósito para el que se esté usando el test. Se advierte al profesional que si el propósito es evitar en lo posible que casos con depresión no sean detectados, entonces los puntos de corte debieran reducirse. Al proceder así se minimizarían los falsos negativos (e incrementarían los falsos positivos).
- El Manual no contiene baremos. Los responsables de la adaptación de la prueba han publicado baremos para las poblaciones general española y de estudiantes universitarios. Se pueden encontrar en Sanz y Vázquez (1998) y en otras referencias bibliográficas del Manual.

## Referencias

Muñiz, J. y Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles del Psicólogo* 31, 108-121.