

EVALUACIÓN DE LA BPR - BATERÍA DE PRUEBAS DE RAZONAMIENTO

RESUMEN DE LA VALORACIÓN DEL TEST

Descripción general

Características	Descripción
Nombre del test	BPR. Batería de Pruebas de Razonamiento
Autor	L.S. Almeida y G.C. Lemos
Autor de la adaptación española	P. Elosua y L.S. Almeida
Variable(s)	Razonamiento general mediante seis pruebas que evalúan capacidades cognitivas específicas
Área de aplicación	Psicología educativa
Soporte	Papel y Lápiz

Valoración general

Características	Valoración	Puntuación
Materiales y documentación	Excelente	5
Fundamentación teórica	Excelente	5
Adaptación	Excelente	5
Análisis de los ítems	Buena-Excelente	4,5
Validez: contenido	-	-
Validez: relación con otras variables	Buena-Excelente	4,3
Validez: estructura interna	Excelente	5
Validez: análisis del DIF	Excelente	5
Fiabilidad: equivalencia	-	-
Fiabilidad: consistencia interna	Excelente	5
Fiabilidad: estabilidad	Adecuada	3
Fiabilidad: TRI	-	-
Fiabilidad: inter-jueces	-	-
Baremos e interpretación de puntuaciones	Buena-Excelente	4,5

Comentarios generales

La Batería de Pruebas de Razonamiento (BPR) es un conjunto de pruebas que se puede aplicar tanto individual como colectivamente. La duración de la aplicación oscila entre los 45 minutos y la hora. Se puede aplicar a personas entre los 9 y los 18 años. Una característica interesante es que la BPR se organiza en tres niveles educativos: Nivel 1 (de 9 a 12 años, de 4º a 6º de Educación Primaria); Nivel 2 (de 12 a 15 años, de 1º a 3º de Educación Secundaria Obligatoria); Nivel 3 (de 15 a 18 años, de 4º de Educación Secundaria Obligatoria a 2º de Bachillerato y Ciclos Formativos de Grado Medio). Obviamente, las diferentes pruebas están adaptadas a la edad de las personas evaluadas. En el Nivel 1 se evalúa Razonamiento Abstracto, Razonamiento Verbal, Razonamiento Numérico y Razonamiento Práctico. En los Niveles 2 y 3 se evalúan las mismas dimensiones que en el Nivel 1 excepto el Razonamiento Práctico e incluyendo el Razonamiento Espacial y el Razonamiento Mecánico. En todos los niveles se obtiene una puntuación en Razonamiento General, resultado de las puntuaciones obtenidas en las diferentes dimensiones. En general, la Batería de Pruebas de Razonamiento (BPR) es una batería interesante y actual para la evaluación del razonamiento, mediante sus seis pruebas que evalúan capacidades cognitivas específicas. Como ventaja, cada aptitud puede evaluarse aisladamente o junto con otras para obtener el índice de Razonamiento general que refleja el nivel en el factor general de razonamiento. La novedad de esta batería frente a otras que existen en el mercado es la inclusión de una prueba de Razonamiento práctico en el nivel más bajo y de Razonamiento mecánico en los niveles 2 y 3.

Los materiales presentan una cuidada edición y especialmente atractiva para la evaluación de niños de menor edad.

Esta batería es el resultado de la adaptación al español de la versión inédita en portugués. El proceso de traducción y adaptación sigue las directrices propuestas por la International Test Commission.

En el análisis de los ítems se presenta información sobre diversos estudios acerca de las características psicométricas de los ítems: dificultad o media, variabilidad y discriminación.

En relación a la fiabilidad, se comprueba la consistencia interna mediante la utilización del alfa ordinal. Para ello se emplea la totalidad de la muestra (dividida por niveles). Por otro lado, también se aportan evidencias de la estabilidad de la medida mediante la correlación test-retest. En este caso no se describe el procedimiento usado para la selección de la submuestra.

Respecto al estudio de validez de las puntuaciones se buscan evidencias internas y externas. En relación a las primeras se comprueba la estructura dimensional mediante Análisis Factorial Exploratorio y Confirmatorio, la invarianza factorial en función del sexo y la relación de las diferentes dimensiones entre sí. En este último análisis se presentan las correlaciones observadas entre las variables que componen cada uno de los niveles y entre estas y el índice de Razonamiento general. En relación a las evidencias externas se comprueba la existencia de diferencias en las puntuaciones en las diferentes dimensiones en función del sexo y la edad, se correlaciona con otros tests de aptitudes e inteligencia y finalmente también se relaciona con el rendimiento académico.

En relación a la baremación se emplean puntuaciones transformadas a escalas CI y la corrección automatizada proporciona toda la información necesaria para su correcta interpretación.

Sería muy interesante que se analizara desde la Teoría de la Respuesta a los Ítems, ya que podría indicar en qué niveles de habilidad es más adecuada la utilización de la batería.

Otro punto fuerte de la Batería de Pruebas de Razonamiento es la adaptación al euskera de sus tres niveles.

Para finalizar, la BPR podría ser perfectamente adecuada para detectar dificultades de aprendizaje, y, en general, para determinar si el alumnado tiene necesidades de apoyo educativo y de qué tipo, dadas las distintas pruebas que la componen. También se podría utilizar para evaluar las altas capacidades y discapacidades, como se indica en el manual.

ANÁLISIS DETALLADO DE LA PRUEBA

1. DESCRIPCIÓN GENERAL DEL TEST

1.1. Nombre del test:

- BPR. Batería de Pruebas de Razonamiento.

1.2. Nombre del test en su versión original:

- Bateria de Provas de Raciocinio

1.3. Autor del test original:

- L.S. Almeida y G.C. Lemos

1.4. Autor de la adaptación española:

- P. Elosua y L.S. Almeida

1.5. Editor del test en su versión original:

- No publicada

1.6. Editor de la adaptación española:

- TEA Ediciones

1.7. Fecha de publicación del test original:

- No publicado

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

- 2016

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

- 2016

1.10. Área general de la(s) variable(s) que pretende medir el test:

- Inteligencia, Aptitudes, Habilidades, Potencial de aprendizaje, Razonamiento General.

1.11. Breve descripción de la(s) variable(s) que pretende medir el test:

- La Bateria de Pruebas de Razonamiento (BPR) pretende la evaluación del razonamiento, mediante seis pruebas que evalúan capacidades cognitivas específicas relacionadas con el Razonamiento abstracto (habilidad para captar y aplicar relaciones lógicas entre elementos figurativos sin carga cultural; razonamiento lógico-deductivo), el Razonamiento verbal (capacidad para captar relaciones entre palabras –inducción- y de aplicar tales relaciones a nuevos contextos –deducción-, partiendo de la comprensión verbal o del significado de las palabras), el Razonamiento numérico (habilidad para captar y aplicar las relaciones lógicas entre los números), el Razonamiento práctico (capacidad de análisis y de síntesis de información compleja para la resolución de problemas que requieren atención, comprensión, organización y razonamiento), el Razonamiento espacial (habilidad para inferir movimiento, anticipar formas en un espacio dado y

dar continuidad a transformaciones y rotaciones de figuras que requieren la tridimensionalidad) y el Razonamiento mecánico (habilidad para codificar y organizar la información descriptiva de los problemas con el fin de dar solución a una cuestión). Cada aptitud puede evaluarse aisladamente o junto con otras para obtener un índice de Razonamiento general que refleja el nivel en el factor general de razonamiento

- El rango de edad de la población diana va desde los 9 hasta los 18 años. En función de la edad de la persona evaluada le corresponderá un nivel diferente. Así, la BPR se organiza en 3 niveles con pruebas específicas adaptadas a los rangos de edad para los que están destinadas (que se corresponden con cuadernillos diferentes):
 - Nivel 1: de 9 a 12 años (de 4º a 6º de Educación Primaria). En este nivel se aplicarán las pruebas de razonamiento abstracto, razonamiento verbal, razonamiento numérico y razonamiento mecánico o práctico.
 - Nivel 2: de 12 a 15 años (de 1º a 3º de Educación Secundaria Obligatoria) y nivel 3: de 15 a 18 años (de 4º de Educación Secundaria Obligatoria a 2º de Bachillerato y Ciclos Formativos de Grado Medio). En los niveles 2 y 3 En los Niveles 2 y 3 se aplicarán las pruebas de razonamiento abstracto, razonamiento verbal, razonamiento numérico, razonamiento espacial y razonamiento mecánico o práctico. Si bien las dimensiones evaluadas en los niveles 2 y 3 son las mismas, los ítems que las evalúan difieren para adaptarlas a las características propias de la edad. Tras obtener una puntuación en cada una de las dimensiones también se aporta una puntuación en Razonamiento General.

1.12. Áreas de aplicación:

- Psicología educativa, psicología clínica y psicología forense.

1.13. Formato de los ítems:

- Respuesta construida (solo en la escala de Razonamiento Numérico), Elección múltiple (acierto/fallo).

1.14. Número de ítems:

- Nivel 1
 - Razonamiento abstracto (Ra): 20
 - Razonamiento verbal (Rv): 20
 - Razonamiento numérico (Rn): 15
 - Razonamiento práctico (Rp): 46
- Nivel 2 y 3
 - Razonamiento abstracto (Ra): 25

- Razonamiento verbal (Rv): 25
- Razonamiento espacial (Re): 20
- Razonamiento numérico (Rn): 20
- Razonamiento mecánico (Rm): 25

Cada una de las pruebas contiene además dos ítems de ensayo.

1.15. Soporte:

- Papel y lápiz

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

- Nivel B.

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

- Está destinado a un rango de edades entre los 9 años y los 18 años.
 - Nivel 1: de 9 a 12 años (de 4º a 6º de Educación Primaria).
 - Nivel 2: de 12 a 15 años (de 1º a 3º de Educación Secundaria Obligatoria).
 - Nivel 3: de 15 a 18 años (de 4º de Educación Secundaria Obligatoria a 2º de Bachillerato y Ciclos Formativos de Grado Medio).
- Existe una versión en euskera.

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

- Existen tres versiones del cuestionario en función de la edad de las personas evaluadas. Desde los 9 a los 12 años (nivel 1), de los 12 a los 15 años (nivel 2) de los 15 a los 18 años (nivel 3).

1.19. Procedimiento de corrección:

- Automatizada por ordenador (plataforma de corrección on-line de TEA).

1.20. Puntuaciones:

- La puntuación directa de cada una de las pruebas indica el número de respuestas correctas. En los ítems en los que se solicitan varias respuestas todas deben ser correctas al no considerarse respuestas parcialmente correctas.
- A partir de las puntuaciones obtenidas en cada una de las dimensiones se estima una puntuación en Razonamiento General. Dicho cálculo se apoya en la relación existente entre las diferentes pruebas y el factor general, y la fórmula está implementada en la plataforma de corrección TEAcorrige. Por lo que la corrección se tiene que realizar usando la plataforma online TEAcorrige. Para ello es necesario disponer de un PIN de corrección que se adquiere al comprar las hojas de respuestas. Tras aplicar la prueba se introducen las respuestas en la plataforma.
- Una vez se tienen las puntuaciones directas se ofrece también los resultados en puntuaciones derivadas con media 100 y desviación típica 15.

1.21. Escalas utilizadas:

- Puntuación Directa, CI de desviación (Media 100 y desviación típica 15).

1.22. Posibilidad de obtener informes automatizados:

- No. La plataforma de corrección proporciona un perfil de resultados, donde aparecen las puntuaciones directas (PD) de cada prueba y la puntuación transformada de CI (PT) correspondiente. También se obtiene el perfil gráfico de resultados, donde se representan las puntuaciones PT obtenidas por el evaluado en cada una de las pruebas. Esa puntuación es el resultado de comparar su puntuación directa con una muestra normativa de su mismo curso y, en caso de que se considere necesario, ajustada al sexo del evaluado.

1.23. Tiempo estimado para la aplicación del test:

- En aplicación individual: 30 minutos Nivel 1 y 37 minutos Niveles 2 y 3 (suma de tiempos indicados en el Manual para cada una de las pruebas); 45-60 minutos en la página web (tiempo estimado incluyendo instrucciones y ejemplos).
- En aplicación colectiva: 30 minutos Nivel 1 y 37 minutos Niveles 2 y 3. (suma de tiempos indicados en el Manual para la realización de cada una de las pruebas); 45-60 minutos en la página web (tiempo estimado incluyendo instrucciones y ejemplos).

1.24. Documentación aportada por el editor:

- Manual, cuadernillos de aplicación para los tres niveles, hojas de respuestas y pin para corrección de 25 usos a través de TEACorrige.

1.25. Precio de un juego completo de la prueba:

- Juego completo en castellano (Manual, 5 cuadernillos de cada nivel, Kit corrección 25 usos): 130,40 euros. (11/09/2017).
- Juego completo en euskera: 240,60 euros. (11/09/2017).

1.26. Precio y número de ejemplares del paquete de cuadernillos:

Cuadernillo nivel 1 (10 ejemplares): 69,82 euros

Cuadernillo nivel 2 (10 ejemplares): 69,82 euros

Cuadernillo nivel 3 (10 ejemplares): 69,82 euros

(11/09/2017)

Cuadernillo Nivel 1 euskera (10 ejemplares): 121 euros

Cuadernillo Nivel 2 euskera (10 ejemplares): 121 euros

Cuadernillo Nivel 3 euskera (10 ejemplares): 121 euros

1.27. Precio y número de ejemplares del paquete de hojas de respuesta:

- Paquete de corrección (25 hojas de respuesta y PIN para 25 correcciones): 14,02 euros

1.28. Precio de la corrección y/o elaboración de informes por parte del editor:

- Paquete de corrección (25 hojas de respuesta y PIN para 25 correcciones): 14,02 euros

2. VALORACIÓN DE LAS CARACTERÍSTICAS DEL TEST

2.1. 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	Excelente	5
2.5. Desarrollo de los ítems del test	Buena-Excelente	4,5
2.6. Calidad de las instrucciones para el participante	Excelente	5
2.7. Calidad de las instrucciones (administración, puntuación, interpretación)	Excelente	5
2.8. Facilidad para registrar las respuestas	Excelente	5
2.9. Bibliografía del manual	Excelente	5
2.10. Datos sobre el análisis de los ítems	Buenos-Excelentes	4,5

2.11. Validez

2.11.1. Evidencias de validez de contenido

Contenido	Valoración	Puntuación
2.11.1.1. Calidad de la representación del contenido o dominio	-	-
2.11.1.2. Consultas a expertos	-	-

2.11.2. Evidencias de validez basadas en la relación entre las puntuaciones del test y otras variables

2.11.2.1. Evidencias de validez basadas en la relación con otras variables

Contenido	Valoración	Puntuación
2.11.2.1.1. Diseños empleados	Correlaciones con otros tests, Diferencias entre grupos	
2.11.2.1.2. Tamaño de las muestras	Varios estudios con muestras de tamaño moderado o con alguna muestra grande y otras pequeñas	4
2.11.2.1.3. Procedimiento de selección de las muestras	Incidental	
2.11.2.1.4. Calidad de los tests empleados como criterio o marcador	Buena- Excelente	4,5
2.11.2.1.5. Promedio de las correlaciones con otros tests que miden constructos similares	Buena	4
2.11.2.1.6. Promedio de las correlaciones con otros tests que miden constructos no relacionados	Buena	4
2.11.2.1.7. Resultados de la matriz multirasgo-multimétodo	-	-
2.11.2.1.8. Resultados de las diferencias intergrupo	Excelente	5

2.11.2.2. Evidencias de validez basadas en la relación entre las puntuaciones del test y un criterio

Contenido	Valoración	Puntuación
2.11.2.2.1. Criterios empleados	-	
2.11.2.2.2. Calidad de los criterios empleados	-	-
2.11.2.2.3. Relación temporal entre test y criterio	-	
2.11.2.2.4. Tamaño de las muestras	-	-
2.11.2.2.5. Procedimiento de selección de las muestras	-	
2.11.2.2.6. Promedio de las correlaciones del test con los criterios	-	-

2.11.3. Evidencias de validez basadas en la estructura interna

Contenido	Valoración	Puntuación
2.11.3.1. Resultados del análisis factorial	Excelente	5
2.11.3.2. Funcionamiento diferencial de los ítems	Excelente	5

2.11.4. Acomodaciones en la administración del test

Contenido	Valoración	Puntuación
2.11.4. El manual del test informa sobre las acomodaciones en la administración del test	Sí	

Comentarios generales sobre evidencias de validez

El estudio de validación de las puntuaciones de la batería se realizó atendiendo a dos aspectos diferentes:

- Evidencias de validez basadas en la estructura interna de la batería. Para ello, se comprobaron las relaciones entre los ítems de la batería y su concordancia con el modelo teórico propuesto para la construcción del mismo. Se han utilizado fuentes de evidencia basadas en la estructura interna de interrelaciones entre las pruebas, análisis factoriales exploratorios de los ítems, análisis factoriales confirmatorios para contrastar el modelo teórico jerárquico de partida.
- Evidencias de validez basadas en la relación con otras variables. Para ello se emplearon dos tipos de variables:
 - Otros cuestionarios de inteligencia y aptitudes: Se correlacionó las puntuaciones obtenidas en las diferentes dimensiones de la batería con las obtenidas en el BADyG, Batería de Aptitudes Diferenciales y Generales, Matrices Progresivas de Raven y Test de Dominós D-48. En general, los resultados van en la dirección esperada. Así, el índice de Razonamiento General mostró buenas correlaciones con todas las escalas generales de inteligencia analizadas tanto para el Nivel 2 como para el Nivel 3 (promedio de correlación 0,60). Por otro lado hay que tener en cuenta que las correlaciones entre cada una de las pruebas de la BPR y las pruebas del BADyG, Matrices Progresivas y Test de Dominós D-48 oscilan en un rango entre -0,01 y 0,72. Los tamaños muestrales fueron de 41 participantes en el Nivel 2 y de 38 en el

Nivel 3). Por otro lado no se especifica cómo se obtuvo la muestra en la que se realizan los análisis, ni tampoco se aportan evidencias para el Nivel 1.

- Rendimiento Académico: Se correlacionan las puntuaciones obtenidas en la batería con las calificaciones en Lengua y Matemáticas. De nuevo, y en general, las correlaciones van en el sentido esperado, aunque tal y como se comenta en el Manual “las correlaciones entre las pruebas de la BPR con la asignatura Lengua muestran un patrón dispar”. En este caso los tamaños muestrales son más grandes (n=124 en el Nivel 2 y n=250 en el Nivel 3), si bien no se especifica cómo se obtuvieron dichas muestras ni tampoco se aportan datos referidos a las pruebas de Nivel 1.
- Diferencias entre grupos de edad y sexo, y la evolución de las puntuaciones en función del curso.

Así mismo se comprobó tanto la equivalencia estructural y métrica en función del sexo (invarianza configural y métrica).

Todas las evidencias apuntan en la misma dirección avalando que la versión española cumple con los objetivos para la que fue construida.

En cuanto a las acomodaciones, a la hora de aplicar la batería a personas con discapacidad, se indica que el evaluador deberá tener una serie de consideraciones generales relacionadas con el entorno y con la propia BPR.

2.12. Fiabilidad

Contenido	Valoración	Puntuación
2.12.1. Datos aportados sobre fiabilidad	Varios coeficientes de fiabilidad (para cada escala y subescala), coeficientes de fiabilidad para diferentes grupos de personas (para cada nivel de la batería), errores típicos de medida (para cada escala y subescala)	

2.12.2. Equivalencia formas paralelas

Contenido	Valoración	Puntuación
2.12.2.1. Tamaño de las muestras	-	-
2.12.2.2. Puesta a prueba de los supuestos de paralelismo	-	-
2.12.2.3. Promedio de coeficientes de equivalencia	-	-

2.12.3. Consistencia interna

Contenido	Valoración	Puntuación
2.12.3.1. Tamaño de las muestras	Varios estudios con muestras grandes	5
2.11.3.2. Coeficientes de consistencia interna presentados	Alfa ordinal	
2.11.3.3. Promedio de los coeficientes de consistencia	Excelente	5

2.12.4. Estabilidad (test-retest)

Contenido	Valoración	Puntuación
2.12.4.1. Tamaño de las muestras	Un estudio con una muestra moderada y otro con una muestra pequeña	2,5
2.12.4.2. Coeficientes de estabilidad	Adecuada-Buena	3,5

2.12.5. Cuantificación de la precisión mediante TRI

Contenido	Valoración	Puntuación
2.12.5.1. Tamaño de las muestras	-	-
2.12.5.2. Coeficientes proporcionados	-	
2.12.5.3. Tamaño de los coeficientes	-	-

2.12.6. Fiabilidad inter-jueces

Contenido	Valoración	Puntuación
2.12.6.1. Tipos de coeficientes presentados	-	
2.12.6.2. Promedio de los coeficientes	-	-

Comentarios generales sobre fiabilidad

Para el cálculo de la fiabilidad de las puntuaciones de la batería se han empleado dos estrategias diferentes. En primer lugar se evalúa la consistencia de cada una de las dimensiones que componen la batería empleando el coeficiente alfa ordinal. La fiabilidad

estimada fue buena en los tres niveles y en todas las escalas. Así mismo se informa del error típico de medida de cada dimensión. También se estima la *estabilidad temporal* mediante un test-retest, con un lapso de tiempo de dos semanas. Para ello se aplicó el Nivel 1 de la batería a 150 participantes y el Nivel 2 a 71. Los resultados aportan correlaciones entre 0,63 y 0,88 para las escalas en el Nivel 1 (promedio 0,75) y entre 0,57 y 0,86 para las escalas en el Nivel 2 (promedio 0,71). En el Manual se explica que los resultados obtenidos del coeficiente pueden ser debidos a posibles efectos de aprendizaje y memoria entre la primera y segunda aplicación (los valores promedios de las pruebas fueron superiores en la segunda aplicación).

2.13. Baremación e interpretación de las puntuaciones

2.13.1. Interpretación normativa de las puntuaciones

Contenido	Valoración	Puntuación
2.13.1.1. Calidad de las normas	Varios baremos dirigidos a diversos estratos poblacionales	4
2.13.1.2. Tamaño de las muestras	Grande-Muy grande	4,5
2.13.1.3. Aplicación de estrategia de tipificación continua	Sí	
2.12.1.4. Procedimiento de selección de las muestras	Incidental	
2.12.1.5. Actualización de baremos	Excelente	5

2.13.2. Interpretación referida a criterio

Contenido	Valoración	Puntuación
2.13.2.1. Adecuación del establecimiento de los puntos de corte	-	-
2.13.2.2. Procedimiento empleado para fijar los puntos de corte		-
2.13.2.3. Procedimiento de obtención del acuerdo inter-jueces		-
2.13.2.4. Valor del coeficiente de acuerdo inter-jueces	-	-

Comentarios generales sobre baremación e interpretación de las puntuaciones

Los baremos están disponibles en función del sexo y curso escolar desde 4º de Primaria hasta Bachillerato, incluyendo los Ciclos Formativos de Grado Medio, basándose en el sistema educativo español. Además, para la obtención de los datos de baremación se contactó con centros de enseñanza de diferentes regiones españolas (Andalucía, Aragón, Cantabria, Castilla y León, Cataluña, Extremadura, Galicia, Navarra, La Rioja y País Vasco). La muestra total es de 2698 participantes (1156 en el Nivel 1, 878 en el Nivel 2 y 664 participantes en el Nivel 3).

El sistema de corrección automático proporcionado por TEA para esta batería informa tanto sobre la puntuación directa como una puntuación transformada tipo CI (media=100, desviación típica=15). También se indica el intervalo de confianza de la puntuación obtenida a un Nivel del Confianza del 95%. Además, en el informe que proporciona el sistema de corrección automático también se representa gráficamente las puntuaciones obtenidas en cada una de las pruebas evaluadas. En el gráfico existen diferentes tonalidades para señalar el nivel de la puntuación (extremo bajo, bajo, promedio, alto, extremo alto). En el Manual se incluye un apartado en el que se describen las puntuaciones transformadas que se obtienen con los baremos de la BPR y la interpretación de las mismas. Por otro lado, se aportan casos prácticos para que el profesional pueda ver ejemplos de cómo la batería se incluyó en diferentes procesos de evaluación y se utilizaron sus resultados para orientar en la intervención.