

Evaluación de

Matrices Progresivas

∞ Raven 2 ∞

RESUMEN DE LA VALORACIÓN DEL TEST

Descripción general

Característica	Descripción
Nombre del test	Matrices progresivas de Raven 2 (Raven's 2)
Autor	Pearson
Autor de la adaptación española	Departamento de I+D de Pearson Clinical & Talent Assessment: Ana Hernández, Cristina Aguilar, Érica
Variable(s)	Aptitud cognitiva
Áreas de aplicación	Psicología clínica, educativa, del trabajo y las organizaciones, forense e investigación
Soporte	Papel y lápiz; Soporte digital

Valoración general

Característica	Valoración	Puntuación
Materiales y documentación	Excelente	5
Fundamentación teórica	Excelente	4,5
Adaptación	Excelente	5
Análisis de ítems	Adecuada	3
Validez: contenido	Excelente	4,25
Validez: relación con otras variables	Excelente	4,25
Validez: estructura interna	No se aporta	--
Validez: análisis del DIF	No se aporta	--
Fiabilidad: equivalencia	No se aporta	--
Fiabilidad: consistencia interna	Excelente	4,5
Fiabilidad: estabilidad	Excelente	5
Fiabilidad: TRI	Excelente	4,5
Fiabilidad inter-jueces	No se aporta	--
Baremos e interpretación de puntuaciones	Excelente	4,75

Comentarios generales

El Raven's 2 Matrices Progresivas es la adaptación a población española del Raven's 2 Progressive Matrices Clinical Edition. Mediante este test se pretende evaluar de forma no verbal la aptitud cognitiva. El test consiste en formas geométricas identificables con independencia del nivel educativo, y es evaluado sin emplear respuestas orales o escritas. De este modo se minimizan los diferentes aspectos culturales que pueden estar afectando a los resultados.

La aplicación del test puede realizarse en personas desde los 4 años hasta los 69 años y 11 meses, pudiendo hacerse la aplicación tanto en papel y lápiz como informatizadamente (en este caso se disponen de dos versiones, una larga y otra abreviada). La corrección del test se debe realizar a través de la plataforma online Q-global proporcionada por Pearson Clinical & Talent Assessment.

A nivel formal, el Raven destaca por la excelente calidad de los materiales, por la calidad de la documentación aportada. Cabe destacar especialmente la sencillez con la que pueden registrarse las respuestas proporcionadas por los evaluados.

Los valores estimados de la fiabilidad, evaluada desde la Teoría Clásica de los Tests como consistencia interna (dos mitades) como estabilidad temporal (test-retest) se pueden considerar buenas y excelentes en los diferentes grupos de edad evaluados. Desde el marco de la Teoría de la Respuesta a los Ítems (fiabilidad marginal), empleada para evaluar la fiabilidad de los tests informatizados, se obtiene una fiabilidad buena para la versión corta y excelente para la larga.

Las evidencias de validez del Raven's 2 son adecuadas. Se muestran evidencias de validez relacionadas con el contenido, con el proceso de respuesta y en relación con otras variables. Todas ellas son adecuadas, por lo que se defiende el uso del test para la evaluación de la aptitud cognitiva.

Todos los test tienen margen de mejora. Aunque se ha mostrado que en el desarrollo del test se ha tratado de minimizar el efecto cultural en las respuestas al mismo, para futuras ediciones del test se recomienda incrementar el número de estudios sobre las evidencias de validez con muestra específicamente española.

ANÁLISIS DETALLADO DE LA PRUEBA

1. DESCRIPCIÓN GENERAL DEL TEST

1.1. Nombre del test:

Matrices Progresivas de Raven 2 (Raven's 2)

1.2. Nombre del test en su versión original:

Raven's 2 Progressive Matrices Clinical Edition

1.3. Autor/es del test original:

Pearson

1.4. Autor/es de la adaptación española:

Departamento de I+D de Pearson Clinical & Talent Assessment: Ana Hernández, Cristina Aguilar, Èrica Paradell, Frédérique Vallar

1.5. Editor del test en su versión original:

NSC Pearson Inc, 2018

1.6. Editor de la adaptación española:

Pearson Educación 2019

1.7. Fecha de publicación del test original:

2018

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

2019

1.9. Fecha de la última revisión del test:

--

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

Inteligencia, Aptitudes

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

Evalúa, a través de ítems no verbales la aptitud eductiva, es decir, la capacidad de razonar y resolver problemas complejos. Dado que la aptitud eductiva es uno de los componentes fundamentales de la inteligencia general o factor g, el test proporciona una estimación del funcionamiento cognitivo general de la persona evaluada.

1.12. Áreas de aplicación:

Psicología clínica, educativa, del trabajo y las organizaciones, forense e investigación

1.13. Formato de los ítems:

Elección múltiple

1.14. Número de ítems:

Formato papel y lápiz: 60 ítems (distribuidos en 5 grupos aplicables en función de la edad de 12 ítems)

Formato digital largo: 36 ítems (al evaluar edades comprendidas entre los 4 años y los 8 años y 11 meses), 48 ítems (para las edades comprendidas entre los 9 años y los 69 años y 11 meses)

Formato digital breve: 24 ítems (con independencia de la edad del evaluado)

1.15. Soporte:

Papel y lápiz y online (versión larga y reducida)

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:

Población general

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

Existen tres formas del test que el usuario puede emplear dependiendo de sus intereses: formato papel y lápiz, y formato digital (con sus dos versiones larga y reducida). En el caso de la aplicación digital, el sistema selecciona para cada evaluado un conjunto de ítems único seleccionados de entre los que conforman el banco de ítems. La correlación entre las puntuaciones obtenidas en la versión digital larga, reducida y en papel es de 0,93.

1.19. Procedimiento de corrección:

Mediante plataforma online

1.20. Puntuaciones:

Puntuación directa: Sumatorio del número de ítems que han sido contestados correctamente

Percentiles

Puntuaciones estandarizadas: Puntuación típica transformada (CI, Media =100, Desviación típica = 15) y Puntuación típica transformada normalizada (Eneatipos)

Calificación cualitativa, dividiendo las puntuaciones en Muy bajo, Bajo, Medio-bajo, Medio, Medio-alto, Alto y Muy alto) en función de la puntuación obtenida en el CI.

Puntuación en aptitud: Cálculo mediante el modelo de Rasch (Teoría de las Respuesta a los Ítems) del nivel de habilidad de la persona.

Puntuación de edad equivalente: Edad promedio en años y meses en los que es más frecuente obtener una puntuación en aptitud.

1.21. Escalas utilizadas:

Percentiles, Puntuaciones típicas, Eneatipos

1.22. Posibilidad de obtener informes automatizados:

La plataforma online genera el informe del perfil del evaluado

1.23. Tiempo estimado para la aplicación del test:

Entre 20 y 45 minutos en función de la forma aplicada

1.24. Documentación aportada por el editor:

Manual, cuaderno de estímulos y hoja de respuestas

1.25. Precio de un juego completo de la prueba:

275€ Incluye: Manual, cuaderno de estímulos, hoja de respuestas (25 unidades) y 25 usos de corrección online (permite también la aplicación online)

Se pueden comprar el Manual (135€) y el Cuaderno de estímulos (145€) de forma independiente.

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

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

64,25€ (Incluye: 25 usos aplicación/corrección - Aplicación Manual y 25 Hojas de respuestas)

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

Se adquiere juntamente con las hojas de respuesta

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	4,5
2.4. Adaptación del test	Excelente	5
2.5. Desarrollo de los ítems del test	Excelente	4,5
2.6. Calidad de las instrucciones para el participante	Excelente	5
2.7. Calidad de las instrucciones (administración, puntuación, interpretación)	Excelente	4,5
2.8. Facilidad para registrar las respuestas	Excelente	4,5
2.9. Bibliografía del manual	Buena	4
2.10. Datos sobre el análisis de los ítems	Adecuados	3

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	Excelente	4,5
2.11.1.2. Consultas a expertos	Buena	4

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 entre las puntuaciones del test y otras variables:

Contenido	Valoración	Puntuación
2.11.2.1.1. Diseños empleados	Correlación con otras variables Diferencias entre grupos	
2.11.2.1.2. Tamaño de las muestras	Adecuada pero con algunas carencias	2
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	3,5
2.11.2.1.5. Promedio de las correlaciones con otros tests que miden constructos similares	Excelente	5
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 multirrasgo-multimétodo	No se aporta	--
2.11.2.1.8. Resultados de las diferencias intergrupo	Buena	4

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	Correlación con versiones anteriores de la escala (escalas SPM, CPM y APM)	
2.11.2.2.2. Calidad de los criterios empleados	Buena	4
2.11.2.2.3. Relación temporal entre test y criterio	Concurrente	
2.11.2.2.4. Tamaño de las muestras	Adecuada pero con algunas carencias	2
2.11.2.2.5. Procedimiento de selección de las muestras	Incidental	
2.11.2.2.6. Promedio de las correlaciones del test con los criterios	Excelente	5

2.11.3. Evidencias de validez basadas en la estructura interna:

Contenido	Valoración	Puntuación
2.11.3.1. Resultados del análisis factorial	No se aporta	--
2.11.3.2. Funcionamiento diferencial de los ítems	No se aporta	--

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	

2.11.5. Comentarios generales sobre evidencias de validez:

El Raven's 2 ha mostrado múltiples evidencias de validez que respaldan su uso como una estimación de la aptitud cognitiva general. En el manual se muestran diferentes evidencias de validez:

- En relación con el contenido: Diferentes expertos se aseguraron de la representatividad y relevancia de los ítems que componen el test.

- En relación con los procesos de respuesta: Se realizaron diferentes estudios tratando de eliminar procesos de respuesta irrelevantes para la evaluación de la aptitud cognitiva (agudeza visual, complejidad en el diseño, presión debida al límite de tiempo para responder...)

- En relación con otras variables: Para ello se correlacionó las respuestas del Raven's 2 con las versiones anteriores del Raven (SPM, CPM, y APM) así como con el Kaufman Brief Intelligence Test- 2 (KBIT-2), mostrando adecuadas evidencias de validez tanto convergente como divergente (con las puntuaciones verbales ofrecidas por el KBIT-2). Así mismo se obtuvieron diferencias en función de la pertenencia a grupos especiales (alta capacidad intelectual vs. discapacidad intelectual).

Toda la información aportada en el manual muestra unas adecuadas evidencias de validez del test. Es cierto que, aunque las autoras de la adaptación española hacen énfasis en la no contaminación cultural/verbal del test, la mayoría de los trabajos sobre las evidencias de validez se han realizado en población extranjera (de hecho el test KBIT-2 no está adaptado a población española), por lo que sería recomendable, para futuras ediciones, incrementar los estudios de evidencias de validez específicos en población española.

2.12. Fiabilidad:

Contenido	Valoración	Puntuación
2.12.1. Datos aportados sobre fiabilidad	Varios coeficientes de fiabilidad Coeficientes de fiabilidad para cada diferentes grupos de personas Estudios de fiabilidad marginal desde la Teoría de la Respuesta a los Ítems Error típico de medida para diferentes grupos de personas	

2.12.2. Equivalencia formas paralelas:

Contenido	Valoración	Puntuación
2.12.2.1. Tamaño de las muestras	No se aporta	--
2.12.2.2. Puesta a prueba de los supuestos de paralelismo	No se aporta	--
2.12.2.3. Promedio de coeficientes de equivalencia	No se aporta	--

2.12.3. Consistencia interna:

Contenido	Valoración	Puntuación
2.12.3.1. Tamaño de las muestras	Buena	4
2.12.3.2. Coeficientes de consistencia interna presentados	Dos mitades	
2.12.3.3. Promedio de coeficientes de consistencia	Excelente	5

2.12.4. Estabilidad (test-retest):

Contenido	Valoración	Puntuación
2.12.4.1. Tamaño de las muestras	Adecuada con carencias	2
2.12.4.2. Coeficientes de estabilidad	Excelente	5

2.12.5. Cuantificación de la precisión mediante TRI:

Contenido	Valoración	Puntuación
2.12.5.1. Tamaño de las muestras	No se aporta	--
2.12.5.2. Coeficientes proporcionados	Fiabilidad marginal	
2.12.5.3. Tamaño de los coeficientes	Escala digital larga: Excelente Escala digital corta: Buena	4,5

2.12.6. Fiabilidad inter-jueces:

Contenido	Valoración	Puntuación
2.12.6.1. Tipos de coeficientes presentados	No se aporta	
2.12.6.2. Promedio de los coeficientes	No se aporta	--

2.12.7. Comentarios generales sobre evidencias de fiabilidad:

Las autoras proporcionan múltiples indicadores de la fiabilidad del test, tanto desde la perspectiva clásica como desde la Teoría de la Respuesta a los Ítems. Desde la perspectiva clásica se evalúa la fiabilidad como consistencia interna mediante el procedimiento de dos mitades y como estabilidad en la medida mediante test-retest. En ambos casos se realizan agrupaciones por edad obteniendo valores de fiabilidad considerados como buenos y excelentes.

Desde el marco de la Teoría de la Respuesta a los Ítems, para las versiones informatizadas tanto larga como breve, se calcula la fiabilidad marginal obteniendo valores considerados como buenos (para la versión reducida), y como excelentes (para la versión larga).

También se ha calculado la fiabilidad para poblaciones especiales como son las personas con alta capacidad intelectual como con discapacidad intelectual, obteniéndose valores considerados como buenos.

Los estudios desde la TRI como en poblaciones especiales fueron realizados en muestras estadounidenses y europeas. Aunque se ha realizado todo un proceso para tratar de que el test esté libre de contaminación cultural, sería recomendable replicar los estudios en los que se emplea muestra internacional con muestra española.

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	Excelente	4,5
2.13.1.2. Tamaño de las muestras	Bueno	4
2.13.1.3. Aplicación de estrategia de tipificación continua	No	
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	No se aporta	--
2.13.2.2. Procedimiento empleado para fijar los puntos de corte	No se aporta	
2.13.2.3. Procedimiento de obtención del acuerdo inter-jueces	No se aporta	
2.13.2.4. Valor del coeficiente de acuerdo inter-jueces	No se aporta	--

2.13.3. Comentarios generales sobre baremación e interpretación de las puntuaciones:

En el manual del Raven's 2 aparecen las normas de interpretación que son sencillas de interpretar. Los baremos se han obtenido a partir de una muestra amplia y proporcionan múltiples tipos de puntuaciones (directas, percentiles, CI, Eneatipos, puntuación de aptitud y puntuación de edad equivalente). Todas estas puntuaciones son proporcionadas tras la corrección automatizada a través de una plataforma online. Es destacable que los autores del manual proporcionan tres casos en los que ejemplifican la interpretación de las puntuaciones proporcionadas por el test.