Consejo General de la Psicología
ESPAÑA

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 de

Batería de Actividades mentales Diferenciales y Generales, Nivel E1, Renovado

BADyG/E1-r

Prueba valorada en la ronda de evaluaciones del año 2019

RESUMEN DE LA VALORACIÓN DEL TEST

Descripción general

Característica	Descripción
Nombre del test	BADyG/E1-r: Batería de actividades mentales Diferenciales y Generales, Nivel E1, renovado
Autor	Carlos Yuste Hernanz y David Yuste Peña
Autor de la adaptación española	-
Variable(s)	Inteligencia, habilidades
Áreas de aplicación	Psicología educativa
Soporte	Papel y lápiz

Valoración general

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

Nota. El signo – se interpreta como que no se aporta información o bien que no procede.

Comentarios generales

La batería BADyG/E1-r es un instrumento de evaluación de las actividades mentales básicas ampliamente utilizado en el contexto educativo desde sus primeras versiones hace ya más de 30 años. El instrumento es aplicable tanto de forma individual como colectiva al alumnado de 1º y 2º curso de Educación Primaria, precisándose entre 1 hora 15 minutos y 1 hora 45 minutos para su aplicación.

Los autores en esta edición aportan nuevos baremos y actualizan y completan la fundamentación en un Manual Teórico-Técnico Completo (MTTC), a la que se puede acceder a través de la página web. Ya desde su anterior actualización el instrumento permite la corrección mediante el Gestor comercial BADyG que además de agilizar el proceso ofrece informes tanto colectivos como individuales. Es un instrumento sencillo tanto en lo relativo a su administración como a su corrección e interpretación que viene apoyada por propuestas para la estimulación en los casos en que es aconsejable.

La bateria BADyG/E1-r mide la inteligencia a través de aspectos de actividades mentales adquiridas culturalmente y rememorativas, pero también de operaciones de razonamiento más abstractas y elaborativas. Consta de 6 pruebas básicas: relaciones analógicas, problemas numérico-verbales, matrices lógicas, órdenes verbales complejas, cálculo numérico y figuras giradas; y 3 pruebas complementarias: memoria visoauditiva, alteraciones en la escritura y discriminación de diferencias. También ofrece 4 factores globales: Factor Verbal, Factor Numérico, Factor Visoespacial y Razonamiento Lógico (RR); y 1 factor general: Inteligencia General (IG). Este instrumento de evaluación permite estimar altas capacidades ya que puede discriminar hasta el Cociente Intelectual de 153, teniendo en cuenta la puntuación en IG y la edad cronológica.

Las evidencias de validez aportadas en la documentación son adecuadas y los resultados son apropiados para la finalidad del test.

Concretamente, las evidencias de validez de contenido se expresan en la definición del dominio, que se explica de manera minuciosa y con detalle en el MTTC. Se echa en falta en la definición cuestiones como el ocio, la comunicación social, el aprendizaje escolar, el desarrollo profesional o las operaciones de introspección. Para esta definición se han consultado tests prestigiosos que pretenden medir la inteligencia y se han tenido en cuenta las condiciones reales de administración. La calidad de los materiales del test (objetos,material impreso, software) es muy buena, de comprensión sencilla y empleabilidad fácil. Las normas de aplicación e instrucciones están bien especificadas, así como lo población a la que va dirigido y la interpretación que se realiza de los resultados. Sería adecuado exponer con más detalle el proceso llevado a cabo para la creación de los ítems, además de la parte cuantitativa que los autores quieren seguir mejorando en las próximas revisiones, presentar el análisis cualitativo llevado a cabo por especialistas, que se aconseja sean diferentes a las personas que desarrollaron los ítems.

Las evidencias de validez basadas en la estructura interna son buenas, se han llevado a cabo diferentes análisis exploratorios utilizando la técnica de análisis en componentes principales: en el 2001 se obtuvo un factor (Inteligencia general) y en el último de 2016 se obtuvieron 3 factores (razonamiento, claramente definido, mientras verbal y numérico quedan menos diferenciados).

Respecto a las evidencias de validez en relación con otras variables, en el manual se aportan datos de la relación de las puntuaciones del test con un criterio relevante, el rendimiento académico. Este estudio, pese a ser modesto, obtiene resultados orientativos e interesantes. Respecto a las evidencias que en el manual se denominan de tipo convergente, debería presentarse un análisis cuantitativo para poder considerarse como tales, si no, recomendamos incluir esta evidencia en el apartado de evidencias de contenido o proceso de respuesta. También podría llevarse a cabo un estudio para aportar evidencias de tipo discriminante, una opción sería un estudio para diferenciar el constructo que evalúa el test y que los autores denominan "Actividades mentales" de otros constructos en el ámbito de la medida de la inteligencia. Por otra parte, en el MTTC se presenta un estudio para valorar las diferencias en función del género.

En los análisis realizados para el estudio de la fiabilidad, se observa un importante incremento en el tamaño de las muestras respecto al primer análisis realizado en el año 2001. Los resultados obtenidos en el último análisis de consistencia interna son buenos y similares a los anteriores, pero más completos. En global, podemos valorar los resultados como buenos al presentar valores medios comprendidos entre ,80 y ,85. Para el futuro, proponemos la posibilidad de estudiar la fiabilidad

basada en la teoría de respuesta a los ítems, y destacamos que el test ya cuenta con dos características favorables para tal propósito, como son disponer de los ítems ordenados por dificultad y la capacidad de recoger datos de muestras muy grandes. Por otro lado, los autores argumentan el no uso del test-retest como índice de estabilidad en la medida, ya que la inteligencia es un constructo dinámico, pero sin proponer ninguna alternativa para aumentar las evidencias de fiabilidad, como podría ser un estudio de medidas paralelas o equivalentes. En este sentido, realizar una prueba de dos mitades es un modo clásico y válido para aportar interesantes evidencias de fiabilidad al tratarse de dos pruebas completas aunque reducidas, pero sigue siendo evidencia basada en las interrelaciones entre los ítems del test.

En la presente revisión se ha realizado una actualización de los baremos incrementando el número de participantes tanto de 1º como de 2º de Enseñanza Primaria, también ha aumentado la variabilidad de la procedencia (número de comunidades) y se ha equilibrado por tipo de centro (público, privado y concertado). Este trabajo ha permitido aumentar la representatividad respecto al informe anterior aunque sigue habiendo desequilibrios entre las muestra de los dos cursos y la media hace que se considere una muestra grande, aunque la muestra de 1º de Educación Primaria es la mitad que la de 2º de Educación Primaria.

Es un factor a resaltar el hecho de que propongan usar acomodaciones in situ para la correcta evaluación de personas con limitaciones, lo cual favorece la equidad y tiene en cuenta las posibles diferencias étnicas y/o culturales. Aun así, estas deberían estar especificadas previamente para que no recaiga toda la responsabilidad en el psicólogo o la psicóloga que administra el test.

ANÁLISIS DETALLADO DE LA PRUEBA

1. DESCRIPCIÓN GENERAL DEL TEST

1.1. Nombre del test:

BADyG/E1-r: Batería de Actividades mentales Diferenciales y Generales, Nivel E1, renovado

1.2. Nombre del test en su versión original:

-

1.3. Autor/es del test original:

Autor principal: Carlos Yuste Hernanz y David Yuste Peña

1.4. Autor/es de la adaptación española:

-

1.5. Editor del test en su versión original:

CEPE (Ciencias de la Educación Preescolar y Especial)

1.6. Editor de la adaptación española:

-

1.7. Fecha de publicación del test original:

1980, renovación 2001

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

-

1.9. Fecha de la última revisión del test:

2019

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

Inteligencia, habilidades

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

El test consta de 6 pruebas específicas: analogías verbales (Rv), series numéricas (Rn), matrices lógicas (Re), completar oraciones (Vv), problemas numéricos (Nn) y encajar figuras (Ge).

Estas pruebas específicas se agrupan en 4 estimaciones globales: razonamiento lógico RR (Rv, Rn, Re), factor verbal VV (Rv, Vv), factor numérico NN (Rn, Nn) y factor visioespacial EE (Re, Ge).

Estas estimaciones globales, a excepción de razonamiento lógico, dan lugar a la estimación de Inteligencia General (IG). A partir de la cual se puede obtener el coeficiente intelectual (CI). Por otra parte, este test también mide la eficacia (EF).

1.12. Áreas de aplicación:

Psicología Educativa

1.13. Formato de los ítems:

Elección múltiple, respuesta abierta, respuesta dicotómica

1.14. Número de ítems:

El test costa de 174 ítems para 1º de primaria y 186 ítems para 2º de primaria.

Cada una de las 6 pruebas básicas cuenta con 18 ítems, al igual que las 3 pruebas complementarias. A excepción de una de las pruebas básicas (Memoria visoauditiva inmediata) que tiene 30 ítems y otra que varía en número de ítem en función del curso (18 ítems para 1º de E. Primaria y 30 ítems para 2º de E. Primaria).

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:

El BADyG/E1 se aplica al alumnado de 1º y 2º de primaria (6-7 años) de la población escolarizada del territorio español.

El test también puede dirigirse a grupos minoritarios pero con determinadas especificaciones. Concretamente, si va dirigido a la población:

• que habla y entiende el español, pero que no es su lengua materna, se le podrán administrar 4 pruebas básicas: Rn, Rv, Re, Ge.

• emigrante, que no domina el español, se le podrán aplicar 2 pruebas básicas: Re y Ge.

• perteneciente a ciertos grupos de integración y con alguna discapacidad funcional, será el/la psicólogo/a quien deba valorar si pueden participar en una administración colectiva.

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

Papel y lápiz.

1.19. Procedimiento de corrección:

Manual y de forma automatizado por ordenador (gestor BADyG).

1.20. Puntuaciones:

Las puntuaciones directas de las seis pruebas básicas es la suma de las respuestas correctas, a excepción de las pruebas de Calculo Numérico (Nn) y Memoria Visoauditiva Inmediata (Ma), que es restar los errores a los aciertos.

En todas las escalas el mínimo es 0, aunque haya más errores que aciertos, como en las escalas Nn y Ma. Y el máximo posible es el número de ítems de la escala. Respecto a la interpretación, todas las escalas a menor puntuación significan peor habilidad, a excepción de la puntuación de Alteraciones en la Escritura (Ae) que tiene un significado positivo, ya que indica ausencia de alteraciones en la escritura.

A partir de las puntuaciones directas de las pruebas básicas se pueden calcular las puntuaciones de los factores globales. Y, a partir de los factores globales, a excepción del razonamiento lógico, la Inteligencia General, y gracias a ésta, la puntuación del CI.

La puntuación de eficacia se el porcentaje de aciertos obtenido (cociente entre las respuestas acertadas y las emitidas).

Para terminar, el cálculo de las omisiones, se realiza simplemente sumando las respuestas omitidas en las seis pruebas básicas.

1.21. Escalas utilizadas:

Centiles, Eneatipos y Puntuaciones estandarizadas CI de desviación (Media 100 y desviación típica 15).

1.22. Posibilidad de obtener informes automatizados:

Si existe la posibilidad de obtener informes automatizados, mediante la web www.gestorbadyg.com con la cual podemos obtener seis tipos de informes.

El colectivo muestra el listado completo de participantes de un grupo con todas las puntuaciones directas y percentiles obtenidas en las distintas pruebas. También se muestra la media aritmética y la desviación típica, tanto del curso como del grupo de baremación. En los listados aparecen abreviadas todas las escalas evaluadas y abajo se clarifica el nombre completo de cada escala. Este informe es útil para el/la psicólogo/a (principalmente) y para el/la profesor/a del grupo porque permite tener una visión global de todo el grupo.

El individual contiene la puntuación directa, percentiles y una representación gráfica para cada una de las pruebas. Este informe es útil para la persona evaluada y sus responsables legales, porque permite valorar los puntos fuertes y débiles de su actividad mental, de su inteligencia actual.

El descriptivo muestra una valoración textual de tres aspectos: las estimaciones globales, las pruebas aptitudinales básicas y las pruebas complementarias, añadiéndose por último, una apreciación pedagógica global de las pruebas básicas. Presenta la misma utilidad que el informe individual.

El comparativo de medias de las puntuaciones directas, tanto del baremo como de los distintos grupos del curso seleccionado. Cuando la media del grupo se encuentra por encima aparece subrayada por una línea verde y cuando queda por debajo aparece subrayada por una línea roja. Este informe es útil para el psicólogo y los profesores, ya que les permite identificar puntos fuertes y débiles de cada grupo y poder decidir futuras actuaciones a partir del resultado obtenido.

También está el informe de distribución donde se presenta la distribución del grupo entre los 7 niveles de aptitud establecidos. Los niveles van desde muy alta hasta muy bajo. Además se indica con una línea roja en qué aptitudes la media del curso se encuentran por debajo de la media del baremo. Este informe es útil para el/la psicólogo/a y, especialmente, para el profesorado de asignaturas que incluyan actividades visoperceptivas, ya que proporcionan información sobre el ritmo que marcar durante el curso, en función del grupo y la materia que se imparte.

Finalmente encontramos el informe de verdaderos valores de IG y CI en el que podemos ver los límites superiores e inferiores entre los que se encuentran estas puntuaciones para todo el alumnado.

En general podemos observar que los informes son bastante claros y fáciles de entender, ya que nos dan la información de los baremos directamente y estos tienen una interpretación sencilla. Además algunos informes complementan la información aportando datos gráficos o incluyendo colores que facilitan la interpretación.

1.23. Tiempo estimado para la aplicación del test:

Aplicación individual o colectiva: 1h y 15' (pruebas básicas) y 1h y 45' (pruebas complementarias).

1.24. Documentación aportada por el editor:

Manual impreso, información técnica complementaria (Manual Técnico completo a través del Gestor BADyG), cuadernos de elementos-preguntas y cuadernos de respuestas.

1.25. Precio de un juego completo de la prueba:

El material se distribuye suelto. Manual técnico Badyg E1: 41,10€. Licencia Badyg E1 con 60 usos: 48,00€ (2/11/2019).

1.26. Precio y número de ejemplares del paquete de cuadernillo y hojas complementarias:

Cuaderno de preguntas-respuestas: 2,71€ c/u. Hojas de presentación de memoria no fungible: 3,80€ paquete de 10 unidades.

1.27. Precio de la corrección online:

Posibilidad de corrección online de la prueba, utilizando 1 uso por cada corrección que incluye la elaboración de informes (0,80€). Hay renovaciones de usos: 30 usos 19,50€ y 300 usos 149,00€ El editor ofrece la posibilidad de corrección por su parte.

*** FE DE ERRORES EN PÁGINA 19**

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	Buena	4
2.3. Fundamentación teórica	Excelente	5
2.4. Adaptación del test	-	-
2.5. Desarrollo de los ítems del test	Buena	4
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	4

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	Buena - Excelente	4,5
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 entre las puntuaciones del test y otras variables:

Contenido	Valoración	Puntuación
2.11.2.1.1. Diseños empleados	Diferencias entre grupos	
2.11.2.1.2. Tamaño de las muestras	Un estudio con una muestra grande (N> 500)	3
2.11.2.1.3. Procedimiento de selección de las muestras	No se aporta información en la documentación	
2.11.2.1.4. Calidad de los tests empleados como criterio o marcador	-	-
2.11.2.1.5. Promedio de las correlaciones con otros tests que miden constructos similares	-	-
2.11.2.1.6. Promedio de las correlaciones con otros tests que miden constructos no relacionados	-	-
2.11.2.1.7. Resultados de la matriz multirrasgo-multimétodo	-	-
2.11.2.1.8. Resultados de las diferencias intergrupo	Bueno	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	Rendimiento académico en lengua, matemáticas e inglés.	
2.11.2.2.2. Calidad de los criterios empleados	Buena	4
2.11.2.2.3. Relación temporal entre test y criterio	Predictivo	
2.11.2.2.4. Tamaño de las muestras	Un estudio con una muestra moderada (100 ≤ N <200)	2
2.11.2.2.5. Procedimiento de selección de las muestras	No se aporta información en la documentación	
2.11.2.2.6. Promedio de las correlaciones del test con los criterios	Buena (,45 ≤ r < ,55)	4

2.11.3. Evidencias de validez basadas en la estructura interna:

Contenido	Valoración	Puntuación
2.11.3.1. Resultados del análisis factorial	Adecuada	3
2.11.3.2. Funcionamiento diferencial de los ítems	-	-

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:

Si existe la posibilidad de obtener informes automatizados, mediante la web www.gestorbadyg.com con la cual podemos obtener seis tipos de informes.

El colectivo muestra el listado completo de participantes de un grupo con todas las puntuaciones directas y percentiles obtenidas en las distintas pruebas. También se muestra la media aritmética y la desviación típica, tanto del curso como del grupo de baremación. En los listados aparecen abreviadas todas las escalas evaluadas y abajo se clarifica el nombre completo de cada escala. Este informe es útil para el/la psicólogo/a (principalmente) y para el/la profesor/a del grupo porque permite tener una visión global de todo el grupo.

El individual contiene la puntuación directa, percentiles y una representación gráfica para cada una de las pruebas. Este informe es útil para la persona evaluada y sus responsables legales, porque permite valorar los puntos fuertes y débiles de su actividad mental, de su inteligencia actual.

El descriptivo muestra una valoración textual de tres aspectos: las estimaciones globales, las pruebas aptitudinales básicas y las pruebas complementarias, añadiéndose por último, una apreciación pedagógica global de las pruebas básicas. Presenta la misma utilidad que el informe individual.

El comparativo de medias de las puntuaciones directas, tanto del baremo como de los distintos grupos del curso seleccionado. Cuando la media del grupo se encuentra por encima aparece subrayada por una línea verde y cuando queda por debajo aparece subrayada por una línea roja. Este informe es útil para el psicólogo y los profesores, ya que les permite identificar puntos fuertes y débiles de cada grupo y poder decidir futuras actuaciones a partir del resultado obtenido.

También está el informe de distribución donde se presenta la distribución del grupo entre los 7 niveles de aptitud establecidos. Los niveles van desde muy alta hasta muy bajo. Además se indica con una línea roja en qué aptitudes la media del curso se encuentran por debajo de la media del baremo. Este informe es útil para el/la psicólogo/a y, especialmente, para el profesorado de asignaturas que incluyan actividades visoperceptivas, ya que proporcionan información sobre el ritmo que marcar durante el curso, en función del grupo y la materia que se imparte.

Finalmente encontramos el informe de verdaderos valores de IG y CI en el que podemos ver los límites superiores e inferiores entre los que se encuentran estas puntuaciones para todo el alumnado.

En general podemos observar que los informes son bastante claros y fáciles de entender, ya que nos dan la información de los baremos directamente y estos tienen una interpretación sencilla. Además algunos informes complementan la información aportando datos gráficos o incluyendo colores que facilitan la interpretación.

2.12. Fiabilidad:

Contenido	Valoración	Puntuación
2.12.1. Datos aportados sobre fiabilidad	Varios coeficientes de fiabilidad (para cada escala o 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 de tamaño moderado o com muestras grandes y otras pequeñas	4
2.12.3.2. Coeficientes de consistencia interna presentados	Coeficiente alfa o KR-20	
2.12.3.3. Promedio de coeficientes de consistencia	Buena (,80 ≤ r < ,85)	4

2.12.4. Estabilidad (test-retest):

Contenido	Valoración	Puntuación
2.12.4.1. Tamaño de las muestras	-	-
2.12.4.2. Coeficientes de estabilidad	-	-

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	-	-

2.12.7. Comentarios generales sobre evidencias de fiabilidad:

El estudio de la fiabilidad del test se ha llevado a cabo a través de la consistencia interna, mediante el coeficiente alfa de Cronbach y del procedimiento de las dos mitades (coeficiente de Spearman-Brown y de Guttman). Se obtuvieron datos de cada una de las pruebas, así como del factor general y los 4 factores globales. El estudio se llevó a cabo con muestras de tamaño grande (612 alumnos/as de 1º de Enseñanza Primaria y 1143 alumnos/as de 2º de Enseñanza Primaria).

En el estudio del 2014 se obtuvieron valores muy elevados en IG (,92 - ,94) y RR (,87 - ,89) para todo el alumnado participante (1º y 2º respectivamente). Por lo que respecta a los coeficientes de fiabilidad de los factores globales los valores oscilaban entre ,72 y ,91, siendo los mas bajos los del factor verbal y los más altos los del factor numérico. Por último, los coeficientes de consistencia interna de las pruebas específicas oscilaron entre ,62 y ,89, también se presenta al igual que en los globales, lo valores más bajos para las órdenes verbales complejas y los más altos para cálculo numérico. También se presentan los datos del análisis del 2001, pero la muestra era inferior y no hay datos de todos los factores globales, pero en general los valores son similares. En global, podemos valorar los resultados como buenos al presentar valores medios comprendidos entre ,80 y ,85. Con la finalidad de seguir mejorando el test, sería muy interesante en el futuro estudiar la fiabilidad de las diferentes combinaciones que recomienda el test para dirigirse a grupos minoritarios.

Siguiendo los criterios de Nunally y Bernstein, si el test se pretende usar para selección de personal y para hacer una interpretación de las puntuaciones individuales de gran impacto, el coeficiente alfa debería de ser igual o superior a ,90. En este caso, obtenemos esos valores en las puntuaciones de inteligencia general y aunque no lleguen a estos niveles los valores de los factores globales o de las pruebas básicas, no importa, porque estos nunca se usan de modo individual si no que funcionan conjuntamente. De todos modos, la mejora de estos valores aumentaría la capacidad predictiva de este test sobre el rendimiento académico del alumnado. Para hacer esta predicción de manera adecuada se requiere que las puntuaciones sean lo más precisas y con el menor grado error posible.

Realizar una prueba de dos mitades es un modo clásico y válido de aportar interesantes evidencias de fiabilidad al tratarse de dos pruebas completas aunque reducidas; y en este estudio, quizá este método sea más adecuado que el alfa de Cronbach, ya que es difícil que este test cumpla los requisitos para aplicarla. Pero por ello, sería necesario, además, aportar evidencias de fiabilidad complementarias, mediante un diseño de medidas equivalentes o incluso mediante la Teoría de Respuesta a los Ítems, puesto que el test ya cuenta con dos factores favorables para tal propósito: disponer de los ítems ordenados por dificultad y la capacidad de recoger datos de muestras muy grandes. Todo ello sin olvidar que, según se dice en la manual, en estudios anteriores los resultados obtenidos aplicando diseños test-retest fueron similares a los resultados obtenidos con el diseño de dos mitades.

El punto a matizar en el texto del manual es la conclusión sobre estabilidad temporal (que

denominan estabilidad cronológica) basada en la estabilidad de los valores del coeficiente alfa de Cronbach obtenidos en los años 2001 y 2017. Los coeficientes de consistencia interna evalúan en qué medida hay concordancia entre las diferentes puntuaciones de las personas evaluadas en distintas partes del test. Que existan valores del coeficiente de alfa similares en dos momentos diferentes no permite asegurar de que las puntuaciones sean fiables con el paso del tiempo.

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 poblaciones	4
2.13.1.2. Tamaño de las muestras	Grande (600 ≤ N < 1000)	4
2.13.1.3. Aplicación de estrategia de tipificación continua	No	
2.12.1.4. Procedimiento de selección de las muestras	No se aporta información en la documentación	
2.12.1.5. Actualización de baremos	Excelente (menos de 10 años)	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	-	-

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

A partir de las puntuaciones obtenidas, el manual nos proporciona baremos adecuados para dar una interpretación normativa a las puntuaciones de la población escolar.

Se ha realizado una actualización de los baremos incrementando el número de participantes (tanto de 1º de Enseñanza Primaria pasando de 385 a 612 como de 2º de Enseñanza Primaria, de 446 a 1.143), ha aumentado la variabilidad de la procedencia (número de comunidades) y se ha equilibrado por tipo de centro (público, privado y concertado). Sin embargo, para asegurar una correcta representatividad de la población diana, deberían reflejarse las proporciones reales de alumnado de 1º y 2º de Enseñanza Primaria en los diferentes tipos de centros educativos (público, privado y concertado).

Por lo que se refiere a los procedimientos para facilitar la interpretación, el manual aporta baremos actualizados diferenciados para cada curso, las puntuaciones se interpretan mediante eneatipos y percentiles, facilitando la comprensión. Además, en la corrección online permite obtener el CI según la edad utilizando la fecha de nacimiento del/de la alumno/a.

Finalmente, queremos enfatizar que esta baremación tiene un trabajo muy costoso detrás. Los autores han llevado a cabo una importante actualización de los baremos, de un modo altamente operativizado, todos estos aspectos son muy positivos y determinantes a la hora de aportar credibilidad a los resultados y a su interpretación, ya que garantizan que la muestra es más representativa que en las versiones anteriores.

*** FE DE ERRORES**

Con fecha 19 de mayo de 2020 se han modificado los epígrafes 1.25, 1.26 y 1.27 incluidos en la primera versión del documento, página 10, con objeto de subsanar un error en la política de precios ajeno a la empresa editora del test.