

Evaluación del

Evaluación Conductual

de la Función Ejecutiva

∞BRIEF-2∞

RESUMEN DE LA VALORACIÓN DEL TEST

Descripción general

Característica	Descripción
Nombre del test	BRIEF-2. Evaluación Conductual de la Función Ejecutiva
Autor	Gerard A. Gioia, Peter K. Isquith, Steven C. Guy y Lauren Kenworthy
Autor de la adaptación española	M ^a Jesús Maldonado, M ^a de la Concepción Fournier, Rosario Martínez-Arias, Javier González, Juan Manuel Espejo y Pablo Santamaría (Dpto. de I+D+i de TEA Ediciones)
Variable(s)	Función ejecutiva
Áreas de aplicación	Neuropsicología, psicología clínica, psicología educativa, psicología forense, servicios sociales, investigación
Soporte	Papel y lápiz, informatizado. Corrección por internet

Valoración general

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

Comentarios generales

El cuestionario BRIEF-2 de Evaluación Conductual de la Función Ejecutiva es una prueba destinada a la evaluación de los aspectos más cotidianos, conductuales y observables de las funciones ejecutivas de niños, niñas y adolescentes de entre 5 y 18 años. El cuestionario puede ser respondido por madres, padres o familiares (BRIEF-2 Familia) y por profesorado (BRIEF-2 Escuela), quienes han de indicar la frecuencia con la que determinados comportamientos resultan problemáticos. Consta de 63 ítems que se responden con una escala de frecuencia (nunca, a veces, frecuentemente) en un tiempo de aproximadamente 10 minutos.

El BRIEF-2, editado por TEA, es la adaptación española del instrumento original estadounidense con el mismo nombre (Gioia et al., 2015), siendo este a su vez una versión mejorada del BRIEF (Gioia et al., 2000). El BRIEF surgió ante la necesidad de pruebas para la valoración de aspectos del funcionamiento ejecutivo en jóvenes, que no estaban recogidos en otras pruebas existentes.

El BRIEF-2 proporciona puntuaciones en varias escalas relacionadas con funciones ejecutivas: inhibición, supervisión de sí mismo, flexibilidad, control emocional, iniciativa, memoria de trabajo, planificación y organización, supervisión de su tarea y organización de materiales. La combinación de estas da lugar a los cuatro índices de regulación conductual, emocional, cognitiva y global de función ejecutiva. De este modo, la prueba proporciona el perfil de afectación de las distintas facetas de las funciones ejecutivas, alteraciones que con frecuencia están presentes en personas con dificultades de aprendizaje, problemas de atención/hiperactividad, trastornos del espectro del autismo, trastornos del lenguaje, prematuridad, trastornos conductuales de origen orgánico y otros trastornos del desarrollo. La obtención de puntuaciones altas en cualquiera de las escalas del BRIEF-2 indica la presencia de problemas en el área que representa dicha escala. Los resultados que proporciona, interpretados junto con otra información relevante, pueden ser de utilidad para el diagnóstico y la planificación de intervenciones profesionales en los ámbitos educativo, clínico, social, forense o de investigación.

Entre los puntos fuertes del BRIEF-2 cabe destacar:

- La calidad de los materiales y documentación aportada, así como la cantidad y la calidad de las referencias utilizadas en la fundamentación teórica. Se ha de destacar que el test incluye dos manuales, uno en el que se incluyen las instrucciones para la aplicación, corrección e interpretación de las puntuaciones, y otro en el que se incluye la información más técnica (ficha técnica, proceso de creación y adaptación, fundamentación psicométrica). El separar esta información en dos tomos, más manejables e independientes, resulta de gran utilidad y practicidad. Ha de destacarse, además, que ambos manuales incluyen información abundante y pormenorizada sobre la gran mayoría de los aspectos esenciales; la información se presenta de forma atractiva clara y sencilla; y, además, se tienen en cuenta la perspectiva de género en el uso del lenguaje (i.e., lenguaje inclusivo) de una forma racional.
- La amplia muestra de tipificación con población general y los diversos estudios con muestras clínicas.
- La cantidad y diversidad de evidencias proporcionadas sobre la fiabilidad y validez de la prueba, así como la explicación asequible de los diversos conceptos abordados.

Entre las limitaciones o aspectos a mejorar destaca que la mayor parte de los análisis proceden exclusivamente de la teoría clásica de los test y que se echa en falta en el manual información más minuciosa sobre el análisis de variables latentes (AFC). Se considera un aspecto a mejorar la provisión de más detalles acerca de las normas de corrección y los baremos. Del mismo modo, sería deseable realizar estudios de invarianza en relación con el test original.

ANÁLISIS DETALLADO DE LA PRUEBA

1. DESCRIPCIÓN GENERAL DEL TEST

1.1. Nombre del test:

BRIEF-2. Evaluación Conductual de la Función Ejecutiva.

1.2. Nombre del test en su versión original:

BRIEF-2. Behavior Rating Inventory of Executive Function, Second Edition.

1.3. Autor/es del test original:

Gerard A. Gioia, Peter K. Isquith, Steven C. Guy y Lauren Kenworthy.

1.4. Autor/es de la adaptación española:

María Jesús Maldonado Belmonte, M^a de la Concepción Fournier del Castillo, Rosario Martínez Arias, Javier González Marqués. Juan Manuel Espejo-Saavedra Roca y Pablo Santamaría.

1.5. Editor del test en su versión original:

PAR. Psychological Assessment Resources.

1.6. Editor de la adaptación española:

TEA Ediciones.

1.7. Fecha de publicación del test original:

2015.

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

2017.

1.9. Fecha de la última revisión del test:

2017.

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

Función ejecutiva.

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

El BRIEF-2 (Evaluación Conductual de la Función Ejecutiva) es un cuestionario diseñado para la evaluación de las funciones ejecutivas en niños, niñas y adolescentes de entre 5 y 18 años. Las funciones ejecutivas son un conjunto de procesos responsables de dirigir, guiar y controlar las funciones cognitivas, emocionales y conductuales, especialmente en aquellos momentos en los que es necesaria una solución de problemas activa por parte de la persona. Las funciones ejecutivas pueden verse afectadas en múltiples situaciones (p. ej., trastornos del neurodesarrollo, daño cerebral adquirido, trastornos del aprendizaje, exposición a tóxicos, problemas emocionales). Puntuaciones altas indican la presencia de problemas en el área que representa la escala en la que se obtenga.

Consiste en un cuestionario de 63 ítems con un formato de respuesta con tres opciones de frecuencia (nunca, a veces, frecuentemente), que se cumplimenta en un breve periodo de tiempo (10 minutos aproximadamente). Dispone de dos versiones que comparten la misma estructura: (1) BRIEF-2 Familia, para ser respondida por familiares (padre, madre u otros) y (2) BRIEF-2 Escuela, para ser respondida por profesorado.

La administración es individual. Para el BRIEF-2 Familia es recomendable contar con la colaboración de los dos progenitores por separado para obtener ambos resultados y perfiles (si hubiere que elegir uno, se recomienda optar por aquel con mayor contacto con la persona evaluada durante los últimos seis meses). En el caso del BRIEF-2 Escuela, puede realizar la evaluación cualquier adulto con contacto prolongado o frecuente con la persona evaluada en un ambiente escolar o educativo.

Consta de nueve escalas clínicas (inhibición, supervisión de sí mismo/a, flexibilidad, control emocional, iniciativa, memoria de trabajo, planificación y organización, supervisión de la tarea, organización de materiales). La combinación de estas escalas da lugar a cuatro índices generales: índice de regulación conductual, índice de regulación emocional, índice de regulación cognitiva, índice global de función ejecutiva. Además, incluye tres escalas de validez: infrecuencia, inconsistencia y negatividad.

Escalas de validez:

- I. *Infrecuencia*: indica el grado en el que se han dado respuestas muy infrecuentes a determinados ítems (p. ej., afirmar que la persona evaluada no sabe contar hasta cuatro o tiene problemas para encontrar la puerta de casa).
- II. *Inconsistencia*: indica en qué medida se ha contestado de forma incongruente o inconsistente a ítems que son similares.
- III. *Negatividad*: mide el grado en el que se responde de un modo inusualmente negativo a determinados ítems en comparación con muestras clínicas.

Índice de regulación conductual: informa del grado de dificultad para regular y supervisar las conductas de manera efectiva. Se compone de dos escalas clínicas:

- I. *Inhibición*: evalúa la presencia de problemas para controlar impulsos, regular el comportamiento adecuadamente y frenar la conducta en el momento oportuno.
- II. *Supervisión de sí mismo*: evalúa la presencia de problemas para darse cuenta y ser consciente del impacto de la propia conducta en otras personas.

Índice de regulación emocional: informa del grado de dificultad para regular respuestas emocionales, especialmente como respuesta a situaciones cambiantes. Está compuesto por dos escalas clínicas:

- III. *Flexibilidad*: evalúa la presencia de problemas para cambiar libremente de una situación, actividad o aspecto de un problema a otro si las circunstancias así lo requieren. Dificultad para realizar transiciones, cambiar el foco atencional y solucionar problemas de forma flexible.
- IV. *Control emocional*: evalúa la presencia de problemas para regular o modular adecuadamente las respuestas emocionales.

Índice de regulación cognitiva: informa del grado de dificultad para controlar y gestionar sus procesos cognitivos y resolver problemas de manera eficaz. Está compuesto por cinco escalas clínicas:

- V. *Iniciativa*: evalúa la presencia de problemas para iniciar tareas o actividades de forma autónoma o independiente o para generar nuevas ideas, respuestas o estrategias de resolución de problemas.
- VI. *Memoria de trabajo*: evalúa la presencia de dificultades para mantener temporalmente la información en la mente a fin de conseguir completar una tarea o de mantener una actividad.
- VII. *Planificación y organización*: evalúa la presencia de problemas para anticiparse a situaciones futuras, ordenar y priorizar la información, plantear objetivos y secuenciar los pasos necesarios para lograrlos, y comprender y comunicar las ideas principales o los conceptos clave.
- VIII. *Supervisión de la tarea*: evalúa el grado en que se tienen dificultades para revisar su trabajo, valorar su ejecución durante y después de realizar la tarea, y asegurarse la consecución del objetivo.
- IX. *Organización de materiales*: evalúa la presencia de problemas para mantener de forma ordenada y organizada la zona de estudio, trabajo, juego y sus cosas (p. ej., escritorio, armario, mochila, habitación).

Índice global de función ejecutiva: puntuación resumen a partir de las nueve escalas clínicas del BRIEF-2. Resulta útil como medida resumen de la presencia de problemas en las funciones ejecutivas.

La prueba dispone de baremos en puntuaciones T de población general diferenciados por sexo, rango de edad e informantes (familia o escuela). En su tipificación se han incluido muestras de distintas procedencias, regiones geográficas y niveles socioeconómicos. La corrección es informática, a través de una clave de acceso (PIN) y aplicación por internet (TEAcorrige), que proporciona el perfil de afectación de las distintas facetas de las funciones ejecutivas. El perfil detalla las puntuaciones T obtenidas en cada una de las escalas e índices de la prueba, lo que permite comparar a la persona evaluada con otras de su misma edad y sexo con un desarrollo típico. Además, permite detectar puntos fuertes y puntos débiles dentro del mismo perfil.

Su aplicación es recomendada para la evaluación en un amplio número de contextos y con distintas finalidades en los ámbitos clínico, educativo, social, forense y de investigación. Los resultados pueden ser útiles para el diagnóstico (si bien no está diseñado para diagnosticar, aplicado de forma aislada, ningún trastorno o patología), la planificación de los objetivos de intervención y la selección de las técnicas más adecuadas para cada caso concreto. Además, está concebido para su aplicación en diversas ocasiones, permitiendo comparar los cambios a lo largo del tiempo y hacer un seguimiento.

1.12. Áreas de aplicación:

Neuropsicología, psicología clínica, psicología educativa, psicología forense, servicios sociales, investigación.

1.13. Formato de los ítems:

Respuesta graduada / tipo Likert.

1.14. Número de ítems:

La prueba consta de 63 ítems en total, pertenecientes a los siguientes factores:

BRIEF-2 Familia:

Inhibición= 8 ítems
Supervisión de sí mismo= 4 ítems
Flexibilidad= 8 ítems
Control emocional= 8 ítems
Iniciativa= 5 ítems
Memoria de trabajo= 8 ítems
Planificación y organización= 8 ítems
Supervisión de la tarea= 5 ítems
Organización de materiales= 6 ítems

Infrecuencia= 3 ítems
*Inconsistencia= 8 pares de ítems
*Negatividad= 8 ítems

BRIEF-2 Escuela:

Inhibición= 8 ítems
Supervisión de sí mismo= 5 ítems
Flexibilidad= 8 ítems
Control emocional= 8 ítems
Iniciativa= 4 ítems
Memoria de trabajo= 8 ítems
Planificación y organización= 8 ítems
Supervisión de la tarea= 6 ítems
Organización de materiales= 5 ítems

Infrecuencia= 3 ítems
*Inconsistencia= 8 pares de ítems
*Negatividad= 8 ítems

* Los ítems de las escalas inconsistencia y negatividad también forman parte de las escalas clínicas.

1.15. Soporte:

Papel y lápiz. Informatizada. Corrección por internet.

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

Nivel C.

Si bien el manual indica que la aplicación y corrección resulta clara y sencilla (pudiendo ser aplicado por auxiliares u otros profesionales que hayan sido formados en la aplicación de pruebas y que estén suficientemente familiarizados con las normas de aplicación de la prueba), se requiere que los aplicadores trabajen siempre bajo la supervisión de un profesional cualificado. La interpretación de las puntuaciones y los perfiles obtenidos debe realizarse por profesionales titulados con formación específica en el área de las funciones ejecutivas y la evaluación con tests psicológicos.

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

La prueba se puede administrar a estudiantes de entre 5 y 18 años.

La muestra de tipificación incluye población general y casos clínicos (TDAH, tempo cognitivo lento, trastornos del espectro del autismo, dificultades de aprendizaje, ansiedad, daño cerebral adquirido, epilepsia, neurofibromatosis tipo 1, leucemia linfocítica aguda, tumor cerebral, diabetes tipo 1), procedentes de distintas regiones geográficas, niveles socioeconómicos y niveles educativos, si bien los baremos solo realizan distinciones por sexo y edad.

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

El BRIEF-2 es la adaptación española de la segunda edición de la prueba original estadounidense BRIEF (Gioia et al., 2000). Los objetivos que se siguieron en la creación del BRIEF-2 fueron fundamentalmente acortar las escalas, incrementar su sensibilidad, añadir una escala de validez y mejorar el paralelismo entre las versiones para la familia y la escuela.

Existen, por tanto, dos formas paralelas del BRIEF-2: BRIEF-2 Familia y BRIEF-2 Escuela, que pueden utilizarse complementariamente o de forma independiente, con el número de informantes que se considere oportuno en cada caso y contexto de evaluación.

Ambas versiones comparten estructura (63 ítems, 9 escalas clínicas, 3 escalas de validez), formato de respuesta e instrucciones. También ofrecen los mismos índices generales (índice de regulación cognitiva, emocional, conductual y global de función ejecutiva).

Del BRIEF-2 existe versión impresa y versión informatizada.

1.19. Procedimiento de corrección:

La corrección es automatizada, mediante una plataforma de corrección on-line.

1.20. Puntuaciones:

La puntuación directa de cada subescala se obtiene sumando los resultados de los ítems que la conforman, si bien no se encuentra en el manual la forma de puntuar cada ítem en función de la respuesta proporcionada (sí se especifica la pertenencia de los ítems a cada subescala). En el manual no se aporta información sobre las expresiones de cálculo de las subescalas clínicas ni de los índices generales.

Las puntuaciones directas de cada una de las nueve subescalas clínicas y de los cuatro índices generales se transforman en puntuaciones T (con media 50 y desviación típica 10), que indican cómo se sitúa la persona evaluada en comparación con la muestra de referencia, informando de la distancia, medida en desviaciones típicas, a la que se encuentra la puntuación directa de la media obtenida por el grupo normativo.

A mayor puntuación T, mayor grado de disfunción ejecutiva. Como pauta general, para todas las escalas clínicas e índices del BRIEF-2 se establece que las puntuaciones T entre 60 y 64 son ligeramente elevadas, puntuaciones T entre 65 y 69 son potencialmente clínicas y las superiores a 70 clínicamente elevadas o significativas.

También puede encontrarse la prevalencia o frecuencia con que aparece cada uno de estos rangos de puntuaciones T tanto en muestras de población general (apéndice A) como en muestras clínicas (apéndice B).

Existen baremos diferenciados por sexo y edad. No obstante, la corrección no se puede realizar manualmente. Las respuestas dadas por los respondientes deben introducirse en la plataforma de corrección por internet (www.teacorrige.com) y el sistema envía un perfil individual con los resultados obtenidos.

El manual hace hincapié en que, como cualquier herramienta de evaluación, la información obtenida con el BRIEF-2 debe interpretarse en el contexto de todos los demás datos.

1.21. Escalas utilizadas:

T (Media 50 y desviación típica 10).

1.22. Posibilidad de obtener informes automatizados:

Sí. El informe presenta información clara sobre los resultados: puntuaciones directas para las escalas de validez (con una interpretación cualitativa que indica si el nivel de validez es normal o los resultados deben tomarse con precaución) y puntuaciones T para las nueve escalas clínicas y los cuatro índices generales. Las puntuaciones T son representadas en un perfil que, mediante colores, permite distinguir claramente cuáles son las puntuaciones ligeramente elevadas, potencialmente clínicas o clínicamente significativas.

La plataforma de corrección requiere introducir las respuestas en la pantalla presentada por el ordenador y seleccionar el baremo. La prueba se corrige de forma inmediata proporcionando un perfil que puede consultarse en pantalla o imprimirse. El proceso completo de corrección se estima en 3-4 minutos.

Si hay 13 o más respuestas en blanco o con dobles respuestas, el sistema informático no permite la corrección ni la obtención del perfil de resultados. Del mismo modo, si una escala tiene dos o más ítems omitidos, no se calcula la puntuación directa ni transformada de dicha escala, ni de los índices que se calculen a partir de ella.

1.23. Tiempo estimado para la aplicación del test:

10 minutos (aplicación individual).

1.24. Documentación aportada por el editor:

Manual de aplicación, corrección e interpretación; manual técnico; información técnica complementaria y actualizaciones; cuestionarios; clave de corrección y aplicación por Internet (PIN).

1.25. Precio de un juego completo de la prueba:

169,25 euros (manuales, kit de corrección Familia, kit corrección Escuela).

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

46,11 euros (25 ejemplares Familia o Escuela, PIN 25 usos).

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

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

46,11 euros (PIN de 25 usos para evaluaciones on-line).

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	Excelente	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	Adecuada	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	Buena	4
2.11.1.2. Consultas a expertos	Se ha consultado a un número moderado de expertos	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	Correlaciones con otros tests, diferencias entre grupos, análisis factoriales exploratorios	
2.11.2.1.2. Tamaño de las muestras	Varios estudios con muestras grandes	5
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	4
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 multirrasgo-multimétodo	No se aporta	---
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	Perfiles de puntuaciones en diferentes grupos clínicos	
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	Varios estudios con muestras de tamaño moderado	4
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	--	---

2.11.3. Evidencias de validez basadas en la estructura interna:

Contenido	Valoración	Puntuación
2.11.3.1. Resultados del análisis factorial	Adecuada-Buena	3,5
2.11.3.2. Funcionamiento diferencial de los ítems	No se aporta información	---

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	No	

2.11.5. Comentarios generales sobre evidencias de validez:

Los estudios de validez se centran en presentar evidencias basadas en el contenido, en la estructura interna de la prueba y en las relaciones con otras variables (puntuaciones con otras pruebas, grupos clínicos diagnósticos).

Las evidencias de validez basadas en el contenido se abordan brevemente en un apartado específico dentro de la sección dedicada a la validez, si bien se remite en este a información más detallada en el capítulo 1 del manual en el que se describe de forma más detallada tanto el proceso de construcción y desarrollo del BRIEF y el BRIEF-2, como el proceso seguido para realizar la adaptación española. Las evidencias de validez de contenido se sustentan en una revisión de la literatura científica y en la consulta a grupos de expertos de tamaño moderado. Se valora positivamente, además de la participación de expertos, la involucración de familias y profesorado en la creación de los ítems, si bien hubiera sido ideal incluir también en el proceso a la población diana a la que se dirige el test (i.e., niños y adolescentes). En el manual se proporcionan referencias para encontrar los resultados del análisis del acuerdo entre jueces (Gioia et al., 2000, 2015). La adaptación española del BRIEF-2 siguió las directrices internacionales para la adaptación de tests de la International Test Commission (Muñiz, Elosua y Hambleton, 2013). Los ítems de la adaptación fueron analizados en distintas etapas de revisión que iterativamente combinaban los análisis estadísticos con criterios racionales de forma análoga al proceso de construcción de la prueba original.

La mayor parte de los estudios de validez restantes se basan en la muestra de tipificación española, conformada por 3.389 protocolos (1.756 de la versión Familia y 1.633 de la versión Escuela). La muestra es amplia, incidental, procedente de las cuatro zonas españolas (norte, sur, este, oeste) y se ajusta a la distribución esperada en cuanto al tipo de centro educativo en el que estaban escolarizadas las personas evaluadas según la distribución poblacional (Ministerio de Educación, Cultura y Deporte, 2015). Ha de tenerse en cuenta que existe entre 103 y 172 personas evaluadas en cada grupo de edad (desde los 5 a los 16 años), si bien solo se han evaluado entre 46 y 70 personas con 17 o 18 años.

En cuanto a las evidencias sobre la estructura interna del test, se presentan las correlaciones ítem-total corregidas (que oscilan entre .27 y .79), las intercorrelaciones entre las escalas (que oscilan entre .26 y .84) y análisis factorial confirmatorio (AFC). Cabe subrayar que el BRIEF-2 se basa en modelos teóricos ampliamente aceptados que postulan las funciones ejecutivas como un constructo de amplio espectro que incluye varios procesos intercorrelacionados y que existen diversos estudios con análisis factoriales exploratorios sobre el BRIEF. Por ello, se considera adecuado optar por análisis factorial confirmatorio para realizar los análisis del BRIEF-2. En el manual se especifica el programa utilizado para el análisis de datos (AMOS 22) y las saturaciones factoriales estandarizadas, los errores y las intercorrelaciones entre factores para las dos versiones de la prueba. Además, se especifican los índices de ajuste del modelo de tres factores (regulación conductual, regulación emocional y regulación cognitiva). Se echa en falta, sin embargo, la inclusión en el manual de información detallada acerca del AFC o referencia a los diversos artículos científicos existentes con estudios internacionales en los que se proporcionan evidencias acerca de la estructura interna de la prueba. Así, por ejemplo, en el manual solo se estima el ajuste de un modelo cuando, para cumplir su función, el AFC requiere del contraste de distintas estructuras. Tampoco se justifica apropiadamente la liberación de correlaciones entre residuales, la necesidad de parcelar los ítems ni si estos cumplen el requisito de unidimensionalidad. Finalmente, dado que se proporciona un índice global de función ejecutiva, sería deseable demostrar la validez de esta puntuación general mediante, por ejemplo, modelos bifactor.

En lo que respecta a las evidencias sobre la relación con otras variables, se proporcionan datos acerca de la validez convergente y discriminante del test, mediante varios estudios con diversas muestras en los que se correlacionan las puntuaciones del BRIEF-2 con otros instrumentos o variables que evalúan constructos similares o próximos (CBCL, BASC-2, Conners 3, ADHD-RS-IV) mientras se hipotetizan correlaciones menores con constructos no relacionados como la ansiedad o las quejas somáticas. Para explorar estas relaciones se llevan a cabo estudios correlacionales y análisis factoriales exploratorios (AFE). Como limitaciones, podría apuntarse que en las matrices de cargas de los AFE se omiten las cargas menores a .30, impidiendo al lector evaluar adecuadamente la validez discriminante del parcel y que se observa cierta redundancia en los constructos subyacentes a las medidas criterio (CBCL, Conners, ADHD-IV), considerándose más deseable evaluar la validez concurrente de las mediante tests que midieran directamente aspectos relacionados con la función ejecutiva (p. ej., atención sostenida, planificación a corto plazo, impulsividad).

2.12. Fiabilidad:

Contenido	Valoración	Puntuación
2.12.1. Datos aportados sobre fiabilidad	Coeficientes de fiabilidad para diferentes grupos de personas	

2.12.2. Equivalencia formas paralelas:

Contenido	Valoración	Puntuación
2.12.2.1. Tamaño de las muestras	Buena	4
2.12.2.2. Puesta a prueba de los supuestos de paralelismo	Buena	4
2.12.2.3. Promedio de coeficientes de equivalencia	Excelente	5

2.12.3. Consistencia interna:

Contenido	Valoración	Puntuación
2.12.3.1. Tamaño de las muestras	Excelente	5
2.12.3.2. Coeficientes de consistencia interna presentados	Coeficiente alfa o KR-20	
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	Varios estudios con muestras de tamaño moderado	4
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	---	---
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:

Se proporcionan varios tipos de evidencias de fiabilidad: consistencia interna (alfa de Cronbach), consistencia temporal o estabilidad test-retest, acuerdo entre evaluadores (fiabilidad interjueces) y equivalencia con el BRIEF.

- Coeficientes alfa de Cronbach: se presentan los coeficientes obtenidos en varios estudios con muestras grandes con un valor promedio de .86, ponderando el promedio por el tamaño muestral de cada estudio y empleando la transformación propuesta por Hakstian y Whalen (1976). Los valores obtenidos para las escalas clínicas de las diferentes formas estuvieron comprendidos entre .74 (iniciativa) y .91 (control emocional) en el caso de la versión Escuela y entre .66 (iniciativa) y .87 (control emocional) en la versión Familia. En las muestras clínicas, los valores fueron más elevados, con valores entre .73 y .93. Todos los valores son mayores de .90 en las muestras clínicas y mayores de .85 en las muestras de población general. Los valores de consistencia interna obtenidos pueden considerarse, por tanto, excelentes. En futuras ediciones, sería deseable la inclusión adicional del cálculo de otros coeficientes de consistencia interna como alfa ordinal u omega ajustados para datos ordinales.
- Fiabilidad test-retest: se calculan coeficientes de correlación entre aplicaciones realizadas con un intervalo temporal de tres semanas con varias muestras de población general de tamaño moderado, extraídas de las muestras de la adaptación española y de la muestra de tipificación estadounidense (Gioia et al., 2015). El valor promedio de los coeficientes de estabilidad es .91, lo que se puede considerar excelente. Complementariamente, se analiza la diferencia en puntuaciones T entre los resultados obtenidos en ambos momentos temporales, encontrándose una magnitud general de cambios pequeña (entre 1.69 y 3.48 puntos).
- Acuerdo entre BRIEF-2 Familia y BRIEF-2 Escuela: las correlaciones entre la versión Familia y la versión Escuela del BRIEF-2 oscilaron entre .40 y .49. Los coeficientes observados por Gioia et al. (2015) entre progenitores oscilaron entre .59 y .70, mientras entre profesorado fueron algo más bajos (.39-.56).
- Equivalencia con el BRIEF: la correlación promedio entre ambas pruebas fue de .93 para la versión Familia y de .92 para la versión Escuela. Para todas las escalas clínicas e índices de ambas versiones, se obtienen coeficientes excelentes (superior a .80), excepto en el caso de supervisión de la tarea /supervisión con muestras clínicas (.71).

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	Moderado	3
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	---	---
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:

La obtención de puntuaciones altas en cualquier de las escalas del BRIEF-2 indica la presencia de problemas en el área que representa dicha escala.

El BRIEF-2 incluye un manual específico dedicado básicamente a proporcionar indicaciones sobre la aplicación y corrección del instrumento, en el que se añade además una guía para la interpretación clínica junto con algunos estudios de caso para comprender los resultados que se obtienen con la prueba. El sistema de corrección informático proporciona los resultados y el perfil del test. Para ello se introducen las respuestas en la pantalla presentada por el ordenador, se selecciona el baremo y la prueba se corrige de forma automática en un tiempo medio de 3-4 minutos.

El primer paso para la corrección implica examinar si se han respondido todos los ítems. En el caso de que existan respuestas en blanco o dobles respuestas, el sistema de corrección informático asigna automáticamente el valor promedio de las respuestas proporcionadas por ese informante en el resto de los ítems de esa misma escala. No es posible la corrección del BRIEF-2 ni se puede obtener el perfil de resultados si hay 13 o más respuestas omitidas en total, ni se pueden obtener un índice si existen 2 o más ítems omitidos en la escala utilizada para calcularlo.

La prueba dispone de baremos en puntuaciones T ($M=50$; $DT=10$) de población general diferenciados por sexo, rango de edad e informante (familia o escuela). Aunque se proponen baremos distintos en función de estas tres variables, no se justifica esta decisión mediante análisis multigrupo. Se asume que la medida es invariante y, en consecuencia, no se escogió Los baremos se obtienen a partir de una muestra de 3.389 participantes (1.756 protocolos del BRIEF-2 Familia y 1.633 del BRIEF-2 Escuela). En 1.436 casos, la misma persona fue evaluada con ambas versiones del BRIEF-2 (en 320 casos solo de dispuso del protocolo de familiares y en 197 solo del protocolo del profesorado). Los protocolos procedían de las cuatro zonas españolas (norte=25,2%; centro=32,3%; sur=22,3%; este=20,2%). El 71 estudiaba en centros públicos, el 23% en concertados y el 6% en centros privados. El 50,2% de las personas evaluadas eran mujeres. En cuanto a la edad, el grupo más numeroso de personas evaluadas fue el conformado por niños y niñas de 6 años, mientras los de tamaño más reducido fueron los conformados por personas con 17-18 años.

Los baremos no se incluyen en el manual, ni tampoco se especifica en este cuáles son los rangos de edad utilizados. Esta información sí se localiza al utilizar el sistema de corrección informático, en el que se puede elegir entre baremos para los rangos de 5-7 años, 8-10 años, 11-13 años y 14-18 años. De este modo, los tamaños muestrales para los distintos grupos de edad incluidos en los baremos oscilan entre 350 y 480 casos, aproximadamente (i.e., tamaño moderado).

Se proporciona al lector información muy detallada acerca de cómo interpretar los resultados. De este modo, se valora muy positivamente la secuencia de interpretación en la que se especifica los pasos a seguir para estructurar y facilitar la interpretación. Se incluyen, además, indicaciones acerca de cómo interpretar la puntuación relacionada con la validez del protocolo y cómo interpretar las puntuaciones T en las escalas clínicas e índices del BRIEF-2. Como pauta general, las puntuaciones T entre 60-64 se consideran ligeramente elevadas, las puntuaciones T entre 65 y 69 se consideran potencialmente clínicas y las puntuaciones iguales o superiores a 70 clínicamente significativas (parece utilizarse así, como criterio general, el de desviación típica por encima de la media). No obstante, dado que se han realizado estudios con diversas muestras clínicas, hubiera sido deseable derivar de estos, de forma empírica, puntos de corte precisos (en lugar de establecer un punto de corte común, en cierta medida, arbitrario).

Otro aspecto destacable es que también puede analizarse la prevalencia o frecuencia con que aparece cada rango de puntuaciones T (mayor o igual a 60, mayor o igual a 65, mayor o igual a 70) tanto en las muestras de población general (apéndice A) como en las muestras con determinados trastornos o patologías (apéndice B).

Además, se incluye información acerca de cómo interpretar el índice de cambio fiable, gracias al cual puede valorarse el grado en el que el cambio observado en las puntuaciones de una persona evaluada en dos momentos temporales va más allá de las esperables oscilaciones debidas al error de medida del instrumento o al efecto de la práctica.

Finalmente, se valora muy positivamente el énfasis que se hace en el manual acerca de que no se trate el BRIEF-2 de forma aislada como prueba diagnóstica, sino que han de interpretarse los resultados en el contexto de otros datos.