

Evaluación de

Test de Inteligencia No Verbal - Revisado

∞ Factor g-R ∞

RESUMEN DE LA VALORACIÓN DEL TEST

Descripción general

Característica	Descripción
Nombre del test	Factor g-R. Test de Inteligencia No Verbal - Revisado
Autor	Raymond B. Catell, Alberta Karen S. Catell y Rudolf H. Weiss, con la colaboración de Bernhard Weiss
Autor de la adaptación española	Dpto. de I+D+i TEA Ediciones
Variable(s)	Factor g o factor general de inteligencia
Áreas de aplicación	Ámbito educativo, clínico, forense y de los recursos humanos
Soporte	Papel y lápiz. Corrección por internet y mecanizada

Valoración general

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

Comentarios generales

El Factor-g-R, Test de Inteligencia No Verbal-Revisado, es la versión más actualizada de los tradicionalmente conocidos como Tests de Factor g. Más concretamente, es una adaptación de la versión alemana CFT 20-R que, a su vez, surgió como una mejora de la última versión y revisión técnica de la Escala 2 de la prueba original de Cattell: el Culture Fair Intelligence Test (Cattell y Cattell, 1960; Cattell, Krug y Barton, 1973). El Factor-G-R engloba y sustituye a la última versión española del Factor g, Escalas 2 y 3 (Cattell y Cattell, 1973).

Responde a la necesidad de evaluar con precisión la inteligencia reduciendo las influencias del lenguaje, la cultura y el nivel social o educativo (i.e., libre de contaminación cultural). Permite evaluar la capacidad para resolver problemas complejos y razonar con contenido abstracto. Se trata de un cuestionario con especial utilidad en diversos contextos y su componente no verbal aporta un valor añadido a la hora de evaluar poblaciones con características específicas (p. ej., población con necesidades educativas especiales, discapacidad intelectual, altas capacidades, barreras idiomáticas). Además, constituye una herramienta útil para la predicción del rendimiento escolar y laboral.

Con casi 45 años de recorrido, se trata de una prueba de reconocido prestigio en la medición de la inteligencia como factor g, sencilla de aplicar, con buenos materiales gráficos, un manual asequible y adaptada a las competencias del profesional de la psicología. La información que ofrece el manual resulta de gran ayuda para la correcta aplicación de la prueba y la interpretación de los resultados, de acuerdo con un modelo teórico de gran relevancia en la investigación psicológica. De hecho, incorpora una serie de ventajas con respecto a versiones previas que permiten concluir que constituye una buena medida del factor de inteligencia fluida, con suficientes evidencias de validez y fiabilidad para su aplicación en población española dentro de un rango de edad especialmente amplio (a partir de los 8 años).

Algunos de los aspectos principales a destacar como puntos fuertes de la prueba son:

- Se basa en una teoría sólidamente fundamentada tanto desde el punto de vista teórico como empírico.
- Los materiales de los que consta la prueba (manual, cuadernillos y hojas de respuesta) presentan un atractivo diseño gráfico (i.e., forma, color, alineación, contraste y espacio). La sencillez en la redacción del manual y las instrucciones facilitan notablemente su aplicación e interpretación.
- Es fácil de aplicar, pues está compuesta por cuatro escalas breves (15, 15, 15 y 11 ítems) con un formato de respuesta múltiple (cinco alternativas). De hecho, se ha simplificado su estructura, al unificar las escalas 2 y 3 de las anteriores versiones en una única forma válida para todas las edades y niveles culturales. Se facilita así la administración del test, permitiendo un considerable ahorro de tiempo en comparación con sus predecesoras. El tiempo de aplicación, incluyendo instrucciones y ejemplos, es muy breve (en torno a 40 minutos en total).
- Incrementa el número de ítems en todos los tests, mejorando la capacidad de discriminación y disminuyendo el posible efecto techo. Los cuatro tests que componen la prueba incluyen ítems que plantean la solución de problemas novedosos, poco habituales y que pueden ser evaluados de forma estandarizada.
- El sistema de corrección por internet proporciona, de forma automática e inmediata, informes claros, detallados, con un perfil gráfico y una breve explicación de los resultados. Las estimaciones de las puntuaciones obtenidas son así más precisas, evitando potenciales errores.
- La muestra de validación es muy amplia, dividida en escolares (estos, a su vez, divididos en cursos escolares: de 4º De Educación Primaria a 2º de Bachillerato) y adultos (agrupados por edades y como grupo general).

- Utiliza un procedimiento de tipificación continua ('continuous norming'), muy adecuado dado que la prueba cubre varios rangos de edad y niveles educativos. Se mitigan así los efectos de las irregularidades muestrales presentes en los distintos grupos considerados para la construcción de los baremos.
- El procedimiento de adaptación se puede considerar, en términos generales, riguroso, aplicando herramientas y estrategias metodológicas adecuadas y describiendo los pormenores de los procedimientos aplicados.
- Proporciona información acerca de la velocidad y potencia de los tests en cada grupo normativo. Además, se muestran los índices de dificultad, homogeneidad y discriminación, junto con los estadísticos descriptivos.
- Ofrece información sobre la consistencia interna y el error típico de medida, considerándose buenos los resultados de ambos en términos generales con respecto a la fiabilidad. Incluye información acerca de la estabilidad temporal de las puntuaciones, teniendo en cuenta un amplio periodo entre las evaluaciones.
- Las evidencias de validez basadas en la estructura interna de la escala se sustentan en un análisis factorial confirmatorio y un análisis de invarianza, que ofrecen resultados óptimos.
- Las evidencias de validez sobre la relación con otras variables aluden a la validez convergente, discriminante y la relación con distintos criterios (entre ellos, el curso escolar, la edad, el nivel educativo de los progenitores, el nivel educativo alcanzado, la finalidad de la evaluación y el rendimiento académico).
- Incluye breve información acerca de la tipificación en población mexicana con una muestra amplia, lo que permite ofrecer baremos específicos de México. En relación con esta tipificación se proporcionan evidencias adecuadas de fiabilidad en términos de consistencia (alfa de Cronbach, alfa ordinal), el error típico de medida y el error máximo para los distintos grupos de escolares y para la muestra general de adultos. En cuanto a la validez, incluye evidencias sobre la estructura interna (intercorrelaciones escalares y análisis factorial confirmatorio) y sobre la relación con criterios externos (grado escolar, rendimiento académico y edad de los adultos).

Entre los aspectos que pudieran mejorarse con respecto a la información aportada en la prueba destacan los siguientes:

- Sería conveniente especificar de forma más detallada el procedimiento de selección de los participantes de los estudios de validación provenientes de la versión original del instrumento, especialmente en lo que se refiere al acceso a las distintas muestras.
- La información relacionada con las evidencias de validez de contenido queda difusa en el manual, por lo que sería recomendable incluir un apartado específico, aunque fuera breve, con referencias con las que poder profundizar sobre esta cuestión.
- Muy relacionado con lo anterior, en el manual se subraya que se han incluido ítems nuevos que mejoran la capacidad de discriminación en el rango superior. Sería deseable especificar cuáles son los ítems nuevos y su proceso de desarrollo/selección.
- La provisión de evidencias convergentes, discriminantes y de estabilidad temporal se abordan por medio de estudios realizados con la versión original. Sería conveniente contar con evidencias basadas con la prueba adaptada en la muestra española.
- Con respecto al estudio de tipificación mexicano, existe cierta descompensación entre el grupo de escolares (con entre 812 y 1296 participantes por curso escolar) y el grupo de adultos (conformado por 360 personas no agrupadas por rangos de edad). En este sentido, aunque los tamaños muestrales son entre moderados y muy grandes, sería interesante incrementar la muestra de adultos.

ANÁLISIS DETALLADO DE LA PRUEBA

1. DESCRIPCIÓN GENERAL DEL TEST

1.1. Nombre del test:

FACTOR g-R. Test de Inteligencia No Verbal - Revisado.

1.2. Nombre del test en su versión original:

CFT20-R. Grundintelligenztest Skala 2 - Revision.

1.3. Autor/es del test original:

Raymond B. Cattell, Alberta Karen S. Cattell y Rudolf H. Weiss, con la colaboración de Bernhard Weiss.

1.4. Autor/es de la adaptación española:

Departamento de I+D+i de TEA Ediciones.

1.5. Editor del test en su versión original:

Hogrefe Verlag GmbH & Co.

1.6. Editor de la adaptación española:

TEA Ediciones.

1.7. Fecha de publicación del test original:

2006.

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:

Inteligencia.

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

El Factor g-R es un instrumento dirigido a la evaluación del factor general de inteligencia (factor g). Incluye tareas no verbales que requieren para su solución percibir la relación entre formas y figuras abstractas. Trata de reducir así, en la medida de lo posible, la influencia de variables culturales o de la experiencia de las personas.

Si bien el planteamiento está lógicamente basado en la teoría de la inteligencia de Cattell, se ha realizado un importante esfuerzo para acercar el contenido al de la teoría con mayor aceptación y apoyo en la actualidad: la teoría Cattell-Horn-Carroll (CHC; Carroll, 1993). Esta teoría formula un modelo jerárquico en tres estratos, situándose en la cúspide la capacidad general o factor g (estrato III, general), por debajo diversas aptitudes intelectuales (estrato II, amplio) y, en la base, factores específicos de cada aptitud intelectual (estrato I, específico). Entre las aptitudes intelectuales del estrato II, destacan los factores de inteligencia fluida y cristalizada (Cattell, 1943). Los ítems incluidos en el Factor g-R plantean, precisamente, problemas típicos de la inteligencia fluida, considerado tradicionalmente el mejor predictor del factor g.

La inteligencia fluida representa la capacidad para manejar abstracciones, reglas, generalizaciones y relaciones lógicas. Se evalúa a través de tareas no verbales novedosas que requieren razonamiento de tipo deductivo e inductivo, a través del establecimiento de inferencias, la clasificación, la generalización, la generación y comprobación de hipótesis, la identificación de relaciones y la transformación de la información.

El Factor g-R consta de 56 ítems, con niveles de dificultad creciente, agrupados en cuatro tests (Series, Clasificación, Matrices y Condiciones) que aprecian distintos aspectos de la inteligencia no verbal o la inteligencia fluida. Los cuatro tests han de aplicarse necesariamente para realizar una evaluación adecuada, no siendo adecuado aplicarlos de forma aislada o realizar combinaciones de ellos que no incluyan los cuatro tests.

Test 1: Series: recoge 15 ítems, formados por figuras incompletas y progresivas, que requieren seleccionar entre las cinco opciones propuestas la respuesta que continúa lógicamente la serie, respetando las reglas marcadas por los elementos previos y que deben ser averiguadas para lograr una solución correcta. El tiempo de aplicación es de 5 minutos.

Test 2: Clasificación: formado por 15 ítems, que requieren identificar, entre un conjunto de cinco figuras, aquella que es diferente o no concuerda con las características más relevantes del resto de estímulos. Su tiempo de aplicación es de 5 minutos.

Test 3: Matrices: incluye 15 ítems, cuya solución requiere completar un cuadro de dibujos, eligiendo de entre las cinco opciones de respuesta aquella figura que encaja o completa la estructura de relaciones abstractas que comparten ubicación y configuración gráfica. El tiempo de aplicación es de 4 minutos.

Test 4: Condiciones: consta de 11 ítems, que exigen la elección de la alternativa que cumple las mismas condiciones a que se atiene un cuadro o figura que se utiliza como referencia. Su tiempo de aplicación es de 11 minutos. La dificultad de esta prueba aumenta en muchos casos debido a que la lógica de respuesta es menos intuitiva que la de los tests anteriores. Se recomienda, por tanto, asegurarse de que la persona identifica y determina la lógica de solución del problema planteado.

La prueba puede ser aplicada de forma individual o colectiva, requiriendo un tiempo total de 41 minutos (que se reducen a 18 de trabajo efectivo si no se tiene en cuenta el tiempo dedicado a instrucciones y ejemplos).

Los principales cambios del Factor g-R con respecto a sus predecesoras son tres. En primer lugar, cuenta con baremos representativos de escolares con edades comprendidas entre los 8 y los 18 años (de 4º de Educación Primaria a 2º de Bachillerato), así como un baremo general y varios por grupos de edad de adultos evaluados en diferentes contextos. En segundo lugar, incluye ítems nuevos que mejoran la discriminación en el rango superior, disminuyendo el efecto techo. Finalmente, se evita la necesidad de aplicar una escala diferente en función de la edad (como sucedía con el anterior Factor g), utilizándose la misma versión para todas las edades, desde los 8 años en adelante.

Puede resultar de utilidad para profesionales de diferentes ámbitos y contextos: educación, clínico, forense y recursos humanos y Psicología del trabajo, principalmente.

1.12. Áreas de aplicación:

Psicología clínica, psicología educativa, psicología forense y psicología del trabajo y las organizaciones.

1.13. Formato de los ítems:

Respuesta dicotómica (acierto/error).

1.14. Número de ítems:

56 ítems (Series= 15; Clasificación= 15; Matrices= 15; Condiciones= 11).

1.15. Soporte:

Papel y lápiz. Corrección por internet y mecanizada.

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

Nivel B.

La aplicación puede ser llevada a cabo por personas con entrenamiento especializado en la evaluación con tests psicológicos, familiarizadas con las tareas de cada test y con destreza en el manejo de grupos numerosos (en aplicaciones colectivas). La interpretación y el uso de las puntuaciones se reserva únicamente a profesionales titulados y cualificados con conocimiento sobre teoría de tests y métodos estadísticos, como psicólogos/as, psicopedagogos/as y profesionales de disciplinas afines.

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

La prueba se puede administrar a población escolar de edades comprendidas entre los 8 y los 18 años (de 4º de Primaria a 2º de Bachillerato) y población adulta.

No se aporta información sobre su posible aplicación a poblaciones específicas o minoritarias, más allá de distinguir entre instrucciones específicas para la evaluación de grupos de escolares e instrucciones para grupos de adultos.

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

Existe una única forma de test para todas las poblaciones a las que es aplicable.

1.19. Procedimiento de corrección:

La corrección no se puede realizar manualmente. Las respuestas dadas por los respondientes deben introducirse en la plataforma de corrección on-line (TEAcorrige) y el sistema envía un perfil individual y un breve comentario sobre los resultados obtenidos.

1.20. Puntuaciones:

Las puntuaciones directas para cada escala se obtienen sumando el número de respuestas correctas en cada test, teniendo en cuenta que cada acierto suma un punto.

Las puntuaciones directas de los cuatro tests se suman para proporcionar una puntuación directa total, que es transformada de forma automatizada por internet en puntuaciones CI (M=100, DT=15) y percentiles.

Además, se especifica el intervalo de confianza de la puntuación CI (al nivel del 95%), aportando los límites superiores e inferiores de la puntuación. Dado que solo se tiene en cuenta para la obtención de estas puntuaciones la puntuación directa total de los cuatro tests, no es posible la interpretación aislada del rendimiento en cada uno de ellos.

1.21. Escalas utilizadas:

Puntuaciones estandarizadas de CI (M= 100, DT= 15) y percentiles.

1.22. Posibilidad de obtener informes automatizados:

Sí, un informe breve, con lenguaje sencillo e información adicional que facilita su interpretación. El informe incluye, en primer lugar, una descripción de la persona evaluada (datos identificativos o sociodemográficos), la persona responsable de la aplicación y el baremo seleccionado. A continuación, se indica de una forma muy visual el número de aciertos en cada uno de los cuatro tests, la puntuación directa total, la puntuación transformada en CI con su correspondiente percentil y los límites de la puntuación, teniendo en cuenta un nivel de confianza del 95%. La puntuación de CI y su intervalo de confianza son representados gráficamente en un perfil que ubica a la persona evaluada con respecto a su grupo de referencia (distinguiendo entre CI muy bajo, bajo, medio-bajo, medio, medio-alto, alto y muy alto). Finalmente, se proporciona un breve texto en el que se detallan y desglosan los resultados obtenidos, explicando sus implicaciones.

1.23. Tiempo estimado para la aplicación del test:

40 minutos, tanto en aplicación individual como colectiva.

1.24. Documentación aportada por el editor:

Manual, cuadernillos, hojas de respuestas y claves de acceso (PIN) al sistema de corrección por internet.

1.25. Precio de un juego completo de la prueba:

80,02 euros (manual, 10 cuadernillos, kit de corrección 25 usos).

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

34,85 euros (paquete de 10 cuadernillos).

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

14,45 euros (25 hojas de respuestas, PIN 25 usos).

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

14,45 (25 hojas de respuesta, PIN 25 usos).

2. VALORACIÓN DE LAS CARACTERÍSTICAS DEL TEST

2.1. Aspectos generales:

Contenido	Valoración	Puntuación
2.1. Calidad de los materiales del test	Buena-Excelente	4,5
2.2. Calidad de la documentación aportada	Buena-Excelente	4,5
2.3. Fundamentación teórica	Excelente	5
2.4. Adaptación del test	Buena	4
2.5. Desarrollo de los ítems del test	No se aporta	---
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	Buena-Excelente	4,5
2.10. Datos sobre el análisis de los ítems	Buena-Excelente	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	Buena	4
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	Correlaciones con otros tests y diferencias entre grupos	
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	Adecuada	3
2.11.2.1.5. Promedio de las correlaciones con otros tests que miden constructos similares	Adecuada	3
2.11.2.1.6. Promedio de las correlaciones con otros tests que miden constructos no relacionados	No se aporta	---
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	Curso escolar, edad, nivel educativo de los progenitores, nivel educativo alcanzado, finalidad de la evaluación y rendimiento académico	
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 grandes	4,5
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	Adecuada	3

2.11.3. Evidencias de validez basadas en la estructura interna:

Contenido	Valoración	Puntuación
2.11.3.1. Resultados del análisis factorial	Buena-Excelente	4,5
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	No	

2.11.5. Comentarios generales sobre evidencias de validez:

Se aportan evidencias de validez sobre la relación con otras variables y sobre la estructura interna de la escala. No se especifica un apartado concreto con evidencias sobre validez de contenido, si bien se remite al proceso de revisión y adaptación del instrumento y al análisis de los ítems.

Las evidencias de validez relacionadas con el contenido resumen brevemente el proceso de adaptación de la prueba: (a) aludiendo a los cambios realizados con respecto al instrumento alemán original (CFT 20-R) y la edición anterior de la adaptación española (Factor "g". Escalas 2 y 3); y (b) analizando los ítems en términos de velocidad y potencia (índice IV de Stafford, 1971) y de dificultad, homogeneidad y discriminación. De tales análisis se obtienen, en términos generales, resultados satisfactorios. Del análisis de las propiedades psicométricas de los ítems (media, desviación típica, máximo y mínimos teóricos de las puntuaciones directas de cada test y en total) se puede concluir que la prueba total permite discriminar a la gran mayoría de personas, pero los subtests Series y Matrices pueden sufrir un ligero efecto techo en cursos avanzados y en adultos.

El estudio de las propiedades psicométricas se lleva a cabo fundamentalmente con población escolar y adulta. En total fueron evaluados 2.930 alumnos con edades comprendidas entre entre 8 y 18 años ($M=12,4$; $DT=2,8$) en 23 centros repartidos en las cuatro zonas españolas (norte=29%, sur=37%, este=15% y centro=19%). Del total del alumnado evaluado se eliminaron aleatoriamente 794 registros para igualar los casos por sexo y tipo de centro (públicos=41%, concertados=52%, privados=7%). Así, la muestra de tipificación de los escolares quedó compuesta por 2.136 casos. La muestra de adultos estuvo formada por 1.940 casos evaluados con fines de investigación (28%) y de selección de personal (72%). Del total, se tomaron 1.732 registros equilibrados por sexo para la elaboración de baremos independientes por intervalos de edad (18-19 años=9%; 20-29 años=41%; 30-39 años=29%; 40-49 años=16%; 50 años o más=5%) y una muestra de 824 casos para la elaboración del baremo general. La selección, dentro de la muestra incidental, es aleatoria y ajustada a los datos del censo de población.

La validación de las puntuaciones de la versión española se apoya fundamentalmente en el estudio de la estructura interna. Los autores realizan un análisis factorial confirmatorio y un estudio de invarianza factorial. Respecto a la estructura interna del instrumento, se muestra información meticulosamente detallada sobre el procedimiento llevado a cabo, así como el estimador empleado y la interpretación de los índices de ajuste obtenidos. Toda vez que la estructura es conocida, se aplica un análisis factorial confirmatorio sobre las escalas. Adicionalmente, los autores plantean varios modelos estructurales para su comparación, aportando mayor solidez a los datos obtenidos y, alcanzando, además, un ajuste óptimo del modelo propuesto a los datos. Los resultados se complementan con un estudio de invarianza factorial utilizando el total de la muestra de validación española.

Las evidencias de validez en relación con otras variables indicadas en el manual aluden a la versión original de la prueba, no habiendo sido llevadas a cabo con la versión adaptada del instrumento. Sería conveniente, por tanto, aportar estudios con la muestra de validación española. Los tamaños muestrales pueden considerarse grandes y los valores obtenidos adecuados en el caso de las evidencias convergentes y discriminantes. Complementariamente, se aportan evidencias sobre la relación con distintos criterios mediante estudios de diferencias entre grupos y la relación de los tests con distintas variables (i.e., curso escolar, edad, nivel educativo de los progenitores, nivel educativo de los adultos, finalidad de la evaluación y rendimiento académico). Por otro lado, los análisis estadísticos aplicados, en el caso de la relación con otras variables, son una base sólida para justificar la elaboración de los baremos pertinentes para cada grupo normativo.

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	---	---
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	4
2.12.3.2. Coeficientes de consistencia interna presentados	Coeficiente alfa y alfa ordinal	
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 pequeñas	2
2.12.4.2. Coeficientes de estabilidad	Adecuada	3

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:

La fiabilidad de la puntuación total del test se aborda desde diferentes perspectivas para todos los grupos normativos (en escolares por curso académico y en adultos por intervalo de edad y la muestra general). En primer lugar, se analiza la fiabilidad entendida como consistencia interna desde un acercamiento clásico (coeficiente alfa de Cronbach) y desde otro más actual y ajustado a la naturaleza de la respuesta de los ítems (de tipo dicotómico, acierto/error). En el caso del coeficiente alfa ordinal, la fiabilidad se puede considerar excelente (entre .88 y .95). Además, el manual aporta información sobre el error típico de medida e intervalos de confianza, proporcionando así información de especial utilidad para la interpretación de las puntuaciones y toma de decisiones al proporcionar el error máximo en diversos niveles de confianza (90%, 95% y 99%).

La información sobre la estabilidad temporal de las puntuaciones se informa a través de una breve descripción de tres estudios realizados con el CFT-20 (versión original anterior al Factor g-R) con muestras pequeñas (entre 31 y 70 participantes) y dos estudios realizados con el test actual original con dos muestras Pequeñas (entre 13 y 38 estudiantes). El valor promedio de los coeficientes de estabilidad temporal puede considerarse bueno, si bien es cierto que las muestras empleadas son reducidas y se aportan escasos datos acerca de sus características y procedimiento de muestreo. Sería interesante proporcionar datos de estabilidad temporal de las puntuaciones con la adaptación y muestra de validación 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	Amplio rango de baremos en función de características relevantes	5
2.13.1.2. Tamaño de las muestras	Muy grande	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	---	---

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

El muestreo se llevó a cabo para dos grupos: muestra de escolares y muestra de adultos. Mientras para la población adulta se siguió un proceso de selección aleatoria "a posteriori" para que fuera representativa de los diferentes estratos censales estimados de población residente en España del Instituto Nacional de Estadística (no se especifica año), el procedimiento de selección en el caso de los escolares es menos riguroso y ajustado a la realidad poblacional, aunque se llevan a cabo criterios de igualación de la muestra por sexo en todos los cursos.

El manual no presenta los baremos para la interpretación de las puntuaciones porque la corrección está mecanizada y es realizada por el editor del test a través de internet. Sin embargo, el manual describe el proceso de construcción de los baremos, que se ajusta a la metodología conocida como tipificación continua. La interpretación de las puntuaciones resulta sencilla, la cual se facilita adicionalmente mediante recomendaciones mostradas en el propio manual y orientaciones aportadas en el informe de perfil obtenido en la puntuación automatizada del instrumento.

Se indican posibles hipótesis interpretativas asociadas a puntuaciones desde muy bajas a muy altas, reflejándose criterios orientativos de clasificación de las puntuaciones con puntos de corte específicos para tener en cuenta en la interpretación. Sumado a ello, el manual aporta instrumentos de evaluación complementarios útiles para obtener una visión más general y comprensiva de la persona evaluada y pueda tomar, de este modo, decisiones fundadas con un mayor rigor en el ámbito educativo, clínico, forense y de los recursos humanos.