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

# Test de Inteligencia No Verbal-4

# TONI-4

Prueba valorada en la ronda de evaluaciones del año 2019

# RESUMEN DE LA VALORACIÓN DEL TEST

## Descripción general

| Característica | Descripción |
|---|---|
| **Nombre del test** | Test de Inteligencia No verbal-4 (TONI-4) |
| **Autor** | Linda Brown, Rita J. Sherbenou, Susan k. Johnsen |
| **Autor de la adaptación española** | Belén Ruiz-Fernández |
| **Variable(s)** | Aptitudes mentales básicas, diferenciales y generales |
| **Áreas de aplicación** | Inteligencia y Aptitudes |
| **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** | Adecuada - Buena | 3,5 |
| **Validez: relación con otras variables** | Adecuada - Buena | 3,9 |
| **Validez: estructura interna** | Excelente | 5 |
| **Validez: análisis del DIF** | Excelente | 5 |
| **Fiabilidad: equivalencia** | Adecuada | 3 |
| **Fiabilidad: consistencia interna** | Adecuada - Buena | 3,5 |
| **Fiabilidad: estabilidad** | Adecuada | 3 |
| **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

El TONI-4 es un test de inteligencia no verbal que procede de sus versiones anteriores (TONI, TONI-2 y TONI-3) y es aplicable a población de entre 6 y 79 años. En el manual del TONI-4 se aportan evidencias de validez y de fiabilidad para la nueva versión, pero también se citan muchas otras fuentes de información, como artículos e informes, que aportan evidencias para las distintas versiones anteriores o para versiones en otros idiomas. Además, el test TONI, con sus sucesivas versiones, ha sido ampliamente usado y existen numerosos artículos científicos que corroboran muchas de las evidencias de validez aportadas en el manual o incluso complementan algunas de las que faltan en el manual. En el manual del test se aportan evidencias de validez basadas en el contenido del test, en el proceso de respuesta al test, en la estructura interna y en la relación con otras variables. En general las evidencias de validez cualitativas son apropiadas y ampliamente documentadas y las evidencias de validez cuantitativas se apoyan en técnicas de análisis de datos acordes con lo que se quiere medir y se obtienen índices de ajuste o coeficientes de correlación elevados. Se realiza una regresión logística para estudiar el funcionamiento diferencial de los ítems en función de sexo y nacionalidad (española/otras), obteniendo resultados no estadísticamente significativos, y se utiliza un análisis factorial confirmatorio para estudiar la invariancia respecto al sexo obteniendo buenos indicadores de ajuste (por ejemplo, CFI y TLI de ,99). En lo que se refiere a la validez basada en la relación con otras variables se aportan datos en relación con el test de Matrices publicado por la misma editorial, encontrando una correlación moderada de ,5. Cabe destacar que la muestra utilizada es pequeña y podría haberse ampliado para que los resultados fueran más robustos. Para las evidencias relacionadas con la fiabilidad, las muestras aportadas son pequeñas (< 200 personas). En general los indicadores de fiabilidad son muy buenos. Sin embargo, la correlación de sus formas paralelas está por debajo de ,8, un resultado inesperado que debería comentarse en el manual. Por lo que se refiere a la muestra de tipificación, el tamaño es muy grande (N = 2.820). Además, aunque no es una muestra seleccionada de forma aleatoria, se ha hecho un esfuerzo para controlar los posibles sesgos intentando obtener la máxima representación de la población. Para asegurar la máxima representatividad, se seleccionaron distintas variables de estratificación (región geográfica, tamaño de la localidad, sexo, edad, nacionalidad, titularidad del colegio y nivel de estudios). La composición de la muestra resultante es parecida a la de población española tanto en el grupo de niños y adolescentes como en el grupo de adultos. En el manual se justifica el amplio rango de edad que abarcan los baremos con el uso de modelos matemáticos basados en la tipificación continua sobre todo para eliminar las discontinuidades debidas a la falta de muestra en algunos tramos de las puntuaciones. Se sugiere que en futuras ediciones del manual se asocie a cada uno de los usos concretos para los que se promueve el test aquellas evidencias de calidad psicométrica que los avalan.

# ANÁLISIS DETALLADO DE LA PRUEBA

## 1. DESCRIPCIÓN GENERAL DEL TEST

**1.1. Nombre del test:**

Test de inteligencia no verbal-4 (TONI-4)

**1.2. Nombre del test en su versión original:**

Test of nonverbal intelligence - Fourth Edition (TONI-4)

**1.3. Autor/es del test original:**

Linda Brown, Rita J. Sherbenou, Susan k. Johansen

**1.4. Autor/es de la adaptación española:**

Belén Ruiz-Fernández

**1.5. Editor del test en su versión original:**

PRO-ED

**1.6. Editor de la adaptación española:**

Departamento de I+D+i TEA Ediciones

**1.7. Fecha de publicación del test original:**

2010

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

2019

**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 y aptitudes

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

El TONI-4, Test de Inteligencia No Verbal – 4, es un instrumento de aplicación individual que evalúa la inteligencia general mediante el razonamiento abstracto y la resolución de problemas libres de factores lingüísticos, motores o culturales, ya que puede ser aplicado de forma verbal y no verbal. Es por ello que el TONI-4 es muy útil en personas que previamente han presentado dificultades a la hora de responder test de inteligencia tradicionales.

El TONI-4 está formado por una única escala que evalúa la inteligencia global, ya que el objetivo que se pretendía conseguir con la construcción de este test era precisamente la obtención de un instrumento que evaluara la inteligencia global y no una medida exhaustiva de los distintos componentes de la inteligencia.

Según el manual, además, sirve para identificar a personas con una posible discapacidad intelectual; comparar los motivos de derivación de otros profesionales para obtener un tratamiento, terapia o acceso a servicios especiales; formular hipótesis que sirvan de guía para una intervención u otras evaluaciones y, finalmente, también se puede utilizar en la selección de personal o para realizar investigaciones.

## 1.12. Áreas de aplicación:

Psicología clínica y Psicologia educativa

## 1.13. Formato de los ítems:

Elección múltiple

## 1.14. Número de ítems:

El test formado por una única escala con 60 ítems

## 1.15. Soporte:

Administración en papel y lápiz,.

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

Nivel C

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

Rango de edad: entre 6 y 79 años

Aplicable a población general

Aplicable a personas con dificultades verbales, auditivas y/o motoras así como aquéllas que no están familiarizadas con la cultura occidental

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

El test se divide en dos formas paralelas (forma A y forma B). Cada una de ellas está compuesta por 60 ítems ordenados de menor a mayor dificultad y con sus respectivos baremos.

## 1.19. Procedimiento de corrección:

Manual y de forma automatizado por ordenador.

## 1.20. Puntuaciones:

El procedimiento para obtener las puntuaciones es informatizado. Este sistema tiene menos errores y hace que la corrección pueda realizarse de manera más ágil. La persona que haya aplicado el test debe introducir los datos en la página web TEAcorrige. Para esto, debe asegurarse de haber establecido una línea base (establecer una línea de conocimiento del test en la que se deben mostrar cinco aciertos consecutivos) o la regla de parada (si la persona evaluada tiene tres errores en cinco pregunta, se da por finalizado el test). A continuación, se registran todas las respuestas, sin contar las de prueba, y sin espacios en blanco entre respuestas. De lo contrario, la corrección del test no podría realizarse. Como es un método de corrección más rápido, se obtienen resultados inmediatos. Aunque ofreciéndose una corrección automática no es imprescindible publicar el algoritmo de cálculo, clarificar la información aproximada sobre cómo se combinan las respuestas a los ítems para obtener la puntuación total ayudaría a interpretar los resultados.

## 1.21. Escalas utilizadas:

Centiles, CI (Media 100 y desviación típica 15), edad mental.

## 1.22. Posibilidad de obtener informes automatizados:

Se da un informe automatizado que contiene una breve interpretación de los resultados.

## 1.23. Tiempo estimado para la aplicación del test:

15 minutos.

## 1.24. Documentación aportada por el editor:

Manual impreso, cuadernos de elementos-preguntas, hojas de anotación y clave acceso PIN para la corrección por internet.

## 1.25. Precio de un juego completo de la prueba:

Juego completo (Manual, Kit de corrección 25 usos, Cuaderno de láminas). 83,20€ (27/01/2020)

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

21,15€ con 25 cuadernillos, PIN 25 usos (27/01/2020)

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

21,15€ con 25 cuadernillos, PIN 25 usos (27/01/2020)

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

21,15€ con 25 cuadernillos, PIN 25 usos (27/01/2020)

## 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** | Buena | 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** | Excelente | 5 |
| **2.11.1.2. Consultas a expertos** | Adecuado, con algunas carencias | 2 |

### 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 | |
| **2.11.2.1.2. Tamaño de las muestras** | Varios estudios con muestras moderadas (200 < N < 500) o pequeñas (N < 200) | 2 |
| **2.11.2.1.3. Procedimiento de selección de las muestras** | 76 personas de 9 a 38 años procedentes de diferentes puntos del territorio español, y muestra de tipificación | |
| **2.11.2.1.4. Calidad de los tests empleados como criterio o marcador** | Excelente | 5 |
| **2.11.2.1.5. Promedio de las correlaciones con otros tests que miden constructos similares** | Adecuado | 3 |
| **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** | 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** | Rendimiento escolar | |
| **2.11.2.2.2. Calidad de los criterios empleados** | Excelente | 5 |
| **2.11.2.2.3. Relación temporal entre test y criterio** | Concurrente | |
| **2.11.2.2.4. Tamaño de las muestras** | Un estudio con una muestra moderada (200 < N < 500) o varios estudios con muestras pequeñas (N < 200) | 2 |
| **2.11.2.2.5. Procedimiento de selección de las muestras** | Muestra de personas de EEUU | |
| **2.11.2.2.6. Promedio de las correlaciones del test con los criterios** | Excelente | 5 |

### 2.11.3. Evidencias de validez basadas en la estructura interna:

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

### 2.11.4. Acomodaciones en la administración del test:

| Contenido | Valoración | Puntuación |
|---|---|---|
| **2.11.4. El manual del test informa sobre las acomodaciones en la administración del test** | Puede aplicarse de forma no verbal y también dando instrucciones verbales en castellano y otros idiomas oficiales en España. También ofrece la posibilidad de modificar el punto de inicio en caso de sospechar que la persona tiene diversidad funcional | |

### 2.11.5. Comentarios generales sobre evidencias de validez:

En primer lugar, cabe destacar que TONI-4 es un test que procede de sus versiones anteriores (TONI, TONI-2 y TONI-3), los cuales ya han aportado muchas evidencias de validez que le dan robustez al test TONI-4. Las personas redactoras del manual del test citan muchos estudios e informes realizados, así como un manual específico donde complementar la información sobre validez. Además, otra fortaleza es que en el manual se señala la similitud de resultados encontrados en diferentes países del mundo con grupos culturales, étnicos y lingüísticos diferentes.

En el manual de TONI-4, se aportan evidencias de validez basadas en el contenido justificando y explicando el proceso de creación de las dos formas. Además, justifica la propia evaluación de dichos ítems con las sucesivas revisiones realizadas del TONI. Sin embargo, la justificación aportada en este manual sobre los cambios realizados y las características de las personas expertas debería documentarse con mayor profundidad, además del suplemento técnico que se cita indicando que en él se puede encontrar más información. Uno de los cambios de TONI-4 respecto a sus anteriores versiones es la inclusión de instrucciones verbales. Este hecho es muy positivo dado que surge de la necesidad de distintas personas profesionales que utilizaban el test y les parecía inadecuado no utilizar instrucciones verbales cuando trabajaban con personas sin problemas auditivos. Para ver si las instrucciones verbales podían condicionar la prueba se realizó un estudio administrando el test verbal y después el test no verbal (con una semana de diferencia) a una muestra de 188 personas. Aunque no se encontraron diferencias estadísticamente significativas en la mayoría de los grupos de edad, se encontró que en personas de 6 a 8 años las instrucciones verbales afectaban al resultado de la prueba de forma estadísticamente significativa con un tamaño del efecto pequeño. Las justificaciones que se dan en el manual no son específicas para este grupo de edad. En el manual también se aportan los índices de dificultad y de discriminación de los ítems. En teste sentido, el rango de dificultad es de ,33-,70 para la forma A y ,32-,71 para la forma B. El rango de discriminación de las formas A y B fue de ,31-,51 y de ,32-,50, respectivamente.

Otro aspecto positivo es que se realiza un análisis del Funcionamiento Diferencial de los Ítems (DIF) y se observa que el test no presenta DIF en general respecto a las variables sexo y nacionalidad (española/otras). Para llegar a esta conclusión se usó un procedimiento de regresión logística comparando el ajuste de distintos modelos de regresión. A pesar de obtener resultados de DIF significativos en el ítem 33 de la forma A para el sexo, en el manual del test se considera que esta diferencia no es relevante basándose en el cambio de ajuste del modelo, aunque sería bueno argumentar más claramente esta decisión.

Para obtener evidencias de validez basadas en la estructura interna, se hizo un análisis factorial y uno de invariancia. En relación con el análisis factorial confirmatorio, se esperaba que todos los ítems pesaran de forma apreciable en un factor general (inteligencia). Los índices de ajuste del modelo a los datos fueron excelentes. Por ejemplo, los índices de ajuste comparativo proporcionan un resultado de ,99 sobre un máximo posible de 1 en ambas formas (A y B). Los índices basados en residuales confirman un buen ajuste del modelo con un valor de ,02 próximo al mínimo de cero. Por otra parte, se utilizó el análisis de invariancia, que evalúa el grado de equivalencia de la medida ofrecida por el test en diferentes subgrupos. En este caso, la muestra se clasificó en subgrupos en función de la variable sexo. Los modelos que se compararon fueron: la invariancia de configuración, que pone a prueba si la inteligencia tiene la misma estructura factorial en ambos subgrupos; la invariancia métrica o débil, que pone a prueba si los pesos factoriales de los ítems son los mismos en ambos subgrupos; la invariancia escalar, que iguala los umbrales de los ítems y, por último, la invariancia estricta, que restringe la varianza residual de cada ítem para que sea la misma en ambos grupos. Los cuatro modelos mostraron un excelente ajuste, con índices de ajuste comparativo superiores a ,98 e índices basados en residuales inferiores a ,06. Con base en estos resultados en el manual se afirma que el TONI-4 muestra invariancia respecto al sexo.

También se aportan evidencias de validez basadas en la relación con otras variables. Para aportar datos de validez de tipo convergente, se realizó un estudio con una muestra de 76 personas, con un rango de edad entre los 9 y los 38 años y con una distribución por sexo equilibrada. A este grupo se le aplicó el TONI-4 junto con otro test de inteligencia publicado por la misma editorial, el test Matrices. Los valores de las correlaciones son de ,50 para el total de la muestra, de ,65 para el subgrupo de niños y niñas (n = 50) y de ,45 para el de personas adultas (n = 26). Aunque la correlación total resultante, según los criterios de Hopkins, se encuentra en el límite para ser considerada grande, sería aconsejable dar la misma importancia en los comentarios al coeficiente grande y al moderado. Por otra parte, estas estimaciones serían más robustas con muestras más grandes.

Por otra parte, usando la muestra de tipificación, se presentan relaciones de las puntuaciones del test con la edad, y diferencias de puntuaciones entre grupos de personas varones/mujeres y nacionalidad española/otra. Se observa una correlación grande con la edad, que va en aumento

hasta los 15 años, estancándose después hasta los 50, donde empieza a disminuir. Esto coincide con lo que sería esperado con lo cual es una evidencia favorable de validez. Es importante remarcar que no se encontraron diferencias importantes entre grupos por sexo y nacionalidad, situándose la media de los grupos alrededor de 100 con una desviación típica de 15 en todos los casos.

En relación con el rendimiento académico se presentan datos obtenidos con una muestra estadounidense. Se realizaron tres pruebas el mismo día que se pasó el TONI-4. El Test of Silent Word Reading Fluency, Test of Silence Contextual Reading Fluency y Woodcock-Johnson III Test of Achievement. Este último se divide en distintas subescalas que miden las habilidades académicas generales, matemáticas y de lectura, mientras que los dos primeros miden la fluencia en la lectura. Los coeficientes de correlación obtenidos van de ,52 a ,77, con lo que podemos considerar que se aportan evidencias favorables de validez en relación con los distintos criterios utilizados.

## 2.12. Fiabilidad:

| Contenido | Valoración | Puntuación |
|---|---|---|
| **2.12.1. Datos aportados sobre fiabilidad** | Varios coeficientes de fiabilidad, coeficientes de fiabilidad para diferentes grupos de personas. Error típico de medida para diferentes grupos de personas. | |

### 2.12.2. Equivalencia formas paralelas:

| Contenido | Valoración | Puntuación |
|---|---|---|
| **2.12.2.1. Tamaño de las muestras** | Un estudio con una muestra moderada (200 < N < 500) o varios estudios con muestras pequeñas (N < 200) | 2 |
| **2.12.2.2. Puesta a prueba de los supuestos de paralelismo** | - | - |
| **2.12.2.3. Promedio de coeficientes de equivalencia** | Buena | 4 |

### 2.12.3. Consistencia interna:

| Contenido | Valoración | Puntuación |
|---|---|---|
| **2.12.3.1. Tamaño de las muestras** | Un estudio con una muestra moderada (200 < N < 500) o varios estudios con muestras pequeñas (N < 200) | 2 |
| **2.12.3.2. Coeficientes de consistencia interna presentados** | Alfa de Cronbach | |
| **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** | Un estudio con una muestra moderada (100 < N < 200) o varios estudios con muestras pequeñas (N < 100) | 2 |
| **2.12.4.2. Coeficientes de estabilidad** | Bueno | 4 |

### 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:

En el TONI-4 se informa de 4 tipos de fiabilidad: consistencia interna, formas alternativas (utilizando tanto las aplicaciones consecutivas como las demoradas) y el test-retest. Para valorar la consistencia interna se utilizó el coeficiente alfa de Cronbach en las dos formas del test (A y B) y en cada uno de los grupos de edad. La muestra de tipificación en la que se basan es muy grande, pero al presentarse los datos según grupo de edad el tamaño de las muestras asociadas a cada coeficiente es menor a 200. En todos los grupos los coeficientes alfa superan el valor ,90 y la media calculada es de ,94, un resultado de fiabilidad que puede considerarse excelente según el modelo de evaluación de tests.

Se valora muy positivamente la publicación del error típico de medida máximo para diferentes intervalos de confianza: 90%, 95% y 99%. Los valores publicados son aplicados a las puntuaciones CI, tienen un rango entre 5 y 11 y pueden considerarse en general muy buenos aunque lógicamente se amplían con el aumento del intervalo de confianza. Una ayuda más es que en la correción automática las puntuaciones de las personas se dan con sus intervalos de confianza.

También se aplicaron las formas A y B del TONI-4 de forma consecutiva a la muestra de tipificación. Los coeficientes de correlación se presentan por grupos de edad y se encuentran alrededor de ,7, un resultado que se considera bueno según el modelo español de evaluación de tests. En edades más tempranas el coeficiente de fiabilidad es menor llegando a ser de ,60, un resultado que sería oportuno comentar en el manual incluyendo una explicación de la variabilidad observada y una valoración sobre si era de esperar.

Para calcular la estabilidad temporal se aportan los distintos coeficientes de correlación de la aplicación demorada de las formas alternativas y de la misma forma. Como es usual, los retest se hicieron con muestras menores (103 personas jóvenes y 85 adultas). El promedio de los coeficientes de estabilidad es bueno (,66 ≤ r < ,74) para la aplicación demorada de las formas alternativa y excelente (,79 ≤ r < ,82) para la misma forma.

Para los distintos coeficientes de fiabilidad se dan coeficientes de correlación corregida y sin corregir. Cabe destacar que las diferencias entre el coeficiente de correlación corregido y sin corregir aportados por este manual son pequeñas.

## 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** | Excelente | 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** | Buena | 4 |

### 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:

Por lo que se refiere al tamaño de la muestra, esta es una muestra muy grande (N = 2.820) que intenta obtener la máxima representación de la población española. Para asegurar la máxima representatividad, se seleccionaron distintas variables de estratificación (región geográfica, tamaño de la localidad, sexo, edad, nacionalidad, titularidad del colegio y nivel de estudios avanzados). La muestra resultante es parecida a la muestra de la población española en relación con la edad. En el manual se justifica el amplio rango de edad que abarcan los baremos con el uso de modelos matemáticos de tipificación continua especialmente para poder eliminar las discontinuidades debido a la falta de muestra en algunos de los puntos de edad.

Los baremos en percentiles y puntuación CI permiten pues comparar la puntuación obtenida de la persona evaluada con un conjunto de personas representativas de la población. Gracias al uso de la estrategia de tipificación continua se presentan baremos diferenciados según la edad: 2 baremos de 6 en 6 meses para la edad de seis años, 10 baremos a razón de uno por año entre los 7 y los 18 años y 6 baremos, uno por década, entre los 19 y los 79 años.

En el test se indican los diferentes puntos de referencia sobre la puntuación para poder interpretarla fácilmente. Los puntos de referencia establecidos vienen dados por la puntuación media del CI (100) y su desviación estándar (15). Se eligieron estas medidas por ser las más usadas en el ámbito de la inteligencia. Estos puntos son: CI < 89, lo que indicaría que la persona se encuentra por debajo de la media; CI > 111, indicaría que la persona estaría por encima de la media; 90 > CI < 110, la persona se encuentra en la media y, por último, CI >120, siendo este un indicativo de persona con altas capacidades.

El último tipo de puntuaciones que se proporcionan es la edad mental equivalente que se obtiene mediante la media de los grupos en intervalos de 6 meses. En el manual del test se especifica que estos datos solamente se aportan porque algunas instituciones educativas los usan, pero se avisa que pueden resultar no adecuados y engañosos.