



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 LA NEPSY-II

RESUMEN DE LA VALORACIÓN DEL TEST

Descripción general

Características	Descripción
Nombre del test	NEPSY-II
Autor	Marit Korkman, Ursula Kirk y Sally Kempt
Autor de la adaptación española	Centro de Rehabilitación Neurológica-FIVAN y Laboratorio de Diversidad, Cognición y Lenguaje, del Departamento de Psicología Evolutiva y de la Educación de la Universidad de Sevilla; en colaboración con el Departamento I+D de Pearson Clinical & Talent Assessment
Variable(s)	Funciones cognitivas
Área de aplicación	Psicología clínica, psicología educativa, neuropsicología
Soporte	Administración oral, papel y lápiz, manipulativo

Valoración general

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

Comentarios generales

La batería NEPSY-II constituye una excelente elección para la valoración exhaustiva del funcionamiento neuropsicológico de niños y adolescentes a través de 6 dominios funcionales: Atención y función ejecutiva, Lenguaje, Memoria y aprendizaje, Percepción social, Procesamiento visoespacial y Sensoriomotor.

La evaluación con la NEPSY-II aporta información útil para realizar el diagnóstico de múltiples trastornos infantiles y ayuda a planificar la intervención. La NEPSY-II se basa en la NEPSY (Korkman, Kick y Kemp, 1998), una de las primeras pruebas neuropsicológicas basada en el enfoque de Luria y desarrollada específicamente para niños, pero con un contenido más amplio y una mejora de las propiedades psicométricas.

Entre sus puntos fuertes se pueden destacar los siguientes:

1. La calidad de los materiales de la prueba es buena. El 'Manual clínico y de interpretación' ofrece una información completa en cuanto a la fundamentación teórica, el tratamiento psicométrico realizado y la corrección e interpretación de los resultados.

2. La NEPSY-II dispone de 4 tipos de puntuaciones distintas: puntuaciones principales, puntuaciones de procesamiento, puntuaciones de comparación y observaciones conductuales. Las puntuaciones principales representan las variables clínicas fundamentales de la prueba. Las de procesamiento miden capacidades y habilidades más específicas o tasas de error que pueden no ser relevantes para todos los niños, pero proporcionan información adicional. Las de comparación sirven para comparar las distintas puntuaciones de una misma prueba, o las puntuaciones de distintas pruebas. Esto permite que el examinador pueda comparar estadísticamente las funciones cognitivas, desde un nivel más complejo a uno más básico. De este modo el examinador puede obtener una información muy precisa de los procesos que subyacen e influyen en el rendimiento de las pruebas.

3. El proceso de adaptación española en su conjunto es muy adecuado, además de muy reciente. Se ha obtenido una muestra de tipificación de 639 evaluados, en una prueba de administración individual, y procurando que fuese representativa de la población en cuanto a edad, sexo, nivel educativo de los padres, tipo de población y zona geográfica.

4. En el apartado de validez, cabe destacar las evidencias de validez basadas en la relación con otras variables. La prueba presenta correlaciones adecuadas con otras pruebas neuropsicológicas, como la NEPSY, la CMS y la D-KEFS, y otras pruebas que evalúan capacidad cognitiva general, como la WISC-IV y la WNV. Asimismo se evaluaron grupos especiales (p.ej., niños con TDAH, con trastorno autista, etc.) y se compararon con grupos controles obtenidos de forma aleatoria y equiparados por sexo y edad de la muestra de tipificación. Estos resultados ponen de manifiesto la sensibilidad diferencial de la NEPSY-II para identificar los déficits cognitivos generales y específicos de trastornos habitualmente diagnosticados en la infancia y que interfieren en el rendimiento académico del niño.

5. Se ofrecen buenos indicadores de consistencia interna en distintos grupos de edad. Se informa además de la fiabilidad inter-jueces (como tasa de acuerdo) y de la estabilidad de las puntuaciones.

6. En el 'Manual de aplicación y corrección' se explica de forma detallada, con el uso de ejemplos, cómo se ha de administrar y corregir el test. El 'Manual clínico y de interpretación' incluye información muy precisa sobre cómo interpretar los resultados. La muestra de tipificación española se divide en 22 baremos según edad, obtenidos mediante un procedimiento de tipificación continúa.

7. Se puede utilizar la NEPSY-II para realizar bien evaluaciones generales (para una visión global del estado neuropsicológico del niño, con un tiempo de aplicación de unos 45 minutos a 1 hora), evaluaciones diagnósticas y selectivas (evaluación en profundidad de las áreas seleccionadas por el examinador, el tiempo y el contenido variable, según objetivos), o evaluaciones comprensivas (para una evaluación neuropsicológica exhaustiva, aplicando todos los subtests correspondientes a la edad del niño, el tiempo de aplicación puede variar de 90 minutos hasta 2 y 3 horas). En el 'Manual de aplicación y corrección' (p. 12) se facilitan las pruebas incluidas en la evaluación general, así como las pruebas recomendadas de cada batería para la evaluación diagnóstica de distintos trastornos del desarrollo y dificultades del aprendizaje. Por otra parte, las pruebas y subpruebas que componen la batería pueden aplicarse en cualquier orden.

En cuanto a las debilidades o aspectos mejorables a tener en cuenta en sucesivas ediciones, se podrían señalar los siguientes:

1. En el desarrollo de la NEPSY-II, tanto en su versión estadounidense como española, se ha realizado un enorme esfuerzo por seguir los Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association y National Council on Measurement in Education [NCME], 1999). Sin embargo, en ocasiones se echan en falta los resultados de los análisis realizados (p.ej., al tratar la validez de contenido, al exponer la evidencia basada en los procesos de respuesta o al hablar sobre las propiedades psicométricas de los ítems), posiblemente esto sea debido a que la inclusión de esta información supondría una mayor extensión de los Manuales (189 páginas el de Aplicación y 255 el de interpretación).

2. Se ofrecen coeficientes de consistencia de las puntuaciones en las pruebas por grupos de edad y estos son buenos. Sin embargo, sería conveniente incluir los tamaños de las muestras, así como señalar si son españolas o estadounidenses. Por otro lado, se debería utilizar otros estimadores de fiabilidad que eviten los problemas de Cronbach, como por ejemplo, el Omega.

3. Si bien los distintos tipos de puntuaciones de la NEPSY-II ofrecen información muy rigurosa sobre los procesos cognitivos del niño evaluado, también complejizan la aplicación, corrección e interpretación de la prueba. Tal vez convendría añadir una forma de puntuación 'abreviada', que podría formar parte de la evaluación general.

4. En la p. 32 del 'Manual de aplicación y corrección', en la Figura 2.8 hay una errata, ya que el rango de percentil correspondiente a un total de 3 errores en Tristeza (prueba Reconocimiento

de emociones) cometidos por un niño de 12 años es, según los baremos, 11-25 y no 26-50 como viene en la figura.

5. Algunos tamaños muestrales para los estudios de las relaciones entre la NEPSY-II y otras variables son escasos, especialmente en los estudios realizados con los grupos especiales en los que los tamaños oscilan entre los 18 y los 29 casos. La aportación de estos estudios es sin duda un valor añadido de la prueba, por ello convendría en un futuro conseguir una mayor representatividad de las muestras.

6. No se presentan datos sobre el funcionamiento diferencial de los ítems, aspecto que sería interesante añadir en futuras versiones.

ANÁLISIS DETALLADO DE LA PRUEBA

1. DESCRIPCIÓN GENERAL DEL TEST

1.1. Nombre del test:

- NEPSY-II.

1.2. Nombre del test en su versión original:

- NEPSY-II.

1.3. Autor del test original:

- Korkman, Kirk y Kemp (1998) y Korkman, Kirk y Kemp (2007).

1.4. Autor de la adaptación española:

- Centro de Rehabilitación Neurológica-FIVAN y Laboratorio de Diversidad, Cognición y Lenguaje, del Departamento de Psicología Evolutiva y de la Educación de la Universidad de Sevilla; en colaboración con el Departamento I+D de Pearson Clinical & Talent Assessment.

1.5. Editor del test en su versión original:

- NCS Pearson Inc, 2007.

1.6. Editor de la adaptación española:

- Pearson Educación, 2014.

1.7. Fecha de publicación del test original:

- NEPSY: 1998, NEPSY-II: 2007.

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

- 2014.

1.9. Fecha de la última revisión del test en su adaptación española:

- 2014.

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

- Aptitudes, neuropsicología.

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

- La NEPSY-II permite realizar una evaluación global de las aptitudes cognitivas en dominios neuropsicológicos clave de niños y jóvenes de edades comprendidas entre 3 y 16 años y 11 meses. Concretamente, evalúa 6 dominios que a su vez incluyen distintas subescalas, conformando un total de 32 subtests: 1) Atención y función ejecutiva, 6 subescalas: atención auditiva y flexibilidad cognitiva, clasificación de animales, estatua, fluidez de diseños, inhibición, relojes; 2) Lenguaje, 7 subescalas: comprensión de instrucciones, denominación e identificación de partes del cuerpo, fluidez verbal, procesamiento fonológico, repetición de pseudopalabras, secuencias oromotoras y velocidad de denominación; 3) Memoria y aprendizaje, 7 subescalas: interferencia de lista de palabras, memoria de caras (demorada), memoria de diseños (demorada), memoria de listas (demorada), memoria de nombres (demorada), memoria narrativa y repetición de frases; 4) Sensoriomotor, 4 subescalas: secuencias motoras manuales, golpeteo con la punta de los dedos, imitación de configuraciones de la mano y precisión visomotora; 5) Percepción social, 2 subescalas: reconocimiento de emociones y teoría de la mente; y 6) Procesamiento visoespacial, 6 subescalas: construcción de cubos, copia de diseños, descubrimiento de rutas, flechas, puzles de imágenes y puzles geométricos.
- Si se tienen en cuenta las 4 pruebas demoradas (complementarias) del dominio de Memoria y aprendizaje, el número total de pruebas que compone la NEPSY-II asciende a 36. Las pruebas son muy diferentes, tanto en la presentación de los estímulos como en los procedimientos de aplicación y el tipo de respuestas obtenidas. El profesional puede aplicar todo el instrumento o seleccionar de entre los 32 subtests aquellos que mejor se ajusten a su objetivo de evaluación. De este modo, se puede utilizar la NEPSY-II para realizar bien evaluaciones generales (para una visión global del estado neuropsicológico del niño a través de los 6 dominios, el tiempo de aplicación varía de unos 45 minutos a 1 hora), evaluaciones diagnósticas y selectivas (evaluación en profundidad de las áreas seleccionadas por el examinador, el tiempo y el contenido varía en función de los subtests seleccionados y nivel de ejecución del niño), o evaluaciones comprensivas (para una evaluación neuropsicológica exhaustiva, aplicando todos los subtests correspondientes a la edad del niño, el tiempo de aplicación puede variar de 90 minutos hasta 2 y 3 horas). El primer criterio para la selección de pruebas es la edad del niño ya que las pruebas disponibles para los niños de edades preescolares y escolares son distintas. La evaluación con la NEPSY-II aporta información útil para realizar el diagnóstico de múltiples trastornos infantiles y ayuda a planificar la intervención.

1.12. Áreas de aplicación:

- Psicología clínica, psicología educativa y neuropsicología.

1.13. Formato de los ítems:

- Respuesta construida, respuesta dicotómica y respuesta graduada.

1.14. Número de ítems:

- Son 36 pruebas con un número diferente de tareas y/o estímulos.
- El número de ítems que contesta el participante depende de la edad y de las pruebas administradas. Se presentan a continuación el número de ítems en función de los dominios.

Atención y funciones ejecutivas:

- Atención auditiva y flexibilidad cognitiva (5-16 años): Atención auditiva: 180 ítems y Flexibilidad cognitiva: 180 ítems.
- Clasificación de animales (7-16 años): 20 ítems
- Estatua (3-6 años): N/A.1 tarea.
- Fluidez de diseños (5-12 años): 2 ítems.
- Inhibición: 4 ítems (5-6 años) y 6 ítems (7-16 años)
- Relojes (7-16 años): 10 ítems.

Lenguaje:

- Comprensión de instrucciones (3-16 años): 33 ítems.
- Denominación e identificación de partes del cuerpo (3-4 años): 11 ítems.
- Fluidez verbal: 2 ítems (3-6 años); 4 ítems (7-16 años)
- Procesamiento fonológico: 22 ítems (3-4 años); 45 ítems (5-16 años).
- Repetición de pseudopalabras (5-12 años): 13 ítems.
- Secuencias oromotoras (3-12 años): 14 ítems.
- Velocidad de denominación: 2 ítems (3-4 años); 3 ítems (5-16 años).

Memoria y aprendizaje:

- Interferencia de lista de palabras (7-16 años): 10 ítems.
- Memoria de caras (demorada) (5-16 años): 16 ítems (16 demorada).
- Memoria de diseños (demorada): 4 ensayos (3-4 años); 4 ensayos y 1 demorado (5-16 años).
- Memoria de listas: 7 ensayos y 1 demorado (7-12 años).
- Memoria de nombres (demorada) (5-16 años): 8 ítems (8 demorada).
- Memoria narrativa: 1 historia (3-4 años); 2 historias (5-16 años).
- Repetición de frases (3-6 años): 17 ítems.

Funcionamiento sensoriomotor:

- Secuencias motoras manuales (3-12 años): 12 ítems.
- Golpeteo con la punta de los dedos (5-16 años): 4 ítems.
- Imitación de configuraciones de la mano (3-12 años): 24 ítems.

- Precisión visomotora: 5 ítems (3-4 años); 2 ítems (5-12 años).

Percepción social:

- Reconocimiento de emociones: 16 ítems (3-4 años); 35 ítems (5-16 años).
- Teoría de la mente (3-16 años): 21 ítems.

Procesamiento visoespacial:

- Construcción de cubos (3-16 años): 19 ítems.
- Copia de diseños: 18 ítems (3-16 años); 21 ítems (7-16 años).
- Descubrimiento de rutas (5-12 años): 10 ítems.
- Flechas (5-16 años): 21 ítems.
- Puzles de imágenes (7-16 años): 20 ítems.
- Puzles geométricos: 12 ítems (3-6 años); 20 ítems (7-16 años).

1.15. Soporte:

- Administración oral, papel y lápiz, manipulativo, estímulos auditivos.

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:

- La NEPSY-II es aplicable a niños y adolescentes de edades comprendidas entre los 3 años y los 16 años y 11 meses.
- Niños con necesidades especiales: la NEPSY-II también es aplicable a niños con TDAH, con trastorno de la lectura-dislexia, con trastorno de cálculo, con trastorno de lenguaje, con discapacidad intelectual leve, con trastorno autista, con trastorno de Asperger, con dificultades de audición y niños prematuros. En el 'Manual de aplicación y corrección' (capítulo 2, p. 18) vienen algunas recomendaciones para adaptar la aplicación a las características de estos niños. Asimismo, en el 'Manual clínico y de interpretación' (p. 13) aparecen unas guías que pretenden orientar al examinador en la selección de las pruebas más adecuadas en función del motivo de consulta o problemática del sujeto (dificultades de aprendizaje, problemas de atención, trastornos del lenguaje, problemas de conducta o problemas sociales).

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

- No. Sí diferencian los 3 tipos de evaluaciones anteriormente señaladas (general, selectiva/diagnóstica y completa). Para la general sí indican las pruebas a evaluar según la edad del niño se encuentre en uno de los 2 rangos: 3-4 años o 5-16 años (p. 11, 'Manual de aplicación y corrección').

1.19. Procedimiento de corrección:

- Manual.

1.20. Puntuaciones:

- a. Puntuaciones principales. Representan los aspectos globales o las variables clínicas fundamentales de la prueba. Se expresan como puntuaciones escalares, aunque algunas se expresan como percentiles.
- b. Puntuaciones de procesamiento. Ofrecen información más específica sobre las aptitudes que son necesarias para completar una tarea o que influyen en el rendimiento de la prueba.
- c. Puntuaciones de comparación. Utilizan la métrica de las puntuaciones escalares (media 10 y desviación típica 3) para poder comparar las distintas puntuaciones de la misma prueba, o de distintas pruebas.
- d. Observaciones conductuales. Las observaciones conductuales proporcionan datos cuantitativos sobre conductas que se observan habitualmente en los niños.

1.21. Escalas utilizadas:

- Centiles y CI de desviación (M:10; DT:3).

1.22. Posibilidad de obtener informes automatizados:

- No.

1.23. Tiempo estimado para la aplicación del test:

- Individual: de 45 minutos a 3 horas dependiendo de si la evaluación es general o completa.

1.24. Documentación aportada por el editor:

- Manual.

1.25. Precio de un juego completo de la prueba:

- 795 € sin iva (consultado 22/07/2016).

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

- Cuadernillo Anotación (3-4 años) 60,20 € sin iva. (consultado 22/07/2016).
- Cuadernillo Anotación (5-16 años) 78,50 € sin iva. (consultado 22/07/2016).

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

- Cuadernillo de 25 Respuestas (3-4 años) 35,70 € sin iva. (consultado 22/07/2016).
- Cuadernillo de 25 Respuestas (5-16 años) 45,90 € sin iva. (consultado 22/07/2016).

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

- No aplicable.

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	Buena	4
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	Buena	4
2.9. Bibliografía del manual	Excelente	5
2.10. Datos sobre el análisis de los ítems	-	-

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	moderado de expertos mediante un procedimiento sistematizado	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 con otras variables

Contenido	Valoración	Puntuación
2.11.2.1. Diseños empleados	Correlaciones con otros tests	
2.11.2.1.1. Tamaño de las muestras	Un estudio con una muestra moderada o varios con muestras pequeñas	2
2.11.2.1.2. Procedimiento de selección de las muestras	Incidental	
2.11.2.1.3. Calidad de los tests empleados como criterio o marcador	Buena	4
2.11.2.1.4. Promedio de las correlaciones con otros tests que miden constructos similares	Adecuada	3
2.11.2.1.5. Promedio de las correlaciones con otros tests que miden constructos no relacionados	Buena	4
2.11.2.1.6. Resultados de la matriz multirrasgo-multimétodo	-	-
2.11.2.1.7. 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	Diferentes grupos: TDAH, autismo, dificultades de aprendizaje-lectura-dislexia, discalculia, lenguaje- discapacidad intelectual leve, dificultades de audición y niños prematuros.	
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	Un estudio con una muestra moderada o varios con muestras pequeñas	2
2.11.2.2.5. Procedimiento de selección de las muestras	Incidental	
2.11.2.2.6. Promedio de las correlaciones del test con los criterios	Buena	3

2.11.3. Evidencias de validez basadas en la estructura interna

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

2.11.4. Acomodaciones en la administración del test

Contenido	Valoración	Puntuación
2.11.4. El manual del test informa sobre las acomodaciones en la administración del test		Sí

Comentarios generales sobre evidencias de validez

El 'Manual clínico y de interpretación' dedica el capítulo 6 a las fuentes de evidencia de validez del NEPSY-II según los Standards for Educational and Psychological Testing (American educational Research Association, American Psychological Association y National Council on Measurement in Education [NCME], 1999, 2015).

En el capítulo se exponen los procedimientos de validez usados y los resultados de diversas investigaciones sobre la validez de contenido, de constructo y de criterio del test.

En cuanto a la validez de contenido se hace referencia a la evidencia basada en los procesos de respuesta. Sin embargo, como se ha comentado con anterioridad, se echa en falta mayor información sobre los resultados hallados en los distintos análisis realizados y comentados en el Manual.

Únicamente se examinan las correlaciones para la muestra total entre las puntuaciones de las pruebas. Asimismo, en el anexo E se presentan las correlaciones por grupo de edad, así como de los grupos especiales. Las correlaciones más altas se dieron entre las pruebas del mismo dominio, siendo más bajas las existentes entre las pruebas de distintos dominios. El patrón de correlaciones entre las pruebas de un mismo dominio evidenciaría la validez de la estructura de la NEPSY-II.

Dentro del capítulo 6 se incluye un capítulo sobre la evidencia de validez concurrente. En este apartado se presentan las relaciones de la NEPSY-II con otras medidas neuropsicológicas: NEPSY (N = 109), *Children's Memory Scale* (CMS; Cohen, 1997) (N = 43), *Delis-Kaplan Executive Function System* (D-KEFS; Delis et al., 2001) (N = 49); así como con otras medidas que evalúan capacidad cognitiva general: Escala de inteligencia de Weschler para niños-IV (WISC-IV; Wechsler, 2003, 2011) (N = 51), Escala no verbal de aptitud intelectual

de Wechsler (WNV; Wechsler y Naglieri, 2005, 2011) (N = 62). En general los resultados fueron satisfactorios.

En el apartado sobre la evidencia de validez concurrente se incluye también un subapartado sobre grupos especiales donde se presentan los resultados de los estudios relacionados con estos grupos con el fin de proporcionar pruebas adicionales de validez, basadas en las relaciones entre el test y diversos criterios. Estos resultados son de especial interés ya que la demostración de la utilidad y especificidad clínicas de una escala es fundamental cuando sus resultados forman parte de una evaluación diagnóstica exhaustiva. En este subapartado se presentan medias y desviaciones típicas de cada grupo (especiales/grupo control muestra tipificación de la NEPSY-II), así como las diferencias, incluido el tamaño del efecto. Para la creación del grupo control se seleccionaron aleatoriamente los casos y se emparejó la muestra obtenida con el grupo especial según la edad y el sexo. Si bien las muestras no son representativas de la categoría diagnóstica en su totalidad, los datos especificados en esta sección ponen de manifiesto la sensibilidad diferencial de la NEPSY-II a los déficits cognitivos generales y específicos mostrados por los niños que más frecuentemente son evaluados en diferentes ambos clínicos.

2.12. Fiabilidad

Contenido	Valoración	Puntuación
2.12.1. Datos aportados sobre fiabilidad	Varios coeficientes de fiabilidad, error típico de medida para diferentes grupos, dos mitades y test-retest.	

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	Un estudio con una muestra moderada o varios	2

	con muestras pequeñas	
2.11.3.2. Coeficientes de consistencia interna presentados	Alfa o KR-20 y dos mitades	
2.11.3.3. Promedio de los coeficientes de consistencia	Buena	4

2.12.4. Estabilidad (test-retest)

Contenido	Valoración	Puntuación
2.12.4.1. Tamaño de las muestras	Un estudio muestra pequeña	1
2.12.4.2. Coeficientes de estabilidad	Buena	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	Porcentaje de acuerdo	
2.12.6.2. Promedio de los coeficientes	Excelente	5

Comentarios generales sobre fiabilidad

Los coeficientes de consistencia interna de las puntuaciones en las pruebas son muy adecuados y además son estudiados por grupos de edad. Se realiza un estudio de estabilidad de las puntuaciones aportando asimismo resultados positivos.

Se ofrecen coeficientes de consistencia interna de las puntuaciones en las pruebas por grupos de edad. Se especifica si las muestras son estadounidenses o españolas. También se muestran los errores típicos de las puntuaciones en las pruebas por grupo de edad. Se aportan datos de la estabilidad de las puntuaciones en una muestra pequeña (N = 90). Los resultados de los estudios de fiabilidad indican que la consistencia interna o estabilidad de la mayoría de

las pruebas de la NEPSY-II es de adecuada a excelente. Las pruebas con los coeficientes de fiabilidad más altos son Copia de diseños, Golpeteo con la punta de los dedos, Imitación de configuraciones de la mano, Memoria de listas, Memoria de nombres, Procesamiento fonológico, Puzzles de imágenes, Relojes y Repetición de frases. Una muestra aleatoria de 150 protocolos de la NEPSY-II fueron sometidos a una doble corrección por parte de dos correctores independientes.

La fiabilidad inter-jueces también arrojó resultados muy positivos para la fiabilidad de la NEPSY-II. La fiabilidad inter-jueces fue casi perfecta, con un porcentaje de acuerdo que osciló entre el 76% (Fluidez verbal) y el 99% (Teoría de la mente).

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	4
2.13.1.2. Tamaño de las muestras	Grande	4
2.13.1.3. Aplicación de estrategia de tipificación continua	Sí	
2.12.1.4. Procedimiento de selección de las muestras	Aleatorio, con muestras representativas	
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	-	-

Comentarios generales sobre baremación e interpretación de las puntuaciones

La tipificación sigue el método de tipificación continua en intervalos de 6 meses para edades de 3 a 12 años y 11 meses y de un año para los mayores.