



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 ESCALA DE INTELIGENCIA DE WECHSLER PARA NIÑOS-V (WISC-V)

RESUMEN DE LA VALORACIÓN DEL TEST

Descripción general

Características	Descripción
Nombre del test	Escala de Inteligencia de Wechsler para Niños-V
Autor	David Wechsler
Autor de la adaptación española	Ana Hernández, Cristina Aguilar, Erica Paradell y Frédérique Vallar
Variable(s)	Inteligencia
Área de aplicación	Todas en las que se evalúe la inteligencia en niños y adolescentes
Soporte	Administración oral; manipulativo; lápiz y papel; Plataforma on-line Q-global

Valoración general

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

Comentarios generales

La escala de inteligencia de Wechsler para niños-V (WISC-V) es un instrumento clínico de aplicación individual que evalúa la inteligencia en niños de 6 años y 0 meses a 16 años y 11 meses. El WISC-V se puede considerar como una de las herramientas más importantes para la evaluación de este constructo. La elección de las escalas de inteligencia Wechsler, y concretamente la versión V, se asienta en el amplio respaldo empírico y teórico.

La inteligencia es uno de los factores psicológicos de mayor relevancia para la comprensión de la conducta humana. La escala WISC-V ofrece puntuaciones de los índices primarios de inteligencia que reflejan el funcionamiento intelectual en diferentes áreas cognitivas: comprensión verbal, visoespacial, razonamiento fluido, memoria de trabajo y velocidad de procesamiento. Así mismo ofrece una puntuación general (CI total) y de igual forma puntuaciones en índices secundarios que reflejan aptitudes cognitivas en diferentes agrupaciones que atienden a necesidades clínicas (p. ej. No verbal y capacidad general).

Una de las principales fortalezas del WISC-V es el hecho de que engloba diferentes manifestaciones de la inteligencia atendiendo a la manifestación multidimensional de la misma. También resultan útiles en la evaluación de grupos clínicos en los que hay dificultades atencionales y/o motivacionales, o en aquellos en los que la fatiga puede aparecer pronto.

Entre otra de las principales fortalezas cabe destacar que se incluyen nuevas medidas de razonamiento fluido, visoespacial y memoria de trabajo. Así mismo, se dispone de nuevos y diferenciados Índices de razonamiento fluido y visoespacial y nuevos índices secundarios que permiten ajustar el diagnóstico a situaciones clínicas específicas. Por otro lado, la fiabilidad de las puntuaciones resulta óptima y de igual forma, existen diferentes evidencias sobre la validez del mismo. Por ejemplo, la validez en relación al constructo ha sido constatada, así como en relación a otras fuentes externas.

En futuras revisiones se podría analizar, por ejemplo, la invarianza de medición de las puntuaciones del WISC-V, el DIF en función del género, edad u otra variable de interés así como la función de información bajo la TRI.

ANÁLISIS DETALLADO DE LA PRUEBA

1. DESCRIPCIÓN GENERAL DEL TEST

1.1. Nombre del test:

- Escala de Inteligencia de Wechsler para Niños-V.

1.2. Nombre del test en su versión original:

- Wechsler Intelligence Scale for Children-Fifth Edition.

1.3. Autor del test original:

- David Wechsler.

1.4. Autor de la adaptación española:

- Ana Hernández, Cristina Aguilar, Erica Paradell y Frédérique Vallar.

1.5. Editor del test en su versión original:

- NCS Pearson, Inc. (Pearson Assessment).

1.6. Editor de la adaptación española:

- Pearson Educación, 2014.

1.7. Fecha de publicación del test original:

- WISC: 1949
- WISC-R: 1974
- WISC-III: 1991
- WISC-IV: 2003
- WISC-V: 2014

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

- WISC: 1974
- WISC-R: 1993
- WISC-III: (no adaptado)
- WISC-IV: 2005
- WISC-V: 2015

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

- 2015.

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

- Inteligencia, aptitudes.

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

- La adaptación española del test WISC-V para la evaluación de las capacidades intelectuales consta de un total de 15 pruebas que se organizan en tres niveles de interpretación: la escala total o CI total, los índices primarios (Comprensión verbal, Visoespacial, Razonamiento fluido, Memoria de trabajo y Velocidad de Procesamiento) y los índices secundarios (Razonamiento cuantitativo, Memoria de trabajo auditiva, No verbal, Capacidad general y Competencia cognitiva). Una actualización destacable de la WISC-V es la eliminación del Índice de Razonamiento Perceptivo (IRP) que se ha sustituido por el Índice Visoespacial (IVE) y el Índice de Razonamiento Fluido (IRF).
- Las 7 pruebas fundamentales que contribuyen al CI total son: Semejanzas, Vocabulario, Cubos, Matrices, Balanzas, Dígitos y Claves. En el caso de los índices primarios, las pruebas que contribuyen a cada índice son 10: Semejanzas y Vocabulario (ICV), Cubos y Puzles visuales (IVE), Matrices y Balanzas (Razonamiento fluido), Dígitos y Span de dibujos (IMT), y Claves y Búsqueda de símbolos (IVP). Los índices primarios, junto con el CI total, constituyen las

puntuaciones necesarias para llevar a cabo una descripción y evaluación completas de la capacidad intelectual. Los índices secundarios son puntuaciones compuestas que se obtienen a partir de combinaciones de pruebas o puntuaciones principales y opcionales, y son:

- El índice de razonamiento cuantitativo (IRC) se obtiene de las pruebas Balanzas y Aritmética.
- El índice de memoria de trabajo auditiva (IMTA) se obtiene de las pruebas Dígitos y Letras y números.
- El índice no verbal (INV) es una medida global de la aptitud intelectual que no requiere respuestas expresivas en las pruebas que no tienen contenido verbal, que son Cubos, Puzles visuales, Matrices, Balanzas, Span de dibujos y Claves.
- El índice de capacidad general (ICG) es una estimación de la aptitud intelectual general menos dependiente de la memoria de trabajo y de la velocidad de procesamiento que se obtiene con las pruebas de Semejanzas, Vocabulario, Cubos, Matrices y Balanzas.
- Por último, el índice de competencia cognitiva (ICC) se basa en las pruebas de Dígitos, Span de dibujos, Claves y Búsqueda de símbolos.

1.12. Áreas de aplicación:

- Psicología clínica, psicología educativa, neuropsicología, psicología forense, servicios sociales.

1.13. Formato de los ítems:

- Respuesta Libre: Semejanzas, Vocabulario, Información, Comprensión.
- -Presentación de modelos a imitar: Cubos.
- -Presentación de estímulos:
 - a repetir o reorganizar: Dígitos, Letras y números
 - a asociar o relacionar (dibujos, números o figuras): Claves, Matrices, Puzles visuales, Balanzas, Span de dibujos
- Tachado de símbolos: Cancelación.
- Problemas a resolver: Aritmética
- Identificación de presencia ausencia de estímulos gráficos: Búsqueda de Símbolos

1.14. Número de ítems:

Subtest	Ítems de eliminación	Ítems de presentación / entrenamiento	Ítems efectivos
Cubos		-	13
Semejanzas		1	23
Matrices		2	32
Dígitos			

Orden directo		-	9 (18 intentos)
Orden inverso		2	1 (2 intentos) 9 (18 intentos)
Orden creciente		4	2 (4 intentos) 9 (18 intentos)
Claves			
Claves A (6 y 7 años)		5	75
Claves B (8 a 16 años)		9	117
Vocabulario		-	29
Balanzas		2	34
Puzles visuales		2	29
Span de dibujos		3	26
Búsqueda de símbolos			
Parte A (6 y 7 años)		5	40
Parte B (8 a 16 años)		5	60
Información		-	31
Letras y números		4 (5 intentos)	10 ítems (30 intentos)
Cancelación		1 ítem ejemplo 1 ítem demostración	2 ítems (1 aleatoria, 1 estructurada)
Presentación		48	-
Aleatoria		-	320
Estructurada		-	320
Comprensión		-	19
Aritmética		-	34

1.15. Soporte:

- Administración oral, papel y lápiz y manipulativo.

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:

- El test es aplicable a niños y adolescentes de edades comprendidas entre 6 años 0 meses y 16 años 11 meses. Abarca los niveles educativos de Primaria (6 a 11), Secundaria (12-15) y Bachillerato (primer curso, 16 años). El test es aplicable preferentemente a grupos de individuos de los que convenga determinar su nivel de habilidades cognitivas o su funcionamiento neuropsicológico: Diagnóstico de altas capacidades; Diagnóstico de discapacidad intelectual leve o moderado; Diagnóstico de

Trastornos de aprendizaje; Disfunciones neuropsicológicas. También aplica a poblaciones con alteraciones atencionales, como el TADH y puede ser usado para el diagnóstico de minorías étnicas. Los índices de razonamiento perceptivo y de velocidad de procesamiento, son especialmente adecuados para ello. Para la evaluación de niños con NEE se contemplan modificaciones del procedimiento estándar; se recoge en la página 14 del manual de aplicación y corrección.

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

- No existen formas paralelas ni tampoco versiones abreviadas específicas, aunque el test contempla una aplicación con 7 pruebas, y la aplicación completa con 15 pruebas.

1.19. Procedimiento de corrección:

- Manual y automatizada por ordenador.

1.20. Puntuaciones:

- En general se obtiene una puntuación directa a partir de la suma de los aciertos en cada uno de los ítems. Siguen este procedimiento las pruebas de: Claves, Matrices, Información, Aritmética, Balanzas y Puzles visuales.
- Se dan además las siguientes variaciones:
- En la prueba de Cubos se puntúa acierto/error, pero se pondera el acierto en la respuesta, según el intento en que se logra el éxito (ítems 1 a 3), la dificultad del ítem (4 puntos para los ítems 4 a 8), y según la rapidez de la respuesta (hasta 7 puntos para ítems 9 a 13). La puntuación directa equivale a la suma de puntos obtenidos a cada ítem.
- En la prueba de Span de dibujos se puntúa con 2 puntos si el niño selecciona todos los dibujos mostrados en la página en el orden correcto, con 1 punto si selecciona todos los dibujos mostrados pero en orden incorrecto, y se puntúa como 0 si el niño no selecciona todos los dibujos mostrados, selecciona un dibujo incorrecto, dice que no lo sabe o no responde en aproximadamente 30 segundos.
- En las pruebas de Semejanzas, Vocabulario y Comprensión se pondera la respuesta en función de la precisión o calidad de la respuesta, en una escala Likert de anclajes en 0 = erróneo y 2 = respuesta acertada precisa. La puntuación directa equivale a la suma de puntos obtenidos a cada ítem.
- En las pruebas de Dígitos y Letras y números se puntúa el total de aciertos habidos en cada uno de los intentos de cada ítem (dos y tres intentos para cada ítem respectivamente). La puntuación directa equivale a la suma de puntos dados a cada ítem.
- En la prueba de Búsqueda de Símbolos la puntuación directa se obtiene calculando la diferencia entre el número de aciertos y el de errores, del total de ítems resueltos.
- En las pruebas de Cancelación se obtiene la puntuación directa calculando la diferencia entre aciertos y errores, y bonificando con puntos la rapidez en la respuesta.

1.21. Escalas utilizadas:

- Centiles, puntuaciones típicas, CI.

1.22. Posibilidad de obtener informes automatizados:

- Sí a través de la plataforma online Q-global de Pearson.

1.23. Tiempo estimado para la aplicación del test:

- La aplicación es individual y la duración aproximada para la aplicación de las 10 diez pruebas principales es de 65 minutos y 48 minutos para la aplicación de las 7 pruebas para calcular el CI total.

1.24. Documentación aportada por el editor:

- Manual de aplicación y corrección.
- Manual técnico y de interpretación.
- Otros: Caja con 9 cubos bicolores; Cuaderno de estímulos 1 y 2.

1.25. Precio de un juego completo de la prueba:

- 985,00 € (incluye 25 perfiles on-line, plataforma Q-global). (Fecha de consulta: 07/07/2016).

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

- Cuadernillos de anotación WISC-V (25 unidades); 78,50 €. (Fecha de consulta: 07/07/2016).

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

- Cuadernillo de respuestas 1: 42,50 €. (Fecha de consulta: 07/07/2016).
- Cuadernillo de respuestas 2: 22,50 €. (Fecha de consulta: 07/07/2016).

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

- Cuadernillo de anotación + 25 perfiles on-line (Plataforma Q-global): 103,50€.
- Recarga 25 perfiles on-line (Plataforma Q-Global) - WISC-V: 50 €.
- (Fecha de consulta: 07/07/2016).

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	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	Amplia muestra aleatoria	5

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, diferencias entre grupos, matriz multirragos-multimétodo	
2.11.2.1.1. Tamaño de las muestras	Estudio con muestra grande	3
2.11.2.1.2. Procedimiento de selección de las muestras	Aleatorio, con muestras representativas	
2.11.2.1.3. Calidad de los tests empleados como criterio o marcador	Excelente	5
2.11.2.1.4. Promedio de las correlaciones con otros tests que miden constructos similares	Buena	4
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 multirrago-multimétodo	Excelente	5
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	Amplia variedad y representatividad de poblaciones	
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 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	4

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	-	-

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

Los comentarios sobre las evidencias de validez se exponen a continuación:

Calidad de la representación del contenido o dominio. Las evidencias de validez obtenidas son excelentes o casi excelentes en la mayoría de los casos. Cabe señalar que el tiempo de aplicación se ha reducido mientras que se ha ampliado el constructo medido, se han reducido los requisitos del lenguaje expresivo siendo el nivel de vocabulario menos relevante en la estimación del CI, y se han añadido nuevos y diferenciados índices de razonamiento fluido y visoespacial así como índices secundarios que permiten ajustar el diagnóstico a situaciones clínicas específicas. Además, se han llevado a cabo modificaciones en los procedimientos de aplicación y corrección para mejorar la facilidad de uso de la escala.

Consultas a expertos. En el capítulo 2 del Manual se expone en detalle el procedimiento seguido en la adaptación española para lograr un contenido válido: se han actualizado las bases teóricas considerando los nuevos hallazgos sobre el neurodesarrollo y la neurocognición así como los nuevos modelos sobre la memoria de trabajo y se hicieron pruebas pilotos adicionales para sustentar la elección de los ítems.

Tamaño de las muestras: La validación del constructo se ha realizado mediante una estratificación de la muestra en función de la edad, el sexo, el nivel educativo de los padres, tipo de población y la zona geográfica. Por un lado se utilizó una muestra de 1008 niños para la tipificación real del test, los análisis factoriales, de validez y las diferencias entre grupos. Asimismo, las pruebas de validez incluyeron análisis con grupos especiales españoles que incluyeron niños con alta capacidad intelectual y discapacidad intelectual.

Correlación del test con otros similares (validez convergente) y distintos (validez discriminante). En la adaptación española se han establecido correlaciones entre las pruebas del WISC-V y o WISC-IV, WPPSI-IV, WAIS-IV, KABC-II, VINELAND-II y BASC-2. La adaptación española solo correlacionó las puntuaciones del WISC-V con las del WISC-IV con un grupo de 80 niños españoles. Las correlaciones corregidas entre las pruebas comunes de ambas pruebas fueron de moderadas a altas. La correlación más baja se observa en Búsqueda de símbolos (.45), pero las otras dos pruebas de Velocidad de procesamiento, Claves y Cancelación, muestran correlaciones más altas (.62 y .63 respectivamente).

Datos sobre el sesgo de los ítems. Este análisis no se ha llevado a cabo con datos españoles por diversos motivos. En primer lugar porque un análisis de sesgo basado en TRI requiere acumular una cantidad considerable de sujetos, en torno a 200, que hayan dado respuesta a cada uno de los ítems de cada subtest. Aunque la muestra de tipificación de la adaptación española es lo suficientemente grande, no todas las personas contestaron a todos los ítems, puesto que están ordenados por dificultad y existen reglas de terminación que impiden que, en la mayoría de los casos, no se llegue al final del test. Para hacer un análisis de este tipo se requiere hacer aplicaciones adicionales sin tiempo a una muestra diferente de la de tipificación. Y puesto que el WISC-V español es una adaptación del WISC-V americano y ambas versiones comparten un gran porcentaje de los ítems, aunque ocupen puestos diferentes por su dificultad, no se consideró necesario replicar los análisis de sesgo que sí han sido realizados con la versión original y que permitieron llegar a una versión final de cada subtest con los mejores ítems. Por otro lado, se consideró que, puesto que ya había quedado demostrada la bondad de los ítems en la versión original, resultaba más adecuado estudiar cómo esos ítems funcionaban en la población española mediante los correspondientes análisis de dificultad, discriminación y contribución a la fiabilidad. Los resultados de estos análisis determinaron la ordenación de los ítems de cada subtest así como la eliminación de aquellos que mostraron un mal funcionamiento.

2.12. Fiabilidad

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

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	Buena	4
2.11.3.2. Coeficientes de consistencia interna presentados	Coeficientes de estabilidad corregidos para las pruebas de velocidad de procesamiento, método de las dos mitades corregida para el resto de pruebas. Los coeficientes de fiabilidad promedio se calcularon utilizando la transformación Z de Fisher	
2.11.3.3. Promedio de los 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 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	Coeficiente de correlación intraclase	
2.12.6.2. Promedio de los coeficientes	Excelente	5

Comentarios generales sobre fiabilidad

La consistencia interna del test es alta para las pruebas, las puntuaciones de procesamiento y las puntuaciones compuestas, tanto en los grupos especiales como en la muestra de tipificación. Además, las puntuaciones son muy consistentes al cabo de varias semanas, y la corrección de las pruebas que requieren una valoración más subjetiva (pruebas de Comprensión verbal) ofrece resultados similares en distintos correctores.

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	Excelente	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	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

Procedimiento de selección de las muestras. Se ha conseguido una buena representación según el censo español en relación a la representación de las regiones. Tal y como puede observarse en las tablas del manual del WISC-V, todos los porcentajes reales se ajustan de forma satisfactoria a los porcentajes teóricos basados en el censo de la población española, con ligeras variaciones. La diferencia más grande se observa en la variable tipo de población, para los niños urbanos y para el grupo de edad de 16 años.