PsychCorp

WISC®-V

ESCALA DE INTELIGENCIA
DE WECHSLER PARA NIÑOS-V

Manual técnico
y de interpretación



ALWAYS LEARNING

PEARSON



WISC-V

ESCALA DE INTELIGENCIA DE WECHSLER PARA NIÑOS-V

Manual técnico y de interpretación

David Wechsler

Directores de investigación de la WISC-V

Susan Engi Raiford

James A. Holdnack

PsychCorp

Alemania Australia/Nueva Zelanda Canadá Dinamarca Estados Unidos Francia Holanda India Reino Unido Suecia/Noruega

PsychCorp

Wechsler Intelligence Scale for Children- Fifth Edition. Copyright © 2014 NCS Pearson, Inc.
Copyright de la adaptación española © 2015 NCS Pearson, Inc. Todos los derechos reservados. Adaptación española realizada y distribuida por Pearson Educación, S. A., Ribera del Loira, 28, 1.º, Madrid 28042, con la autorización de NCS Pearson, Inc. (EE. UU.).
Pearson, PsychCorp, Wechsler y WISC son marcas registradas en EE. UU. y/u otros países por Pearson Education, Inc. o afiliados.

Impreso por BOOK PRINT DIGITAL, S.A.

ISBN: 978-84-9035-302-8
Depósito legal: M-29872-2015

Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del «Copyright», bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo públicos.

PEARSON

Pearson Clinical & Talent Assessment

www.pearsonclinical.es

Agradecimientos de Pearson Clinical & Talent Assessment

Para Pearson Clinical & Talent Assessment supone una gran satisfacción poner a disposición de los profesionales de la psicología la *Escala de inteligencia de Wechsler para niños-V* (WISC-V).

Numerosos cambios han sido introducidos en la WISC-V respecto a su predecesora para ofrecer una herramienta basada en los modelos de la inteligencia contemporáneos y la investigación sobre el neurodesarrollo y la neurocognición, que permita una evaluación completa de la aptitud intelectual de los niños de 6 años a 16 años 11 meses.

Los cambios más destacables son los siguientes:

- Se presenta una nueva estructura en tres niveles de interpretación (escala total, escalas primarias y escalas secundarias) que permite adaptar la evaluación a cada situación clínica y obtener una mayor cantidad de información.
- Se incluyen 5 índices primarios (Comprensión verbal, Visoespacial, Razonamiento fluido, Memoria de trabajo y Velocidad de procesamiento) y 5 índices secundarios (Razonamiento cuantitativo, Memoria de trabajo auditiva, No verbal, Capacidad general y Competencia cognitiva).
- Se incorporan tres pruebas nuevas: *Puzles visuales*, que contribuye al Índice visoespacial; *Balanzas*, que contribuye al Índice de razonamiento fluido; y *Span de dibujos*, que contribuye al Índice de memoria de trabajo.
- Se ofrecen nuevas posibilidades de análisis primario y secundario que permiten comparar y evaluar aspectos globales y específicos del rendimiento.

La adaptación española de la WISC-V ha requerido la dedicación y el esfuerzo de muchos profesionales, a los que queremos transmitir nuestro más sincero agradecimiento.

En primer lugar, nos gustaría expresar nuestra gratitud a todos los colegios oficiales de psicología de España que nos permitieron seleccionar a los profesionales más adecuados y nos brindaron un espacio para llevar a cabo la formación.

En segundo lugar, queremos reconocer y agradecer la labor de los psicólogos que participaron en la recogida de datos de las muestras experimental y de tipificación, quienes actuaron con enorme diligencia y profesionalidad, respetando siempre los criterios de estratificación asignados por Pearson:

Fase experimental

Norte

Almudena Gutiérrez Higuero
Ana María Pena Pereira
Andrea López Menéndez
Araceli Álvarez Alegre
Olalla Rodríguez Giráldez

Sur

Alejandra Martín Padrón
Ana Antelo López
Ana Belén González Luna
Ana Isabel Álvarez Torregrosa
Ana María Laguna López
Beatriz Luna Alonso
Davinia Saray García Silva
Jorge Alejandro Fernández Hidalgo

Este

Adrián Ruiz Parra
Gisela Ferré Rey
Javier Berrio de Haro
Justo Rodríguez Pau
Laura Monzonis Domènech
Natalia Bañó Alexandre
Neus Fàbregas Bofill

Centro

Alicia Escuadra Río
Ana Carmen Molina Lafuente
Cristina Gómez Domínguez
Dolores Puentes Parras
Estefanía Barrado Mariscal
Raquel Rodríguez Ramos
Rosa María Salas Muñoz

Fase de tipificación

Norte

Alicia Gutiérrez Baeza
Ana Hernández González
Andrea Revuelta Echevarría
Beatriz González Díaz de Tudanca
Estela Alonso Martínez
Estrella Fernández
Genoveva Montoya Murillo
Laura Conde Bande
María Antonia Maniega Rubio
María Dolores Sánchez
Mercedes Rey Morillo
Patricia Fernández Varela
Paula Cabal
Rosa María Feijoo Rebollo
Rosalía González Freijo
Sarah Berrocoso Cascallana
Ziortza Arana González

Sur

Alicia Pérez Santana
Amanda Librero López
Ana Gracia Naval
Beatriz Aguilar González
Carmen Molina Molina
Consolación Abellán García
Davinia María Suárez Déniz
Diana Caparrós López
Encarni Meroño Sánchez
Félix Morales Hidalgo
Jacobo Chico Reyes
Laura Teruel Jordán
Luz María Amigo Ávila
María Ángeles Morales Carrillo
María Dolores Parra Córdoba
María Gema Muñoz de la Cruz
María Isabel Relaño Luque
María Josefa Campillo Cascales
María Luisa Palancar Belloso
María Márquez Martín
Marta Caride Vázquez
Marta Garcerán Candel
Mercedes Carazo Bueno
Núria Guzmán Sanjaume
Priscila Molina Gómez
Sara González Santos
Vanesa Pérez Aibar

Este

Aida Pons Lluch
Alba Ruiz Raventós
Alejandra Serrano Tomás
Ana Barambio Puig
Ana Belén García-Gascó Rodríguez
Ana María Pérez Pérez
Ana Sierra Álvarez
Anabel Pozo Reixach
Anaïs Barcelona Esteban
Clara Martín Muriana
Claudia Berenice Gómez Segovia
Cristina Leiro Sueiro
Ester Gil Giner
Eva María Sances Cara
Ludmila de Palma Marín
María Cristina Puebla Ledesma
Maria Dolores Benaloy Expósito
Meritxell Beltran Fernández
Natalia Bañó Alexandre
Neus Fàbregas Bofill
Noelia Lillo Moya
Rut Carolina Mizrahi Mass
Sandra Solbes Valero

Centro

Ana Sanz Toledo
Ángel Rosa Alcázar
Beatriz Serrano del Cerro
Carolina del Valle Estévez
Giovanna Falchi
Laura Gómez Díaz
Laura López Sánchez
Lucía Toledo
Luz María Rincón de la Flor
María de las Nieves Fernández Cespedes
María Pilar Usón Carreras
María Teresa Martínez Jiménez
Mónica Carabantes Navea
Nuria Rosell Martínez
Olga Díez Garcié
Pablo Yago Mendívil Nasarre
Raquel Cardona Gambaro
Sandra Rosero García

Por último no queremos olvidarnos de los profesionales Montse Alberte, Katarina Forssén, Emilie Gille, Maria Muñoz, Erica Paradell, Mario Pérez, Toni Rovira, Mireille Simon, Louis-Charles Vannier, así como de los miembros del equipo de diseño de Kurt Johnson y de psicometría de J. J. Zhu, que han trabajado incansablemente durante meses aportando su experiencia para ofrecer a los psicólogos españoles una nueva herramienta que, sin duda, contribuirá a mejorar la evaluación de la inteligencia de los niños y adolescentes.

Ana Hernández y Cristina Aguilar, psicólogas del Departamento de I+D

Frédérique Vallar, psicóloga y directora del Departamento de I+D

Índice

Agradecimientos de Pearson Clinical & Talent Assessment III

Ficha técnica XV

Prólogo XVII

Capítulo 1.
Introducción 1

Corrientes históricas y actuales de la evaluación de la inteligencia 1

Perspectiva histórica y contemporánea de las escalas de Wechsler 2

Capítulo 2.
Diseño y objetivo de la escala 5

Versiones anteriores 5

Contenido y estructura de la WISC-V 6

Pruebas 6

Cambios y modificaciones en el contenido de las pruebas 6

Descripción de las pruebas 7

Terminología y uso de las pruebas 12

Puntuaciones compuestas 13

Contenido de las puntuaciones compuestas 13

Terminología de los índices 14

Puntuaciones de procesamiento 14

Puntuaciones de procesamiento escalares 15

Puntuaciones de procesamiento directas 15

Estructura de la WISC-V 16

Escala total 16

Escalas primarias 16

Escalas secundarias 16

Objetivos de la revisión 18

Actualización de las bases teóricas 18

Modelos estructurales de la inteligencia 18

Investigación sobre el neurodesarrollo y la neurocognición 19

Modelos e investigación sobre la memoria de trabajo 24

Mejora de la adecuación al desarrollo 25

Consignas y redactado de los ítems 25

Criterios de puntuación ... 25
Bonificación por tiempo ... 26
Aumento de la facilidad de uso ... 26
Seguridad de los ítems ... 26
Materiales ... 26
Tiempo de aplicación ... 26
Instrucciones de aplicación y corrección ... 27
Reglas de terminación ... 27
Mejora de las propiedades psicométricas ... 27
Baremos y método de tipificación ... 27
Fiabilidad y validez ... 27
Suelo y techo de las pruebas ... 28
Sesgo de los ítems ... 28
Opciones del nivel de significación de los valores críticos ... 28
Incremento de la utilidad clínica ... 28
Ilustraciones e ítems ... 29
Estructura de la escala ... 29
Metodología de comparación de las diferencias entre puntuaciones ... 30
Requisitos de lenguaje expresivo ... 30
Puntuaciones de los índices secundarios ... 30
Puntuaciones de procesamiento ... 30
Grupos especiales ... 31

Capítulo 3.
Desarrollo de la versión estadounidense de la WISC-IV ... 33
Directrices de investigación ... 33
Principales fases de la investigación ... 33
Desarrollo conceptual ... 33
Equipo asesor ... 34
Investigación profesional ... 34
Estudios semiestructurados de expertos y examinadores ... 34
Estudios piloto ... 34
Experimentación ... 34
Tipificación ... 35
Composición final y evaluación de la WISC-V ... 35
Consistencia de los ítems, las consignas y los estímulos ... 35
Consistencia del orden de aplicación ... 35

Capítulo 4.
Adaptación española de la WISC-V ... 37
Fase experimental ... 37
Selección de los ítems experimentales ... 38
Selección y formación de los examinadores ... 39
Aplicación de la versión experimental ... 39

Control y seguimiento 39
Muestra experimental 40
Análisis de los datos de la fase experimental 41
Fase de tipificación 41
Modificaciones respecto a la WISC-IV 41
Diseño de la muestra de tipificación 42
Selección y formación de los examinadores 43
Aplicación de la versión de tipificación 43
Puntuación de los protocolos 44
Muestra de tipificación real 44
Representatividad de la muestra 45
Creación de los baremos 48
Establecimiento de los puntos de comienzo y de las reglas de terminación 48
Establecimiento de las bonificaciones por tiempo 48
Obtención de las puntuaciones escalares 49
Obtención de las puntuaciones compuestas 49
Obtención de las tasas base 50
Cálculo de los valores críticos para las diferencias entre puntuaciones 50
Puntos fuertes y débiles 50
Comparación entre índices, pruebas o puntuaciones compuestas 51
Cálculo de las puntuaciones de edad equivalente 51

Capítulo 5.
Fiabilidad 53
Fiabilidad y errores de medida 54
Consistencia interna 54
Errores típicos de medida e intervalos de confianza 58
Estabilidad temporal: test-retest 61
Acuerdo entre correctores 63
Resumen 63

Capítulo 6.
Validez 65
Evidencia basada en el contenido 65
Evidencia basada en los procesos de respuesta 66
Evidencia basada en la estructura interna 67
Estudios de intercorrelación 68
Intercorrelaciones de pruebas, puntuaciones de procesamiento y puntuaciones compuestas 68
Análisis factorial confirmatorio 72
Evidencias basadas en las relaciones con otras variables 80
Relaciones con otras medidas 80
Correlaciones con la WISC-IV 81
Correlaciones con la WPPSI-IV 82
Correlaciones con la WAIS-IV 84

Correlaciones con la KABC-II 86
Correlaciones con el Vineland-II 89
Correlaciones con el BASC-2 CPP 91
Grupos especiales 93
Niños con alta capacidad intelectual 95
Niños con discapacidad intelectual leve o moderada 98
Niños con capacidad intelectual límite 101
Niños con trastorno específico del aprendizaje 103
Niños con trastorno por déficit de atención con hiperactividad 110
Niños con comportamiento disruptivo 112
Niños con traumatismo craneoencefálico 113
Niños con trastorno del espectro autista 115
Resumen del rendimiento de los grupos especiales en la WISC-V 120
Consecuencias de la aplicación de tests 120

Capítulo 7.
Interpretación de los resultados 121
Descripción de los resultados 121
Puntuaciones típicas 122
Percentiles 124
Errores típicos de medida e intervalos de confianza 124
Clasificación cualitativa 124
Puntuaciones de edad equivalente 125
Diferencias entre puntuaciones 126
Significación estadística de las diferencias entre puntuaciones 126
Puntos fuertes y débiles 127
Comparación entre índices, pruebas o puntuaciones compuestas 128
Tasas base de las diferencias entre puntuaciones 128
Análisis de dispersión entre pruebas 128
Etapas para el análisis de un perfil 129
Realización del análisis primario de un perfil 129
Paso 1. Descripción del CI total 129
Paso 2. Descripción del ICV 130
Paso 3. Descripción del IVE 130
Paso 4. Descripción del IRF 130
Paso 5. Descripción del IMT 131
Paso 6. Descripción del IVP 131
Paso 7. Evaluación de los puntos fuertes y débiles en los índices 131
Paso 8. Evaluación de las comparaciones entre los índices 132
Paso 9. Evaluación de los puntos fuertes y débiles en las pruebas 135
Paso 10. Evaluación de las comparaciones entre las pruebas 135
Realización del análisis secundario de un perfil 137
Paso 1. Descripción del IRC 137
Paso 2. Descripción del IMTA 138

Paso 3. Descripción del INV 138
Paso 4. Descripción del ICG 139
Paso 5. Descripción del ICC 140
Paso 6. Evaluación de las comparaciones entre las puntuaciones compuestas de los índices secundarios 140
Paso 7. Evaluación de las comparaciones entre las pruebas de los índices secundarios 142
Paso 8. Realización del análisis de procesamiento 143
Realización del análisis cualitativo de las respuestas 147
Resumen 147

Anexo A.
Tablas de intercorrelaciones 149

Anexo B.
Criterios de inclusión para los grupos especiales 155

Anexo C.
Equipo de expertos, revisores y consultores de la versión estadounidense y española 159

Bibliografía 161

Tablas

Tabla 2.1. Pruebas nuevas y cambios en las pruebas de la WISC-V 7
Tabla 2.2. Abreviaciones y categorías de las pruebas 13
Tabla 2.3. Abreviaciones de las puntuaciones compuestas 14
Tabla 2.4. Abreviaciones de las puntuaciones de procesamiento 15
Tabla 2.5. Abreviaciones de las puntuaciones de span 16
Tabla 4.1. Ítems conservados, eliminados y nuevos respecto a la versión estadounidense para la versión experimental española 38
Tabla 4.2. Cambios en la adaptación española de la WISC-V respecto a la WISC-IV 42
Tabla 4.3. Tamaño de la muestra de tipificación por sexo y por grupo de edad 44
Tabla 4.4. Porcentajes teóricos (censo) y reales (muestra de tipificación), según edad y nivel educativo 45
Tabla 4.5. Porcentajes teóricos (censo) y reales (muestra de tipificación), según edad y tipo de población 46
Tabla 4.6. Porcentajes teóricos (censo) y reales (muestra de tipificación), según edad y sexo 46
Tabla 4.7. Porcentajes teóricos (censo) y reales (muestra de tipificación), según edad y zona geográfica 46
Tabla 5.1. Coeficientes de fiabilidad de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas (datos españoles) 55
Tabla 5.2. Coeficientes de fiabilidad de las pruebas y de las puntuaciones de procesamiento para los grupos especiales (datos españoles) 56
Tabla 5.3. Coeficientes de fiabilidad de las pruebas y de las puntuaciones de procesamiento para los grupos especiales (datos estadounidenses) 57

Tabla 5.4. Errores típicos de medida de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas (datos españoles) 59
Tabla 5.5. Datos demográficos del estudio de estabilidad test-retest y del estudio de acuerdo entre correctores (datos españoles) 61
Tabla 5.6. Coeficientes de estabilidad de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas (datos españoles) 62
Tabla 6.1. Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas, para todas las edades (datos españoles) 70
Tabla 6.2. Modelos del análisis factorial confirmatorio de la WISC-V (datos españoles) ... 74
Tabla 6.3. Índices de bondad de ajuste para el análisis factorial confirmatorio (datos españoles) 76
Tabla 6.4 Datos demográficos para los estudios de validez con otras medidas 80
Tabla 6.5. Correlaciones entre la WISC-V y la WISC-IV (datos españoles) 81
Tabla 6.6. Correlaciones entre la WISC-V y la WPPSI-IV (datos estadounidenses) 83
Tabla 6.7. Correlaciones entre la WISC-V y la WAIS-IV (datos estadounidenses) 85
Tabla 6.8. Correlaciones entre la WISC-V y la KABC-II (datos estadounidenses) 87
Tabla 6.9. Correlaciones entre la WISC-V y el Vineland-II (datos estadounidenses) 90
Tabla 6.10. Correlaciones entre la WISC-V y el BASC-2 CPP (datos estadounidenses) 93
Tabla 6.11. Datos demográficos de los grupos especiales 94
Tabla 6.12. Resultados del grupo con alta capacidad intelectual y del grupo control (datos españoles) 96
Tabla 6.13. Resultados del grupo con alta capacidad intelectual y del grupo control (datos estadounidenses) 97
Tabla 6.14. Resultados del grupo con discapacidad intelectual y del grupo control (datos españoles) 99
Tabla 6.15. Resultados del grupo con discapacidad intelectual leve y del grupo control (datos estadounidenses) 100
Tabla 6.16. Resultados del grupo con discapacidad intelectual moderada y del grupo control (datos estadounidenses) 101
Tabla 6.17. Resultados del grupo con capacidad intelectual límite y del grupo control (datos estadounidenses) 102
Tabla 6.18. Resultados del grupo con trastorno específico del aprendizaje-lectura y del grupo control (datos estadounidenses) 105
Tabla 6.19. Resultados del grupo con trastorno específico del aprendizaje-lectura y expresión escrita y del grupo control (datos estadounidenses) 107
Tabla 6.20. Resultados del grupo con trastorno específico del aprendizaje-matemáticas y del grupo control (datos estadounidenses) 109
Tabla 6.21. Resultados del grupo con TDAH y del grupo control (datos estadounidenses) 111
Tabla 6.22. Resultados del grupo con comportamiento disruptivo y del grupo control (datos estadounidenses) 113
Tabla 6.23. Resultados del grupo con traumatismo craneoencefálico y del grupo control (datos estadounidenses) 115
Tabla 6.24. Resultados del grupo con trastorno del espectro autista con deterioro del lenguaje, y del grupo control (datos estadounidenses) 117
Tabla 6.25. Resultados del grupo con trastorno del espectro autista sin deterioro del lenguaje, y del grupo control (datos estadounidenses) 119

Tabla 7.1. Relación entre las puntuaciones escalares, las desviaciones típicas respecto a la media y los percentiles equivalentes. 122
Tabla 7.2. Relación entre las puntuaciones compuestas, las desviaciones típicas respecto a la media y los percentiles equivalentes. 123
Tabla 7.3. Clasificación cualitativa de las puntuaciones compuestas 125
Tabla A.1 Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas para las edades 6:0-7:11. 150
Tabla A.2 Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas para las edades 8:0-9:11 151
Tabla A.3 Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas para las edades 10:0-11:11. 152
Tabla A.4 Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas para las edades 12:0-13:11. 153
Tabla A.5 Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas para las edades 14:0-16:11. 154

Figuras

Figura 2.1. Estructura de la WISC-V 17
Figura 4.1. Distribución de los examinadores de la fase de tipificación 45
Figura 4.2. Comparación entre los datos del censo y los datos de la muestra de tipificación según las características demográficas 47
Figura 6.1. Modelo jerárquico de cinco factores para todas las pruebas (6-16 años) (datos españoles). 78
Figura 6.2. Modelo jerárquico de cinco factores para las pruebas principales (6-16 años) (datos españoles). 79

Ficha técnica

Nombre del test:	*Escala de inteligencia de Wechsler para niños-V*
Nombre original:	*Wechsler Intelligence Scale for Children-Fifth Edition*
Autor:	David Wechsler
Adaptación española:	Departamento de I+D de Pearson Clinical & Talent Assessment: Ana Hernández, Cristina Aguilar, Erica Paradell, Frédérique Vallar
Editor original:	NCS Pearson Inc, 2014
Editor de la adaptación española:	Pearson Educación, 2015
Aplicación:	Individual
Ámbito de aplicación:	Niños de edades comprendidas entre 6 años 0 meses y 16 años 11 meses.
Tiempo de aplicación:	65 minutos aproximadamente para la aplicación de las diez pruebas principales y 48 minutos aproximadamente para la aplicación de las siete pruebas principales (cálculo del CI total).
Objetivo:	Evaluación completa de la aptitud intelectual del niño.
Tipificación:	Muestra representativa de la población española compuesta de 1008 niños, teniendo en cuenta los siguientes criterios de estratificación: edad, sexo, nivel educativo, zona geográfica y tipo de población.
Material:	*Manual técnico y de interpretación*, *Manual de aplicación y corrección*, cuadernillo de anotación, cuaderno de estímulos 1 y 2, cuadernillo de respuestas 1 y 2, plantilla de corrección para *Búsqueda de símbolos*, plantilla de corrección para *Claves*, plantilla de corrección para *Cancelación*, caja con 9 cubos bicolores, lápiz sin goma de borrar, lápiz rojo sin goma de borrar.
Nivel de cualificación requerido:	Nivel C; titulación superior en psicología, psiquiatría o psicopedagogía y experiencia profesional en diagnóstico clínico.
Corrección:	Manual o mediante plataforma on-line.

profundo de sus fundamentos conceptuales. Hace unos años el Colegio Oficial de Psicólogos (COP) hizo una encuesta a un numeroso grupo de profesionales.[6] Quienes respondieron hicieron una alta valoración sobre la utilidad de los tests psicológicos en su práctica cotidiana, pero también reconocieron que la formación universitaria recibida al respecto era mejorable. Esa carencia formativa convierte en imperativa la necesidad de estudiar cuidadosamente el material incluido en los detallados manuales técnicos que acompañan a las pruebas estandarizadas.

La adaptación española de la WISC-V se ha basado en una muestra representativa (según edad, sexo, nivel educativo y zona geográfica) de más de mil niños y adolescentes. Se debe subrayar la inversión y el esfuerzo que supone la preparación de los materiales de evaluación, la recogida de datos y su posterior análisis. La información sobre las garantías técnicas de la adaptación (fiabilidad, validez, etc.), que se incluye en el detallado manual, permite comprobar que se satisfacen los criterios más exigentes.

Además de su uso en la población general, la WISC-V resultará de utilidad para la evaluación de poblaciones especiales: niños con alta capacidad intelectual, con discapacidad intelectual, con trastornos del aprendizaje, con TDAH, con comportamiento disruptivo, con algunas clases de traumatismo craneoencefálico o con trastornos del espectro autista. La información del manual técnico facilita la confección de perfiles (puntos fuertes y débiles, comparación de índices, etc.) según la evaluación realizada. También se ofrecen pautas con respecto a la interpretación cualitativa.

Pearson Clinical and Talent Assessment España pone a disposición de los profesionales un excelente instrumento de evaluación de la capacidad intelectual para la población infantil y adolescente. El cuidado exquisito que se ha puesto en maximizar la calidad del producto es manifiesto durante la lectura del manual técnico, pero también se revela al consultar el equipo de revisores que figuran en los créditos, entre quienes encontramos a Nelson Cowan, Susan Gathercole, Alan S. Kaufman, Matthew R. Reynolds, Timothy Salthouse, W. Joel Schneider y Lisa Suzuki. Todos ellos son científicos de referencia en la investigación de la inteligencia y las capacidades cognitivas.

Escribía mi colega Tom Bouchard en el 2014 que «los números pueden no garantizar la verdad, pero el progreso en la ciencia es imposible sin números y sin medidas, puesto que la retórica y las palabras son insuficientes».[7] Trataba de desmontar, a mi juicio con rotundo éxito, la perspectiva de autores tristemente tan populares como Stephen J. Gould dirigida a denigrar la medida de la inteligencia.

Ninguna medida es perfecta. Pero nuestro mundo mejora cuando logramos medir con mayor precisión, cuando aumenta el grado en el que podemos fiarnos de esas medidas. La WISC-V contribuye a esa deseable empresa.

Roberto Colom

Universidad Autónoma de Madrid

2 de octubre de 2015

6 Muñiz, J. y Fernández-Hermida, J. R. (2010). La opinión de los psicólogos sobre el uso de los tests *Papeles del Psicólogo, 31*(1), 108-121.

7 Bouchard, T. (2014). Genes, Evolution and Intelligence. *Behavior Genetics, 44*(2), 549-577.

Capítulo 1

Introducción

La *Escala de inteligencia de Wechsler para niños-V* (WISC-V) es un instrumento clínico de aplicación individual para evaluar la inteligencia de los niños de 6 años y 0 meses a 16 años y 11 meses (6:0-16:11). Se trata de una versión revisada de la *Escala de inteligencia de Wechsler para niños-IV* (WISC-IV; Wechsler, 2003, 2011 [ed. orig., 2005]). La WISC-V ofrece puntuaciones de los índices primarios, que reflejan el funcionamiento intelectual en determinadas áreas cognitivas (comprensión verbal, visoespacial, razonamiento fluido, memoria de trabajo y velocidad de procesamiento); una puntuación compuesta que representa la aptitud intelectual general (CI total); y puntuaciones de los índices secundarios, que reflejan las aptitudes cognitivas en distintas agrupaciones basadas en necesidades clínicas (p. ej., no verbal y capacidad general).

Esta quinta edición de la escala forma parte de la historia evolutiva de la evaluación de la inteligencia y debe considerarse dentro del contexto de las corrientes históricas y actuales de dicha evaluación. En este capítulo se describen las corrientes de la evaluación de la inteligencia, así como el origen y la evolución de las escalas Wechsler.

Corrientes históricas y actuales de la evaluación de la inteligencia

A principios del siglo xx, la teoría dominante sobre la inteligencia defendía un único constructo de inteligencia general subyacente, en gran parte responsable del rendimiento del sujeto en todas las tareas mentales. Spearman (1904) identificó este constructo como *factor g*. Aunque su teoría de la inteligencia incluía otros factores específicos, los primeros tests de inteligencia pretendían clasificar a los sujetos según su nivel general de funcionamiento cognitivo (Beres, Kaufman y Perlman, 2000; Gottfredson y Saklofske, 2009; Wasserman, 2012).

En 1905, Binet y Simon, como respuesta a la legislación francesa que establecía la educación universal, publicaron una escala de inteligencia con el objetivo de desarrollar métodos que permitieran identificar a aquellos niños que podrían tener dificultades para seguir una educación normal. Continuando con las mejoras que Binet y Simon habían realizado a su escala, así como las llevadas a cabo por Terman a principios del siglo xx, el propio Terman y sus colegas de la Universidad de Stanford presentaron una versión tipificada de dicha escala en Estados Unidos (Terman, 1916).

La entrada de Estados Unidos en la primera guerra mundial generó la necesidad de crear medidas de inteligencia para evaluar y clasificar a los reclutas. Con este fin se creó el *Army Alpha*, un test que incluía un considerable componente verbal. Pero la alta tasa de analfabetismo de algunos reclutas hizo que fuera necesaria una medida no verbal de inteligencia, por lo que se desarrolló el *Army Beta* (Wasserman, 2012). Wechsler reconoció esta doble necesidad de evaluar la inteligencia de forma verbal y no verbal, y, al elaborar su primera escala de inteligencia, la *Escala de inteligencia de Wechsler-Bellevue* (Wechsler-Bellevue; Wechsler, 1939), incluyó puntuaciones tanto para la escala verbal como para la escala manipulativa, además de una puntuación total. La Wechsler-Bellevue también era innovadora porque proporcionaba las puntuaciones CI de desviación, basadas en puntuaciones típicas para todas las edades.

En la década de 1950, la evaluación de la inteligencia se centró en la medición de aspectos más específicos del funcionamiento cognitivo de los sujetos. Durante este periodo, los avances en las técnicas de análisis factorial se aplicaron al campo de la medición para aclarar aun más la natu-

raleza de la inteligencia. Cattell, que había sido alumno de Spearman, formuló la teoría de que la inteligencia se componía de dos factores generales: la inteligencia fluida (*Gf*) y la inteligencia cristalizada (*Gc*; Cattell, 1941, 1957). Más tarde, Horn amplió la teoría de Cattell e incluyó los siguientes factores: percepción visual, memoria a corto y largo plazo, velocidad de procesamiento, aptitud de procesamiento auditivo, razonamiento cuantitativo, lectura y escritura (Horn, 1985, 1988, 1991; Horn y Noll, 1997).

Mientras los investigadores identificaban dominios de inteligencia más específicos, la interpretación de los tests de inteligencia comenzó a centrarse en el rendimiento del sujeto en ámbitos más segmentados del funcionamiento cognitivo, así como en las posibles diferencias de desarrollo en esos dominios (Carroll, 1993; Goldstein y Hersen, 2000; Keith, 1985, 1990; Schneider, Schumann-Hengsteler y Sodian, 2005). A lo largo de los últimos setenta años, gran parte del debate sobre la evaluación de la inteligencia se ha centrado en la existencia de un aspecto global subyacente de la inteligencia, que influye sobre los diversos aspectos del rendimiento intelectual del sujeto (Gottfredson y Saklofske, 2009; Jensen, 1998; Reeve y Charles, 2008; Wasserman, 2012). Carroll (1993, 2012), basándose en los estudios realizados hasta la fecha que evaluaban las pruebas de aptitud cognitiva mediante el análisis factorial, concluyó que la evidencia de un factor general de inteligencia era incuestionable. Por tanto, la tendencia a considerar las aptitudes intelectuales como múltiples y más definidas no supone el rechazo a la existencia de un factor general de inteligencia (Gottfredson y Saklofske, 2009; Johnson, Bouchard, Krueger, McGue y Gottesman, 2004). La mayoría de las teorías sobre la inteligencia consideran que su estructura es jerárquica, con un factor general en el nivel superior y amplios dominios de aptitud cognitiva en los niveles siguientes; estos dominios, a su vez, divididos en aptitudes más específicas y segmentadas. Los resultados basados en los estudios del análisis factorial coinciden en la identificación de entre 8 y 10 grandes dominios de inteligencia cuando se usa un número suficientemente amplio de tareas distintas en el análisis (Carroll, 1993, 2012; Horn y Blankson, 2012; Johnson et al., 2004; Schneider y McGrew, 2012). Existe respaldo empírico en distintos tests para 4 o 5 de estos grandes dominios en varias muestras de niños y adolescentes en edad escolar (Benson, Hulac y Bernstein, 2013; Benson y Taub, 2013; Reynolds, Keith, Flanagan y Alfonso, 2013; Sanders, McIntosh, Dunham, Rothlisberg y Finch, 2007; Weiss, Keith, Zhu y Chen, 2013b).

Perspectiva histórica y contemporánea de las escalas de Wechsler

La historia de la *Escala de inteligencia de Wechsler para niños-V* (WISC-V) se inicia con la Wechsler-Bellevue. Wechsler basó este test en la premisa de que la inteligencia es global, porque caracteriza el comportamiento del sujeto como un todo, y específica, porque está compuesta por aptitudes o elementos distinguibles entre ellos. A partir de su experiencia clínica, Wechsler seleccionó y desarrolló pruebas destinadas a poner de relieve los aspectos cognitivos de la inteligencia que le parecían importantes medir: comprensión verbal, razonamiento abstracto, procesamiento visoespacial, razonamiento cuantitativo, memoria y velocidad de procesamiento. Las teorías y pruebas de inteligencia siguen confirmando que estos dominios son aspectos importantes de la aptitud cognitiva (Carroll, 1993, 2012; Horn y Blankson, 2012, Schneider y McGrew, 2012).

Algunos investigadores consideraron que Wechsler, dado que divide las pruebas en tareas verbales y manipulativas, asumió una estructura bifactorial de la inteligencia. Sin embargo, Wechsler (1958) expuso el objetivo práctico de tal división:

> [La agrupación de las pruebas en las áreas verbal y manipulativa] no implica que estas sean las únicas aptitudes abordadas en los tests. [...] Las pruebas son diferentes medidas de la inteligencia, no medidas de diferentes tipos de inteligencia; y la dicotomía de áreas verbal y manipulativa es solo una de las muchas formas posibles de clasificar los tests.[1]

En revisiones recientes de las escalas de inteligencia de Wechsler, se ha ampliado la medida de dominios específicos de la aptitud cognitiva (p. ej., visoespacial, razonamiento fluido, memoria de trabajo y velocidad de procesamiento), en atención a las teorías sobre la inteligencia, los estudios de análisis factorial y las investigaciones clínicas actuales, sin dejar de ofrecer una medida fiable de la inteligencia global (CI total).

1 Traducción de los adaptadores, cita original en Wechsler, 1958, p. 64.

Aunque la evaluación y división del funcionamiento cognitivo en dominios más específicos presenta ventajas, existen otros aspectos que deben tenerse en cuenta. Primero, las funciones cognitivas están interrelacionadas, funcional y neurológicamente, lo que dificulta la medición de un dominio puro del funcionamiento cognitivo. Incluso las medidas más tradicionales de dominios específicos, como la velocidad de procesamiento, implican que el sujeto tenga la capacidad de comprender la instrucción verbal, discriminar entre los estímulos visuales, procesar la información e indicar la respuesta mediante la intervención de la función motora. Pese a que los resultados de los estudios de análisis factorial sugieren la presencia o ausencia de los dominios evaluados, pueden no reflejar la variedad de aptitudes cognitivas requeridas para realizar una prueba concreta. Dado que las cargas factoriales de cada prueba difieren según la composición y combinación de las pruebas, lo que mide una prueba concreta varía en función de la combinación de pruebas incluidas en el análisis.

Segundo, es ecológicamente válido incluir pruebas que requieren la participación de varias aptitudes cognitivas para realizar correctamente las tareas. Las tareas cognitivas raramente se llevan a cabo de forma aislada. Como indicó Wechsler (1975):

> [...] los atributos y factores de la inteligencia, al igual que las partículas elementales en la física, tienen al mismo tiempo propiedades individuales y colectivas, es decir, parecen comportarse de forma distinta cuando están solas a como lo hacen cuando funcionan conjuntamente.[2]

La validez ecológica de la inteligencia general se demuestra por su capacidad de predecir cosas tales como el rendimiento laboral (Hunt y Madhyastha, 2012; Kuncel, Ones y Sackett, 2010; Schmidt, 2014), la salud mental y física y los hábitos de salud (Johnson, Corley, Starr y Deary, 2011; Wrulich et al., 2014), así como el rendimiento académico y el nivel educativo (Deary y Johnson, 2010; Deary, Strand, Smith y Fernandes, 2007; Johnson, Deary y Iacono, 2009; Kaufman, Reynolds, Liu, Kaufman y McGrew, 2012; Nelson, Canivez y Watkins, 2013). Algunos investigadores se han centrado en la capacidad predictiva del CI total en relación con el rendimiento académico (Glutting, Watkins, Konold y McDermott, 2006; Nelson et al., 2013); mientras que otros han hecho hincapié en la cognición basada en un dominio específico en relación con el rendimiento académico (Calhoun y Mayes, 2005; Fiorello, Hale y Snyder, 2006; Li y Geary, 2013; Mayes y Calhoun, 2004; Vukovic et al., 2014). Los dominios generales y específicos del funcionamiento cognitivo son importantes para comprender los retos que la vida académica y cotidiana plantea a los niños con discapacidades.

Tercero, no sería razonable pensar que una única medida de inteligencia pueda evaluar todos los dominios de forma significativa y práctica (Flanagan, Ortiz y Alfonso, 2007). Wechsler logró seleccionar medidas representativas de una amplia variedad de dominios cognitivos (p. ej., comprensión verbal, visoespacial, memoria), de los que se había demostrado su importancia como aspectos del funcionamiento cognitivo. También advirtió que durante una sesión de evaluación, la fatiga del examinador o del sujeto podía producir resultados inválidos, por ello seleccionó un número suficiente de pruebas que proporcionasen información clínicamente significativa en un tiempo de aplicación razonable. Wechsler pensaba que otros factores relacionados, como el rendimiento académico y las habilidades motoras, podían influir en el rendimiento del sujeto en los tests de inteligencia, pero consideraba que estos aspectos se evaluaban mejor con pruebas específicamente elaboradas para estos dominios. Su sensibilidad hacia los aspectos prácticos de la evaluación de la inteligencia se demuestra en el hecho de que elaboró una escala exclusivamente para la evaluación de la memoria (Wechsler, 1945).

Wechsler creía que los resultados de las pruebas de aptitud cognitiva solo reflejaban una parte de la inteligencia, a la que definió como la «capacidad del sujeto para actuar con determinación, para pensar racionalmente y para relacionarse eficazmente con el entorno».[3] Wechsler era muy consciente de que los resultados de los estudios factoriales representaban solo una parte de la inteligencia, y creía que había otros atributos que contribuían a la inteligencia. Entre estos

2 Traducción de los adaptadores, cita original en Wechsler, 1975, p. 138.

3 Traducción de los adaptadores, cita original en Wechsler, 1944, p. 3.

atributos figuran la planificación y conciencia de los objetivos, el entusiasmo, la dependencia o independencia de campo, la impulsividad, la ansiedad y la perseverancia.

La influencia de Wechsler en el campo de la evaluación de la inteligencia sigue presente en la actualidad. Las recomendaciones actuales para las buenas prácticas indican que, al interpretar los resultados de los tests, los profesionales han de considerar al sujeto como único y tener en cuenta otros atributos además de la inteligencia (Fiorello, Hale y Wycoff, 2012; Kaufman y Lichtenberger, 2006; Sattler, 2008a). Comúnmente, se acepta que los sujetos que obtienen puntuaciones similares en los tests pueden afrontar con distinta eficacia los retos planteados por el entorno, debido a aspectos no relacionados con sus aptitudes cognitivas. A la inversa, dado que algunos aspectos que no guardan relación con la inteligencia influyen en los resultados de los tests (Maroof, 2012), los sujetos con distintos niveles de aptitud intelectual pueden obtener puntuaciones similares. La evaluación de la inteligencia de un sujeto es más compleja que la simple obtención de sus puntuaciones en las pruebas de inteligencia. Como indicó Wechsler (1975):

> Lo que medimos con los tests no es lo que los tests miden: información, percepción espacial, capacidad de razonamiento. Esto es solo un medio para llegar a un fin. Lo que miden los tests de inteligencia es algo mucho más importante: la capacidad del sujeto para comprender el mundo que lo rodea, y sus recursos para afrontar los retos que se le presentan.[4]

Dado que las versiones actuales de las escalas de inteligencia de Wechsler miden tanto el funcionamiento intelectual global como los dominios específicos de la aptitud cognitiva, las bases teóricas de dichas escalas son evidentes. La existencia de una base teórica en las escalas de Wechsler se demuestra claramente en el hecho de que las mismas pruebas, o similares, están presentes en otras escalas de inteligencia, y por las elevadas correlaciones entre las escalas de inteligencia de Wechsler y otras pruebas de aptitud cognitiva. Muchas de las pruebas originales (p. ej., *Cubos* y *Vocabulario*) siguen estando presentes, con alguna modificación, en otras medidas actuales de aptitud intelectual que afirman estar basadas en las teorías de la inteligencia (p. ej., Kaufman y Kaufman, 2004; Roid, 2003). La validez de la escala también se demuestra en sus correlaciones con otras medidas de la aptitud cognitiva (p. ej., *Differential Ability Scales-Second Edition* [DAS-II; Elliott, 2007]; *Kaufman Assesssment Battery for Children, Second Edition* [KABC-II; Kaufman y Kaufman, 2004]; *Stanford-Binet Intelligence Scales, Fifth Edition* [Roid, 2003]; *Woodcock-Johnson IV* [Woodcock, Schrank, McGrew y Mather, en prensa]. Las altas correlaciones que se obtienen entre las escalas de inteligencia de Wechsler y otras pruebas de aptitud intelectual indican que estos instrumentos miden constructos similares.

La utilidad práctica y clínica de las escalas de inteligencia de Wechsler se ve avalada por más de setenta años de investigación en distintos campos y con finalidades muy diversas. Una y otra vez, estas escalas han demostrado su utilidad clínica en la identificación de la discapacidad intelectual y de discapacidades específicas de aprendizaje, asignación a programas especializados, intervención clínica y evaluación neuropsicológica (Beres et al., 2000). Dadas las numerosas demostraciones de su utilidad clínica, parece incuestionable que Wechsler tenía un verdadero concepto de la naturaleza de la inteligencia cuando seleccionó y desarrolló las pruebas de sus escalas.

Con esta exposición se ha ofrecido una perspectiva histórica y contemporánea de la teoría sobre la inteligencia en relación con la evaluación de la inteligencia en niños mediante la WISC-V. Para una revisión más completa sobre la evolución de la teoría de la inteligencia, véanse Flanagan y Harrison (2012), Goldstein y Beers (2003) y Sternberg (2000).

4 Traducción de los adaptadores, cita original en Wechsler, 1975, p. 139.

Capítulo 2

Diseño y objetivo de la escala

En los últimos setenta años, las escalas de inteligencia de Wechsler se han revisado frecuentemente para incorporar los avances realizados en el campo de la evaluación de la inteligencia, para actualizar los baremos según los cambios de la población, para actualizar el contenido de los ítems y adaptarlo a los cambios culturales, y para adaptarlas a las necesidades prácticas y clínicas de la sociedad contemporánea. Para analizar cómo se reflejan estos cambios en el diseño de la WISC-V, es importante considerar las versiones anteriores de la escala, en especial su contenido y estructura. En este capítulo se presentan, además, las mejoras realizadas en la WISC-V en cuanto a las características de las pruebas y el contenido y denominación de las puntuaciones compuestas y de procesamiento. Se describen con detalle las distintas puntuaciones, así como la estructura de la escala. Finalmente, se exponen los objetivos de la revisión, que se relacionan con los cambios llevados a cabo en el contenido y estructura.

Versiones anteriores

La primera versión de la *Escala de inteligencia de Wechsler para niños* (WISC; Wechsler, 1949, 1974b) estaba formada por doce pruebas: *Información, Aritmética, Semejanzas, Vocabulario, Dígitos, Comprensión, Figuras incompletas, Historietas, Cubos, Rompecabezas, Claves* —procedentes de la Wechsler-Bellevue (Wechsler, 1939) y adaptadas para su uso con niños— y *Laberintos* —creada especialmente para esta escala—. Las pruebas se clasificaron en dos escalas, la verbal y la manipulativa, y permitían obtener tres puntuaciones: CI verbal (CIV), CI manipulativo (CIM) y CI total (CIT).

En la siguiente versión, la *Escala de inteligencia de Wechsler para niños-Revisada* (WISC-R; Wechsler, 1974a, 1993), se conservaron las doce pruebas, pero se amplió el rango de edad, que pasó de 5-15 años a 6-16 años. No se introdujeron nuevas pruebas, y se mantuvieron las mismas puntuaciones: CIV, CIM y CIT.

Posteriormente se elaboró la *Wechsler Intelligence Scale for Children-Third Edition* (WISC-III; Wechsler, 1991; no adaptada a la población española), en la que se mantuvieron todas las pruebas de la WISC-R y se incorporó una prueba nueva, *Búsqueda de símbolos*, como medida de la velocidad de procesamiento. En cuanto a las puntuaciones, se mantuvieron las tres tradicionales: CIV, CIM y CIT, pero se añadieron cuatro puntuaciones de índices nuevas que representaban dominios más específicos del funcionamiento cognitivo: Índice de comprensión verbal (CV), Índice de organización perceptiva (OP), Índice de independencia de la distracción (ID) e Índice de velocidad de procesamiento (VP).

La revisión de esta versión dio lugar a la *Escala de inteligencia de Wechsler para niños-IV* (WISC-IV; Wechsler, 2003, 2011 [ed. orig., 2005]), en la que se eliminaron tres pruebas procedentes de la WISC-III (*Historietas, Rompecabezas* y *Laberintos*). El resto de pruebas (*Cubos, Semejanzas, Dígitos, Claves, Vocabulario, Comprensión, Búsqueda de símbolos, Figuras incompletas, Información* y *Aritmética*) se conservaron, y se revisó tanto el contenido de los ítems como los procedimientos de aplicación y corrección. Además, se crearon cinco pruebas nuevas: *Adivinanzas*, como medida de la comprensión verbal; *Matrices* y *Conceptos*, como medidas del razonamiento fluido; *Letras y números*, como medida de la memoria de trabajo; y *Animales*, como medida de la velocidad de procesamiento. Por otra parte, se eliminaron las puntuaciones CIV y CIM; y, para reflejar con más precisión las capacidades medidas, se cambió el nombre del Índice de organización perceptiva

(OP) y del Índice de independencia de la distracción (ID) por el de Índice de razonamiento perceptivo (RP) e Índice de memoria de trabajo (MT), respectivamente. Además, se incluyeron varias puntuaciones de procesamiento para ofrecer información más detallada acerca del rendimiento.

Contenido y estructura de la WISC-V

La WISC-V se distingue de su antecesora porque incorpora los nuevos avances en investigación sobre inteligencia, desarrollo cognitivo, neurodesarrollo, neurociencia cognitiva y procesos importantes para el aprendizaje. En esta sección se describe el contenido y la estructura de la WISC-V, prestando especial atención a las descripciones de las pruebas, de los índices y de las puntuaciones, así como a la organización de la escala.

Pruebas

Cambios y modificaciones en el contenido de las pruebas

La versión estadounidense de la WISC-V (Wechsler, 2014a) está compuesta de 21 pruebas, trece de las cuales proceden de la WISC-IV y ocho son pruebas nuevas. En la versión estadounidense se eliminaron dos pruebas de la WISC-IV: *Adivinanzas* y *Figuras incompletas*. La prueba *Adivinanzas* se eliminó porque su constructo se solapaba con el de *Vocabulario*. La prueba *Figuras incompletas* se suprimió para reducir la presencia de la velocidad en la escala y posibilitar la inclusión de otras medidas relevantes (p.ej., pruebas de razonamiento visoespacial, razonamiento fluido cuantitativo, memoria de trabajo). Además de la supresión e incorporación de pruebas, también se realizaron cambios en los procedimientos de aplicación y corrección de las pruebas conservadas. Es necesario que los examinadores se familiaricen con estos cambios a fin de seguir los criterios estandarizados de aplicación de la escala.

La adaptación española de la WISC-V (Wechsler, 2014) está compuesta de quince pruebas, doce de las cuales proceden de la WISC-IV: *Cubos, Semejanzas, Matrices, Dígitos, Claves, Vocabulario, Búsqueda de símbolos, Información, Letras y números, Animales* (denominada *Cancelación* en la WISC-V), *Comprensión* y *Aritmética;* y tres son pruebas nuevas: *Puzles visuales, Balanzas* y *Span de dibujos*. *Puzles visuales* es una prueba de Visoespacial adaptada de la prueba homónima de la *Escala de inteligencia de Wechsler para adultos-IV* (WAIS-IV, Wechsler, 2008, 2012b). *Balanzas* es una prueba de Razonamiento fluido cuantitativo adaptada de la prueba homónima de la WAIS-IV. *Span de dibujos* es una prueba de Memoria de trabajo diseñada para medir la memoria de trabajo visual y utiliza la interferencia proactiva y la secuenciación para crear los componentes de procesamiento cognitivo necesarios en las pruebas de memoria de trabajo.

En la adaptación española de la WISC-V, al igual que en la versión estadounidense, se eliminaron las pruebas de *Adivinanzas* y *Figuras incompletas* y además se eliminó la prueba *Conceptos*, que no contribuía a ningún índice y reemplazaba parcialmente a *Matrices* o *Balanzas*, para reducir el tiempo de aplicación.

La tabla 2.1 muestra las pruebas nuevas y los cambios realizados en las pruebas de la WISC-V respecto a la WISC-IV.

Tabla 2.1. Pruebas nuevas y cambios en las pruebas de la WISC-V

		Cambios en las pruebas		
	Prueba nueva	Consignas de aplicación	Registro y puntuación	Ítems nuevos
Semejanzas			✓	✓
Vocabulario			✓	✓
Información			✓	✓
Comprensión			✓	✓
Cubos			✓	✓
Puzles visuales	✓			
Matrices		✓		✓
Balanzas	✓			
Aritmética		✓		✓
Dígitos		✓	✓	✓
Span de dibujos	✓			
Letras y números		✓	✓	✓
Claves			✓	✓
Búsqueda de símbolos		✓	✓	✓
Cancelación				✓

Descripción de las pruebas

En las siguientes descripciones de las pruebas se expone la tarea que debe realizar el niño, así como los constructos que mide la prueba.

Pruebas de Comprensión verbal

Semejanzas

En esta prueba, el examinador lee dos palabras que representan objetos o conceptos comunes, y el niño debe describir en qué se parecen. El objetivo de esta prueba es medir el razonamiento verbal y la formación de conceptos. Además, mide la inteligencia cristalizada, el conocimiento léxico, la flexibilidad cognitiva, la comprensión auditiva, la memoria a largo plazo, el pensamiento asociativo y categórico, la capacidad de distinguir entre características esenciales y secundarias y la expresión verbal (Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Sattler, 2008b).

Esta prueba consta de 23 ítems: 11 son nuevos y 12 proceden de la WISC-IV (6 de ellos fueron ligeramente modificados). Además, se ha agregado un ítem de ejemplo. Para los 12 ítems conservados se revisaron los criterios de puntuación.

Vocabulario

Esta prueba está formada por ítems gráficos e ítems verbales. Para los ítems gráficos, el niño ha de nombrar los dibujos que le muestra el examinador. Para los ítems verbales, el niño ha de definir palabras que el examinador lee. El objetivo de esta prueba es medir el léxico del sujeto, así como la formación de conceptos verbales. Además mide la inteligencia cristalizada, los conocimientos adquiridos, la capacidad de aprendizaje, la expresión verbal, la memoria a largo plazo y el nivel de desarrollo del vocabulario. Otras capacidades que pueden medirse en esta prueba son la percepción y comprensión auditiva y el pensamiento abstracto (Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Sattler, 2008b).

Esta prueba consta de 29 ítems: 4 son ítems gráficos y 25 son ítems verbales. De los 18 ítems nuevos, 2 son gráficos y 16 son verbales. De los 11 ítems procedentes de la WISC-IV, 2 son gráficos y 9 son verbales. Para los 11 ítems conservados se revisaron los criterios de puntuación.

Información

En esta prueba, el niño debe responder a una serie de preguntas de conocimientos generales. El objetivo de esta prueba es medir la capacidad del niño para adquirir, almacenar y recuperar conocimientos referidos a hechos generales y, por tanto, se relaciona con la inteligencia cristalizada, la memoria a largo plazo y la capacidad para almacenar y recuperar la información del entorno en que se encuentra el niño y/o de la escuela. Otras habilidades que intervienen en esta prueba son la percepción, la comprensión y la expresión verbales (Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Sattler, 2008b).

Esta prueba consta de 31 ítems: 21 son nuevos y 10 proceden de la WISC-IV (2 de ellos fueron ligeramente modificados). Para los 10 ítems conservados se revisaron los criterios de puntuación.

Comprensión

En esta prueba, el niño debe responder a las preguntas que formula el examinador sobre normas de conducta y reglas sociales. El objetivo de esta prueba es medir el razonamiento y conceptualización verbal, la comprensión y expresión verbal, la capacidad para evaluar y utilizar la experiencia, así como para demostrar el conocimiento y juicio práctico del sujeto. También se relaciona con la inteligencia cristalizada, el conocimiento de las normas de conducta convencionales, el juicio social, la memoria a largo plazo y el sentido común (Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Sattler, 2008b).

Esta prueba consta de 19 ítems: 12 son nuevos y 7 proceden de la WISC-IV (5 de ellos fueron ligeramente modificados). Para los 7 ítems conservados se revisaron los criterios de puntuación.

Pruebas de Visoespacial

Cubos

En esta prueba, el niño debe reproducir, con cubos de uno o dos colores, un modelo construido por el examinador y/o una imagen mostrada en el cuaderno de estímulos. El niño dispone de un tiempo límite para realizar la tarea. El objetivo de esta prueba es medir la capacidad de analizar y sintetizar estímulos visuales abstractos. También se relaciona con el razonamiento y formación de conceptos no verbales, la inteligencia visual general, la percepción y organización visual, el procesamiento simultáneo, la coordinación visomotora, el aprendizaje y la capacidad de distinguir la figura y el fondo en estímulos visuales (Carroll, 1993; Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Sattler, 2008b).

Esta prueba consta de 13 ítems: 5 proceden de la WISC-IV y 8 son nuevos, incluido el último ítem (con forma de X), incorporado para ampliar el techo de la prueba.

Puzles visuales

En esta prueba, el examinador presenta al niño un puzle completado en el cuaderno de estímulos, y el niño debe seleccionar las tres opciones de respuesta que permiten reconstruir el puzle. El niño dispone de un tiempo límite para realizar la tarea. El objetivo de esta prueba es medir la capacidad mental de construcción no motora, en la que interviene el razonamiento visoespacial, la rotación mental, la memoria visual de trabajo, la comprensión de las relaciones parte-todo, y la capacidad de analizar y sintetizar los estímulos visuales abstractos. *Puzles visuales* es similar a otras tareas, como la prueba *Rompecabezas* de la *Escala de inteligencia de Wechsler para preescolar y primaria-IV* (WPPSI-IV, Wechsler, 2012a, 2014b) y el *Revised Minnesota Paper Form Board Test* (Likert y Quasha, 1995), diseñadas para medir la percepción visual, la inteligencia visual amplia, la inteligencia fluida, el procesamiento simultáneo, la visualización y manipulación espacial y la capacidad para anticipar relaciones entre las partes (Carroll, 1993; Groth-Marnat, 2009; Kaufman y Lichtenberger, 2006; Likert y Quasha, 1995; Sattler y Ryan, 2009). Esta prueba mide el procesamiento y la agudeza visuales, las relaciones espaciales, la integración y la síntesis de las relaciones parte-todo, el razonamiento no verbal y el aprendizaje por ensayo-error (Flanagan y Kaufman, 2009; Flanagan, Alfonso y Ortiz, 2012; Sattler y Ryan, 2009).

Puzles visuales es una prueba nueva, adaptada de la WAIS-IV, y consta de 29 ítems.

Pruebas de Razonamiento fluido

Matrices

El niño observa una matriz incompleta y debe seleccionar la opción de respuesta que completa la matriz. Esta tarea requiere que el niño utilice la información visoespacial para identificar la regla conceptual subyacente que relaciona todos los estímulos, y aplique el concepto subyacente para seleccionar la respuesta correcta. El objetivo de esta prueba es medir la inteligencia cristalizada, la inteligencia visual general, la capacidad espacial y la clasificación, el conocimiento de las relaciones parte-todo y el procesamiento simultáneo (Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Sattler, 2008b). Además, la tarea requiere atención a los detalles visuales y memoria de trabajo. Los distintos tipos de ítems que formaban esta prueba en la WISC-IV se redujeron para que el aprendizaje de la tarea sea más eficaz.

Esta prueba consta de 32 ítems: 20 son nuevos y 12 proceden de la WISC-IV. Los 2 ítems de ejemplo son nuevos.

Balanzas

En esta prueba, el niño observa, en el cuaderno de estímulos, una balanza con dos platillos y debe seleccionar la pesa que falta para equilibrar los dos platillos. El niño dispone de un tiempo límite para realizar la tarea. Esta prueba requiere que el niño aplique el concepto cuantitativo de igualdad para comprender la relación entre los objetos y aplicar los conceptos de correspondencia, suma y/o multiplicación para identificar la respuesta correcta. El objetivo de esta prueba es medir el razonamiento fluido y la inducción (Flanagan, Alfonso y Ortiz, 2012; Flanagan y Kaufman, 2009; Sattler y Ryan, 2009). Las tareas de razonamiento cuantitativo implican procesos de razonamiento que pueden expresarse matemáticamente, y se basan principalmente en la lógica inductiva o deductiva (Carroll, 1993). Aunque esta prueba se relaciona con la memoria de trabajo, su implicación es reducida en comparación con las tareas cuantitativas clásicas (como *Aritmética*), puesto que los estímulos se presentan visualmente en el cuaderno de estímulos.

Balanzas es una prueba nueva, adaptada de la WAIS-IV, y consta de 34 ítems.

Aritmética

En esta prueba, el niño debe resolver mentalmente una serie de problemas aritméticos y dar la respuesta en un tiempo límite. Tanto los ítems gráficos como verbales que componen la prueba requieren manipulación de las representaciones mentales, concentración, atención, memoria a corto y largo plazo, capacidad de razonamiento numérico y agilidad mental. Otras capacidades que también pueden intervenir en esta tarea son el procesamiento secuencial; el razonamiento fluido, cuantitativo y lógico; y el conocimiento cuantitativo (Groth-Marnat, 2009; Kaufman y Lichtenberger, 2006; Sattler, 2008b). Además, esta tarea requiere que el procesamiento lingüístico y auditivo estén intactos, incluidas la discriminación y la comprensión auditivas, y en menor grado la expresión verbal.

Esta prueba consta de 34 ítems: 24 de ellos proceden de la WISC-IV y 10 son nuevos. De los ítems conservados, 18 se modificaron sustancialmente y 6 se mantienen sin cambios. Los 18 ítems modificados implican realizar los mismos cálculos numéricos que en la versión anterior de la escala, pero su contenido se actualizó y ahora es más adecuado al desarrollo de los niños. Estos ítems se revisaron, cuando era necesario, para incrementar la intervención de la memoria de trabajo, sin alterar el conocimiento matemático requerido. Los 10 ítems nuevos se incorporaron para mejorar el suelo y el techo de la prueba y la creciente dificultad de los ítems. En esta versión, el examinador debe dar consignas explícitas al niño acerca de la repetición. En los ítems más largos, si es necesario repetirlos, se pausa el cronómetro para controlar la variabilidad en la velocidad de habla del examinador.

Pruebas de Memoria de trabajo

Dígitos

En esta prueba, el examinador lee una serie de números y el niño debe repetir los números en el mismo orden (tarea de orden directo), en orden inverso (tarea de orden inverso) o en orden creciente (tarea de orden creciente).

El paso de una tarea a otra requiere flexibilidad cognitiva y agilidad mental. En todas las tareas interviene el registro de la información, la atención enfocada, la discriminación auditiva y la repetición auditiva. La tarea de *Dígitos* en orden directo mide la repetición auditiva y la capacidad de almacenamiento temporal de la memoria de trabajo. En *Dígitos* en orden inverso interviene la memoria de trabajo, la transformación de la información, la manipulación de las representaciones mentales y la imaginación visoespacial (Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Reynolds, 1997; Sattler, 2008b).

Dígitos en orden creciente es una tarea nueva incorporada a la WISC-V y adaptada de la WAIS-IV. Se asemeja a otras tareas cuyo objetivo es medir la memoria de trabajo y la manipulación de las representaciones mentales (MacDonald, Almor, Henderson, Kempler y Andersen, 2001; Werheid et al., 2002). Se ha incluido para incrementar los requisitos de complejidad cognitiva de la prueba. Las tareas de *Dígitos* en orden inverso y en orden creciente requieren seriar de nuevo la información; la principal diferencia entre ellas es el establecimiento de la serie. En la tarea en orden inverso, el niño debe conservar, en la memoria de trabajo, la posición que los números ocupan en la serie, para poder repetirlos correctamente. En la tarea en orden creciente, debe conservar, en la memoria de trabajo, el valor de cada número y compararlo con el valor del número anterior y el valor del número posterior en la serie; de forma que el niño no sabe qué posición ocupará cada número en la serie hasta que se le han leído todos los números.

La puntuación directa de *Dígitos* se obtiene a partir de las puntuaciones directas de las tareas en orden directo, en orden inverso y en orden creciente. *Dígitos* en orden directo debe aplicarse siempre, dado que, de lo contrario, algunos niños pueden obtener una puntuación más baja en *Dígitos* en orden inverso (posiblemente debido a que no se han beneficiado de un entrenamiento progresivo). La aplicación de *Dígitos* en orden directo garantiza, además, un número suficiente de ítems de baja dificultad para los niños con discapacidad intelectual o baja capacidad cognitiva.

Para cada ítem de *Dígitos* se ofrecen dos intentos distintos, ambos de la misma longitud; y cada tarea está formada por 9 ítems. Se han creado varios intentos nuevos para equilibrar la distribución de los números (0-9). En *Dígitos* en orden directo se incluyó un ítem nuevo de 10 números para ampliar el techo de la tarea; y en *Dígitos* en orden inverso se incluyó un ítem nuevo para suavizar el incremento de la dificultad de los ítems de la tarea. De los intentos de *Dígitos* en orden directo, 17 son nuevos y 1 procede de la WISC-IV; y de los intentos de *Dígitos* en orden inverso, 16 son nuevos y 2 proceden de la WISC-IV. De los dos ítems de ejemplo de *Dígitos* en orden inverso, uno es nuevo y el otro procede de la WISC-IV. Todos los intentos de *Dígitos* en orden creciente son nuevos.

Span de dibujos

En esta prueba, el niño observa una página de estímulos con uno o más dibujos (de objetos de fácil denominación) durante un tiempo límite, y luego debe seleccionar (en orden secuencial, si es posible) los dibujos que ha visto de entre las opciones de una página de respuesta.

El objetivo de esta prueba es medir la memoria de trabajo visual y la capacidad de memoria de trabajo. Otras tareas similares también requieren la atención, el procesamiento visual, la memoria visual inmediata y la inhibición de la respuesta (Flanagan, Alfonso y Ortiz, 2012; Flanagan, Alfonso, Ortiz y Dynda, 2010; Miller, 2010, 2013). Esta prueba se asemeja a otras tareas de memoria de trabajo visual (Hartshorne, 2008; Makovski y Jiang, 2008; Wechsler, 2009a, 2013), pero es relativamente innovadora porque los estímulos son semánticamente significativos. Asimismo el uso de estos estímulos también puede activar la memoria de trabajo verbal. Los requisitos de esta prueba se deben a que se basa en la interferencia proactiva así como en la secuenciación.

Span de dibujos es una prueba nueva, creada para la WISC-V, y consta de 26 ítems.

Letras y números

En esta prueba, el examinador lee una combinación de letras y números, y el niño debe repetirla diciendo primero los números en orden creciente y luego las letras en orden alfabético. Al igual que *Dígitos*, esta prueba requiere ciertos procesos cognitivos básicos, como la discriminación auditiva, la atención focalizada, la concentración, el registro y la repetición auditiva. Además interviene el procesamiento secuencial, la capacidad de comparar estímulos basados en criterios cuantitativos o alfabéticos, la memoria de trabajo y la manipulación de las representaciones mentales. Otras capacidades que también pueden intervenir en ella son el procesamiento de la información, la flexibilidad cognitiva y la inteligencia fluida (Crowe, 2000; Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Sattler, 2008b). Las habilidades de nivel superior representan el control ejecutivo y las funciones de asignación de recursos en la memoria de trabajo.

Esta prueba consta de 10 ítems, cada uno con tres intentos distintos. Se revisaron los intentos para mejorar la dificultad creciente de la prueba y distribuir uniformemente los números y letras. También se eliminó la aparición de la letra U y la cifra 1 en un mismo intento, a fin de evitar respuestas ambiguas por parte de niños catalanohablantes. Por otra parte, también se revisaron las consignas con el objetivo de que la tarea se enseñe progresivamente (es decir, primero, el niño debe repetir los números y después las letras; y, segundo, debe repetir los números en orden creciente y luego las letras en orden alfabético). Por consiguiente, esta prueba incluye dos nuevos ítems de demostración y dos nuevos ítems de ejemplo. Además, se ofrecen dos ítems prerrequisito, que permiten determinar si el niño dispone ya de las habilidades de contar y de alfabetizar necesarias para realizar la tarea.

Esta prueba está formada por 30 intentos, de los cuales 26 son nuevos y 4 proceden de la WISC-IV. De los 3 intentos que componen los ítems de ejemplo, 2 son nuevos y 1 procede de la WISC-IV. Los 2 ítems de demostración son nuevos, y los 2 ítems prerrequisito proceden de la WISC-IV.

Pruebas de Velocidad de procesamiento

Claves

En esta prueba, el niño debe copiar, en un tiempo límite, una serie de símbolos que aparecen emparejados cada uno a una figura geométrica o a un número, de acuerdo con una clave. Además de la velocidad de procesamiento, esta prueba mide la memoria visual a corto plazo, la capacidad de aprendizaje procedimental e incidental, la velocidad psicomotora, la percepción visual, la coordinación visomotora, la capacidad de selección visual, la flexibilidad cognitiva, la atención, la concentración y la motivación. También pueden intervenir el procesamiento secuencial visual y la inteligencia fluida (Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Sattler, 2008b).

Claves se revisó para que la dificultad de los ítems fuera más equitativa a lo largo de la prueba. Así, se controló la presentación de las formas y los números para que aparecieran dos veces en cada fila. También se revisaron los símbolos para mejorar la objetividad de la puntuación (se eliminaron los pares de símbolos rotados). En los criterios de puntuación se especifica el grado aceptable de rotación de los símbolos dibujados por el niño. Además, se simplificaron las consignas.

La parte A de esta prueba consta de 75 ítems; la parte B, de 117 ítems. En la parte A se usan 5 figuras geométricas y 5 símbolos. De las figuras geométricas, 3 proceden de la WISC-IV y 2 se modificaron; y de los símbolos, 4 son nuevos y 1 procede de la WISC-IV. La parte B de esta prueba consta de 9 símbolos: 6 son nuevos y 3 se modificaron.

Búsqueda de símbolos

En esta prueba, el niño observa dos grupos de símbolos (un grupo clave y un grupo de búsqueda) y debe indicar, en un tiempo límite, si alguno de los símbolos del grupo clave es igual a los del grupo de búsqueda. Además de las capacidades visoperceptiva (p. ej., identificación visual y correspondencia) y de velocidad de toma de decisiones, esta prueba se relaciona con la memoria visual a corto plazo, la coordinación visomotora, el control inhibitorio, la discriminación visual, la velocidad psicomotora, la atención sostenida y la concentración. También puede medir la organización perceptiva, la inteligencia fluida, y la planificación y el aprendizaje (Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Sattler, 2008b).

Una de las modificaciones realizadas en esta prueba fue la inclusión, en el grupo de búsqueda, de un símbolo con características similares a las del símbolo objetivo o de una versión del símbolo objetivo rotada 90, 180 o 270 grados. Por otra parte, también se han simplificado las consignas. Asimismo, para que la dificultad de los ítems fuera más equitativa y para aumentar el control de los aciertos por azar, se distribuyó equilibradamente la posición de los símbolos clave a lo largo de la prueba; asimismo se controló la posición y ocurrencia de los símbolos similares y rotados tanto en los ítems que contienen el símbolo clave correspondiente como en los que no lo contienen. Además, se redujo el número de ítems que figuran en cada página y se aumentó un poco el tamaño de los símbolos.

La parte A de esta prueba consta de 40 ítems; la parte B, de 60 ítems. Todos los ítems son nuevos.

Cancelación

En esta prueba, el niño debe buscar y marcar unos estímulos objetivo (animales) de entre un conjunto de estímulos distribuido de forma estructurada o de forma aleatoria, en un tiempo límite. El objetivo de esta prueba es medir la velocidad al realizar la prueba, la velocidad de procesamiento visual y de toma de decisiones, el rastreo visual, y el reconocimiento y la discriminación visoperceptiva (Flanagan y Kaufman, 2009; Sattler, 2008b). También se relaciona con la atención, la concentración y el recuerdo visual (Sattler, 2008a). Las pruebas de cancelación se han usado muy frecuentemente en el campo neuropsicológico como medidas de negligencia visual, inhibición de respuesta y perseveración motora (Lezak, Howieson y Loring, 2004).

Esta prueba se revisó significativamente para equilibrar la dificultad de los cuadrantes de la página en lo que a la distribución de los objetos se refiere. Muchos de los estímulos objetivo y de los distractores son nuevos o se modificaron para crear ocho grupos de estímulos de forma y color más similares. Los dibujos que representaban objetos y animales poco familiares para otras culturas se sustituyeron por objetos y animales más universales. Además, se distribuyeron más sistemáticamente los ocho grupos de estímulos en las filas y columnas, así como los estímulos objetivos y los distractores en cada uno de los cuadrantes de la página. Por otra parte, se simplificaron las consignas.

Esta prueba consta de 2 ítems: el ítem 1 presenta los dibujos de forma aleatoria; y el ítem 2, de forma estructurada. Ambos ítems se revisaron.

Terminología y uso de las pruebas

Las pruebas de la WISC-V pueden agruparse en dos categorías: principales u opcionales (véase la tabla 2.2). Se recomienda aplicar las diez pruebas principales para realizar una evaluación y descripción completas de la capacidad intelectual. Las cinco pruebas opcionales pueden aplicarse junto con las pruebas principales para obtener una representación más amplia del funcionamiento intelectual, así como información adicional para la toma de decisiones clínicas.

Las diez pruebas principales se usan en distintas combinaciones para obtener el CI total, las puntuaciones de los índices primarios y tres de las cinco puntuaciones de los índices secundarios. Para obtener el CI total, se utilizan siete de las diez pruebas principales. En el apartado «Terminología de los índices» de este capítulo se ofrece más información sobre las puntuaciones de los índices primarios y secundarios.

Cuando no se dispone del resultado de una prueba principal o la prueba se considera inválida, puede utilizarse como sustituta una prueba opcional o cualquiera de las tres pruebas principales que no se utilizan para obtener el CI total; (*Puzles visuales*, *Span de dibujos* y *Búsqueda de símbolos*). Dado que la sustitución de pruebas puede incrementar el error de medida del CI total, se han establecido unas limitaciones al respecto, que se describen en el capítulo 2 del *Manual de aplicación y corrección*. Para obtener dos de las cinco puntuaciones de los índices secundarios, se usan algunas combinaciones de pruebas opcionales y principales.

En la WISC-V, la única sustitución permitida es para obtener el CI total. Esta nueva versión de la escala ofrece más puntuaciones compuestas, lo que resulta útil en determinadas situaciones, cuando no se puede obtener el CI total. Por ello, la necesidad de realizar sustituciones se reduce

considerablemente respecto a las versiones anteriores. En la sección «Puntuaciones compuestas» de este capítulo figura más información acerca de las nuevas puntuaciones compuestas.

La tabla 2.2. presenta el nombre y la abreviación de las pruebas de la WISC-V, así como la categoría a la que pertenecen, con la indicación «(CI total)» para señalar las pruebas principales que se utilizan para obtener esta puntuación compuesta.

Tabla 2.2. Abreviaciones y categorías de las pruebas

Prueba	Abreviación	Categoría
Semejanzas	S	Principal (CI total)
Vocabulario	V	Principal (CI total)
Información	I	Opcional
Comprensión	CO	Opcional
Cubos	C	Principal (CI total)
Puzles visuales	PV	Principal
Matrices	M	Principal (CI total)
Balanzas	B	Principal (CI total)
Aritmética	A	Opcional
Dígitos	D	Principal (CI total)
Span de dibujos	SD	Principal
Letras y números	LN	Opcional
Claves	CL	Principal (CI total)
Búsqueda de símbolos	BS	Principal
Cancelación	CA	Opcional

Puntuaciones compuestas

Contenido de las puntuaciones compuestas

Las puntuaciones compuestas incluyen el CI total y los índices, y solo pueden obtenerse después de haber aplicado y puntuado las pruebas que las conforman. Las once puntuaciones compuestas permiten describir el rendimiento del niño en dominios del funcionamiento cognitivo más amplios que los medidos por las pruebas. Las puntuaciones compuestas de la WISC-V y sus abreviaciones se presentan en la tabla 2.3.

Tabla 2.3. Abreviaciones de las puntuaciones compuestas

Puntuación compuesta	Abreviación
Índice de comprensión verbal	ICV
Índice visoespacial	IVE
Índice de razonamiento fluido	IRF
Índice de memoria de trabajo	IMT
Índice de velocidad de procesamiento	IVP
CI total	CIT
Índice de razonamiento cuantitativo	IRC
Índice de memoria de trabajo auditiva	IMTA
Índice no verbal	INV
Índice de capacidad general	ICG
Índice de competencia cognitiva	ICC

Terminología de los índices

Los diez índices disponibles en la WISC-V pueden dividirse en dos categorías: primarios y secundarios. Los índices primarios son puntuaciones compuestas basadas en los factores que habitualmente se obtienen en una evaluación completa de la capacidad cognitiva, y en esta escala son cinco: el Índice de comprensión verbal, el Índice visoespacial, el Índice de razonamiento fluido, el Índice de memoria de trabajo y el Índice de velocidad de procesamiento.

Los índices secundarios son los cinco siguientes: el Índice de razonamiento cuantitativo, el Índice de memoria de trabajo auditiva, el Índice no verbal, el Índice de capacidad general y el Índice de competencia cognitiva. Son puntuaciones compuestas que se obtienen a partir de combinaciones de pruebas principales o de puntuaciones principales y opcionales, y que proporcionan información adicional sobre las capacidades cognitivas de un niño y sobre su rendimiento en la escala. Por ejemplo, el Índice de capacidad general proporciona una estimación de la capacidad general sin incidir demasiado en la memoria de trabajo ni en la velocidad de procesamiento.

En el capítulo 7 de este manual se ofrece información sobre la interpretación de las puntuaciones de los índices.

Sustitución del Índice de razonamiento perceptivo

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). Este cambio proporciona más puntuaciones relacionadas con las capacidades visoespaciales y de razonamiento fluido y, además, permite transmitir con más claridad los resultados obtenidos por un niño a otros profesionales, a los profesores y a los padres o cuidadores.

Puntuaciones de procesamiento

El enfoque de procesamiento de Boston, desarrollado por Edith Kaplan (1988), destaca que la evaluación cualitativa del rendimiento en un test (análisis de los errores, registro de las observaciones conductuales y evaluación de los límites del sujeto) es tan importante como la evaluación cuantitativa de las puntuaciones. Las revisiones recientes de las escalas de inteligencia de Wechsler, como la WISC-IV, la WAIS-IV y la WPPSI-IV, incluyen puntuaciones de procesamiento para proporcionar información adicional acerca del rendimiento del sujeto en una prueba. Las abreviaciones de las puntuaciones de procesamiento de la WISC-V están formadas por la abreviación de la prueba (en mayúscula) y la abreviación específica de cada puntuación (en minúscula). En esta sección se describen los dos tipos de puntuaciones de procesamiento, así como su contenido.

Puntuaciones de procesamiento escalares

Al igual que las puntuaciones de las pruebas, las puntuaciones de procesamiento escalares se obtienen a partir de las puntuaciones directas, que, en la mayoría de los casos, son la suma de las puntuaciones de los ítems. Sin embargo, la forma de puntuar los ítems puede ser distinta del procedimiento habitual (p. ej., en *Cubos* se pueden puntuar los cubos colocados correctamente en vez de puntuar solo la construcción totalmente correcta), o la puntuación directa puede obtenerse solo de parte de las puntuaciones de los ítems (p. ej., solo los ítems de *Dígitos* en orden directo).

La WISC-V ofrece siete puntuaciones de procesamiento escalares, correspondientes a las tres pruebas siguientes: *Cubos*, *Dígitos* y *Cancelación*. Estas puntuaciones ofrecen una información más detallada de los procesos cognitivos que intervienen en el rendimiento del niño en las distintas pruebas. Las puntuaciones de procesamiento escalares no pueden sustituirse por ninguna puntuación de las pruebas ni contribuyen al CI total. El *Manual de aplicación y corrección*, en el apartado «Puntuación» correspondiente a cada prueba, ofrece información sobre cómo calcular estas puntuaciones de procesamiento.

En la tabla 2.4 se presentan las puntuaciones de procesamiento escalares y sus abreviaciones.

Tabla 2.4. Abreviaciones de las puntuaciones de procesamiento

Puntuación de procesamiento	Abreviación
Cubos sin bonificación por tiempo	Cs
Cubos parcial	Cp
Dígitos en orden directo	Dd
Dígitos en orden inverso	Di
Dígitos en orden creciente	Dc
Cancelación aleatoria	CAa
Cancelación estructurada	CAe

La puntuación de procesamiento Cs se basa en el rendimiento del niño en la prueba *Cubos* sin puntos de bonificación por tiempo en la resolución de los ítems. La puntuación de procesamiento Cp se basa en el número total de cubos correctamente colocados en todos los ítems, con puntos de bonificación por tiempo en determinados ítems cuando corresponde. Las puntuaciones de procesamiento Dd, Di y Dc reflejan el rendimiento de un niño en las tres tareas de la prueba *Dígitos* y ofrecen información sobre la variabilidad relativa del rendimiento a lo largo de las tres tareas. Las puntuaciones de procesamiento CAa y CAe indican el rendimiento del niño en los ítems 1 y 2, respectivamente, de la prueba *Cancelación*.

En el capítulo 7 de este manual se ofrece más información sobre las bases teóricas y la interpretación de las puntuaciones de procesamiento escalares.

Puntuaciones de procesamiento directas

La WISC-V ofrece otras puntuaciones de procesamiento, denominadas *puntuaciones de procesamiento directas* o *puntuaciones de span*, que indican el rendimiento máximo en una tarea de span (p. ej., en *Dígitos* en orden directo, el total de números repetidos correctamente en el último intento) A diferencia de las puntuaciones directas, las puntuaciones de procesamiento directas no se obtienen sumando las puntuaciones de los ítems.

Estas puntuaciones de procesamiento directas o puntuaciones de span reflejan aspectos del rendimiento del niño en las pruebas *Dígitos*, *Span de dibujos* y *Letras y números*. Aunque a menudo estas puntuaciones son consistentes con las puntuaciones escalares obtenidas en la prueba correspondiente, las puntuaciones de span pueden reflejar mejor el rendimiento máximo del niño que las puntuaciones escalares cuando el rendimiento varía a lo largo de ítems de longitud similar (p. ej., cuando el niño obtiene puntuaciones de 1 y 0 puntos en intentos de un mismo ítem o de ítems distintos de longitud similar). El *Manual de aplicación y corrección*, en el apartado «Puntua-

ción» correspondiente a cada una de las tres pruebas, ofrece información sobre cómo calcular las puntuaciones de span. En la tabla 2.5 se presentan las puntuaciones de span y sus abreviaciones.

Tabla 2.5. Abreviaciones de las puntuaciones de span

Puntuación de span	Abreviación
Span de *Dígitos* en orden directo	SpanDd
Span de *Dígitos* en orden inverso	SpanDi
Span de *Dígitos* en orden creciente	SpanDc
Span de estímulos de *Span de dibujos*	SpaneSD
Span de respuestas de *Span de dibujos*	SpanrSD
Span de *Letras y números*	SpanLN

Estructura de la WISC-V

La WISC-V se organiza en tres niveles de interpretación: escala total, escalas primarias y escalas secundarias. Cada nivel de interpretación se compone, a su vez, de escalas (p. ej., Escala total, escala Comprensión verbal, escala No verbal) formadas por pruebas, que se utilizan para obtener la información normativa de la puntuación compuesta. La figura 2.1 representa la estructura de la WISC-V. En el capítulo 7 de este manual se ofrece la información interpretativa para todas las puntuaciones.

Escala total

La Escala total está formada por cinco dominios: Comprensión verbal, Visoespacial, Razonamiento fluido, Memoria de trabajo y Velocidad de procesamiento. En la figura 2.1, las siete pruebas que contribuyen al CI total figuran en verde y las pruebas que pueden usarse como sustitutas figuran en negro y cursiva. El capítulo 2 del *Manual de aplicación y corrección*, en el apartado «Sustitución de las pruebas», incluye las normas de sustitución de pruebas. La información normativa del CI total se encuentra en el anexo A del *Manual de aplicación y corrección.*

Escalas primarias

La WISC-V ofrece cinco escalas primarias: Comprensión verbal, Visoespacial, Razonamiento fluido, Memoria de trabajo y Velocidad de procesamiento. Los índices que se obtienen de las escalas primarias son ICV, IVE, IRF, IMT e IVP, respectivamente. En la figura 2.1 figuran en verde las pruebas principales que contribuyen a cada índice. 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. **No se permite sustituir ninguna prueba de los índices primarios.** La información normativa de los índices primarios se encuentra en el anexo A del *Manual de aplicación y corrección.*

Escalas secundarias

La WISC-V ofrece cinco escalas secundarias: Razonamiento cuantitativo, Memoria de trabajo auditiva, No verbal, Capacidad general y Competencia cognitiva. Los índices que se obtienen de las escalas primarias son IRC, IMTA, INV, ICG e ICC, respectivamente. Los índices secundarios son puntuaciones compuestas que se obtienen a partir de combinaciones de pruebas principales o de puntuaciones principales y opcionales, y que proporcionan información adicional sobre las capacidades cognitivas de un niño y sobre su rendimiento en la escala. En la figura 2.1 figuran en verde las pruebas principales o las pruebas principales y opcionales que contribuyen a cada índice. **No se permite sustituir ninguna prueba de los índices secundarios.** La información normativa de los índices secundarios se encuentra en el anexo C del *Manual de aplicación y corrección.*

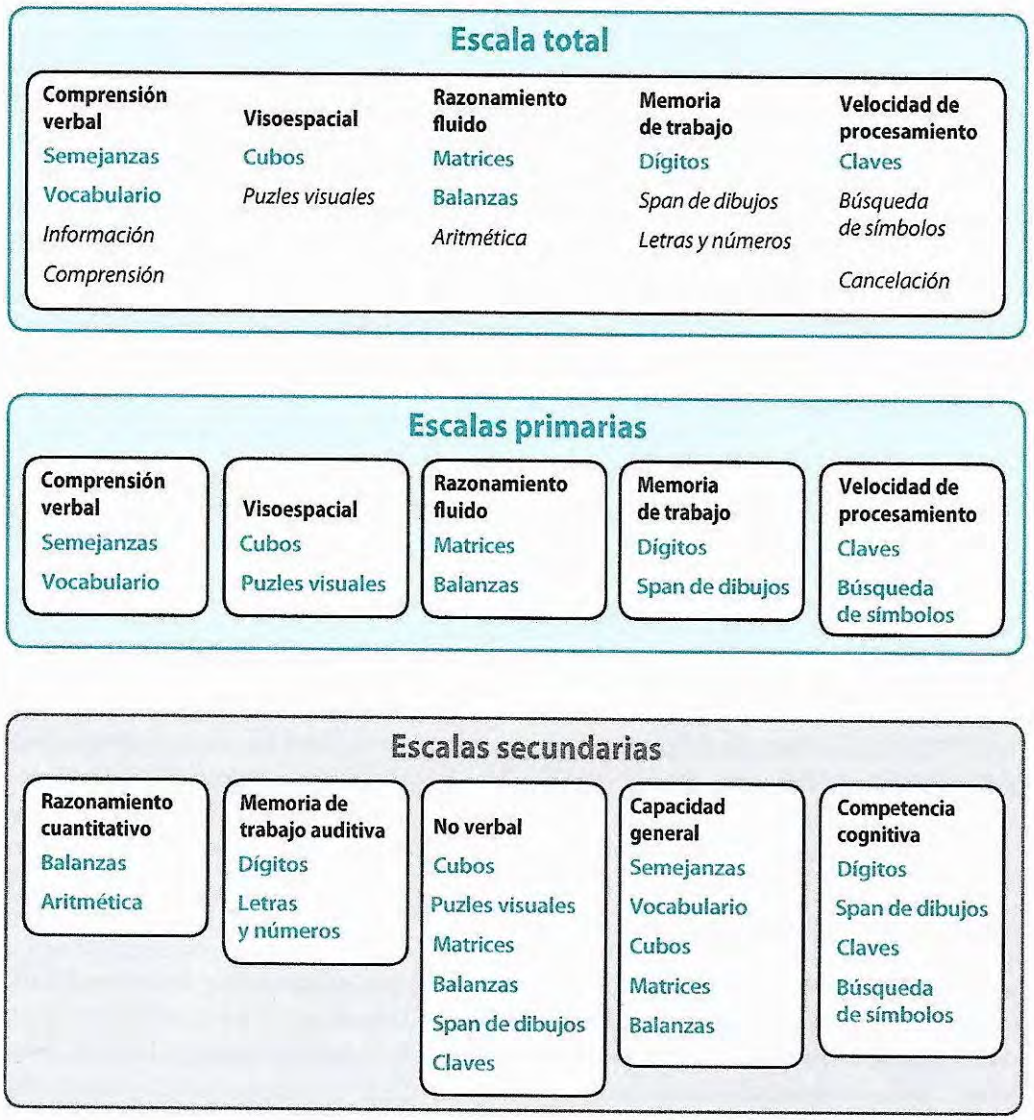


Figura 2.1. Estructura de la WISC-V

Objetivos de la revisión

Los objetivos de esta revisión se establecieron tras considerar distintos modelos estructurales de la inteligencia y de la capacidad cognitiva, la investigación sobre el neurodesarrollo y la neurocognición, los resultados psicométricos, la utilidad clínica y las necesidades prácticas de los profesionales. Asimismo se basan en más de setenta años de investigación con las escalas de inteligencia de Wechsler; en las aportaciones de expertos en los campos de la neuropsicología y el desarrollo cognitivo de los niños; en las sugerencias de los psicólogos que usan las escalas y, también, en el amplio estudio de la literatura existente en los campos de la teoría e investigación de la capacidad cognitiva e intelectual, la neurociencia cognitiva, la evaluación de la inteligencia y el desarrollo cognitivo de los niños. Los cinco objetivos principales fueron los siguientes:

- actualización de las bases teóricas;
- mayor adecuación al desarrollo;
- aumento de la facilidad de uso;
- mejora de las propiedades psicométricas;
- incremento de la utilidad clínica.

Actualización de las bases teóricas

La WISC-V ha recibido la influencia de varias teorías y modelos relevantes sobre la evaluación de la inteligencia. En este apartado se exponen los modelos estructurales de la inteligencia, las investigaciones sobre el neurodesarrollo y la neurocognición, los modelos sobre la memoria de trabajo y los estudios empíricos relacionados con el desarrollo de la WISC-V.

Modelos estructurales de la inteligencia

Las ideas de Wechsler sobre la inteligencia y las progresivas adaptaciones de sus escalas de inteligencia a las teorías e investigaciones sobre la cognición, la inteligencia y la neuropsicología parecen contrarias a la afirmación de que sus escalas carecen de base teórica (Coalson, Raiford, Saklofske y Weiss, 2010; Kaufman, 2010; Raiford y Coalson, 2014). No obstante, algunos autores afirman o insinúan que las escalas de inteligencia de Wechsler carecen de estructura o contexto teórico (Benson, Hulac y Kranzler, 2010; Kamphaus, 1993; Keith, Fine, Taub, Reynolds y Kranzler, 2006). Cuando Wechsler creó la *Escala de inteligencia de Wechsler-Bellevue*, destacaban dos teorías de la inteligencia: la de Spearman (1904), quien defendía la existencia de un factor general y otros específicos, y la de Thorndike (1936), que consideraba que hay varios tipos de inteligencia, como la abstracta, la social y la práctica. Aunque se consideró a Wechsler como un teórico del factor *g* (Kaufman, Flanagan, Alfonso y Mascolo, 2006), quizá por incluir una puntuación resumida para representar la inteligencia general, su definición de inteligencia (véase el capítulo 1 de este manual) no es incompatible ni con la teoría del factor *g* ni con la de los distintos tipos de inteligencia. Wechsler describió un aspecto general de la inteligencia, compuesto por capacidades distintas cualitativamente (Coalson et al., 2010), y, cuando desarrolló sus tests, seleccionó medidas de la capacidad cognitiva que luego resultaron ser importantes según los modelos estructurales de la inteligencia (Carroll, 1993, 2012; Johnson y Bouchard, 2005a, 2005b; Johnson, te Nijenhuis y Bouchard, 2007). Además, incorporó el uso de puntuaciones compuestas alternativas basadas en los estudios del análisis factorial (véanse Cohen, 1957a, 1957b, 1959 y Kaufman, 1975) y en las demostraciones de estudios clínicos que indicaban su utilidad en el diagnóstico diferencial (Coalson et al., 2010; Wechsler, 1958).

Actualmente, los modelos estructurales de la inteligencia ampliamente aceptados se basan en resultados exhaustivos del análisis factorial. Estos estudios demuestran con firmeza que hay una inteligencia general en el nivel superior de un modelo jerárquico, y varias capacidades generales relacionadas y distinguibles en el nivel inferior. En algunos modelos, las capacidades amplias están compuestas de varias capacidades más segmentadas en el nivel más inferior (Carroll, 1993, 2012; Horn y Blankson, 2012; Johnson y Bouchard, 2005a, 2005b; Johnson et al., 2004; Johnson et al., 2007; Salthouse, 2004). Aunque las evidencias aportadas por los modelos estructurales no

coinciden exactamente, la mayoría de los modelos indican que las capacidades de comprensión verbal, visoespacial, razonamiento fluido, memoria de trabajo y velocidad de procesamiento son componentes fundamentales (Carroll, 1993, 2012; Horn y Blankson, 2012; Johnson et al., 2007; Salthouse, 2004). Las escalas de inteligencia de Wechsler (p. ej., Wechsler, 2003, 2008, 2012) han evolucionado consecuentemente, incorporando nuevas medidas de la capacidad visoespacial, el razonamiento fluido, la memoria de trabajo y la velocidad de procesamiento. La WISC-V continúa esta línea evolutiva ofreciendo nuevas medidas de capacidad visoespacial, razonamiento fluido y memoria de trabajo, proporcionando un índice visoespacial y un índice de razonamiento fluido, y mejorando las medidas de comprensión verbal y velocidad de procesamiento, manteniendo las puntuaciones compuestas para ambas.

La teoría no es el único factor que guía el desarrollo de las escalas de inteligencia de Wechsler, pero el modelo de la WISC-V refleja las teorías estructurales, y en su interpretación pueden utilizarse las perspectivas teóricas justificables (p. ej., Flanagan y Kaufman, 2009). Queda fuera del alcance de este capítulo exponer con detalle las teorías de la inteligencia; si se desea más información sobre los modelos que pueden aplicarse al interpretar la WISC-V, véanse Flanagan y Harrison (2012) y Sattler (2008a).

Investigación sobre el neurodesarrollo y la neurocognición

La investigación neuropsicológica sobre el desarrollo del cerebro proporciona información acerca de las relaciones entre la cognición, el comportamiento y el funcionamiento cerebral, así como acerca del impacto que producen los trastornos neurológicos y del neurodesarrollo en estos dominios. Los estudios sobre las relaciones entre el rendimiento mostrado en un test cognitivo y la función y estructura cerebrales permiten comprender mejor el efecto que los fallos en determinadas vías neurales causan en las dificultades cognitivas de los niños con afecciones neurológicas, del neurodesarrollo y psiquiátricas. Varios de los constructos cognitivos medidos por la WISC-V pueden utilizarse para generar y probar hipótesis sobre déficits del procesamiento neuropsicológico, y pueden interpretarse desde la perspectiva del neurodesarrollo (Fiorello et al., 2006; Maricle y Avirett, 2012; McCloskey, Whitaker, Murphy y Rogers, 2012; Miller y Maricle, 2012). Aunque para la WISC-V no se ha dispuesto de estudios de neuroimagen, otros estudios relacionados vinculan tanto la inteligencia general como las capacidades cognitivas específicas medidas por esta escala con aspectos estructurales y funcionales del desarrollo cerebral. En esta sección se destacan los principales cambios del neurodesarrollo que se producen entre los 6 años y los 16 años 11 meses (rango de edad de la WISC-V), así como la relación de estas estructuras cerebrales y su activación y conectividad con la inteligencia general y las capacidades cognitivas específicas.

Neurodesarrollo e inteligencia general

El cerebro se desarrolla continuamente a lo largo de toda la vida: desde el nacimiento hasta la primera infancia, los cambios son muy rápidos y espectaculares; y durante la edad adulta, los cambios son más lentos y moderados. Los principales procesos que se producen en el cerebro a nivel celular durante la infancia y la edad adulta son la mielinización, la arborización (Lenroot y Giedd, 2006) y la poda sináptica (Low y Cheng, 2006). La mielinización es el proceso de formación de una vaina de materia blanca alrededor de los axones, que facilita la transmisión de los impulsos nerviosos a lo largo del axón y, con ello, mejora la velocidad y la sincronía de los patrones de actividad neuronal (Giedd, 2008). La arborización consiste en el incremento de las conexiones axónicas y dendríticas entre las células, lo que aumenta la interconectividad neuronal (Lenroot y Giedd, 2006). La poda sináptica o neuronal es la eliminación selectiva de las ramas y conexiones neuronales para garantizar la formación de circuitos funcionales (Low y Cheng, 2006). La combinación de la mielinización, arborización y poda permite al cerebro seguir aprendiendo durante toda la vida (Goyal y Raichle, 2013). Estos procesos también dan lugar a la aparición de la materia gris y de la materia blanca. La materia gris refleja el número de somas neuronales y su densidad, resultado de los efectos de proliferación y arborización; y la materia blanca refleja el número de axones mielinizados y su tamaño. La materia gris interviene en el procesamiento de la información, y la materia blanca interviene en el flujo eficiente de información (Colom, Karama, Jung y Haier, 2010). Estos procesos conllevan que la maduración cerebral sea más organizada, eficiente y rápida al procesar la información durante la infancia y la adolescencia.

Además de los cambios celulares, el tamaño y la estructura del cerebro cambia a lo largo de toda la vida; no obstante, los cambios estructurales que se producen entre los 6 años y los 16 años 11 meses (rango de edad de la WISC-V) son particularmente importantes para la maduración cerebral y el subsiguiente funcionamiento. Los cambios en las estructuras cerebrales tienen lugar a distintas edades y pueden variar entre los niños y las niñas; su pico de crecimiento se da durante la adolescencia y gradualmente va descendiendo a lo largo de la edad adulta (Giedd, 2008; Lenroot y Giedd, 2006; Schmithorst, Holland y Dardzinski, 2008; Wilde et al., 2009). El desarrollo cerebral de los niños de un mismo grupo de edad puede encontrarse en distintos estadios debido a factores genéticos, biológicos y ambientales. Las expresiones de un gen específico pueden regular la mielinización y la poda a lo largo de la vida (Goyal y Raichle, 2013). Los efectos genéticos parecen más intensos en el desarrollo del lóbulo frontal, en relación con el desarrollo de los lóbulos temporales, parietales y occipitales (Schmitt et al., 2008); y en cuanto a la integridad de la materia blanca, la expresión de la heredabilidad es mayor en los adolescentes que en los adultos, y en los sujetos con alta capacidad cognitiva que en los sujetos con baja capacidad cognitiva (Chiang et al., 2011). Aunque el tamaño y la estructura del cerebro son en gran medida heredables, los factores ambientales también afectan al desarrollo cerebral y pueden desempeñar un papel importante en la expresión de las dificultades cognitivas (Giedd, Schmitt y Neale, 2007). Los factores ambientales, como la exposición a neurotoxinas, la malnutrición, el estrés asociado a la vida urbana, el nivel de estimulación del entorno y las lesiones e infecciones cerebrales, afectan a los procesos cerebrales normales, en especial si se producen en periodos críticos del desarrollo del cerebro (Andersen, 2003; Georgieff, 2007; Lederbogen et al., 2011; Rice y Barone, 2000). Es importante considerar la influencia de la interacción de estos factores genéticos y ambientales al interpretar los resultados de tests cognitivos.

La investigación basada en neuroimágenes indica una relación generalmente positiva entre el desarrollo cerebral y el funcionamiento cognitivo general. Los distintos tamaños y volúmenes de la estructura de todo el cerebro o de una región se han relacionado positivamente con el CI total, la capacidad general y el funcionamiento cognitivo general (Frangou, Chitins y Williams, 2004; Lange, Froimowitz, Bigler, Lainhart y Brain Development Cooperative Group, 2010; Reiss, Abrams, Singer, Ross y Denckla, 1996). En los niños, la correlación entre el volumen total del cerebro y la capacidad cognitiva general es de alta a moderada, con una asociación significativa de la materia gris cortical y subcortical con la cognición, aunque la fuerza de esta asociación puede verse influida por el sexo (Haier, Jung, Yeo, Head y Alkire, 2005; Schmithorst, 2009; Willerman, Schultz, Rutledge y Bigler, 1991). La mayoría de los estudios comparativos de la cognición y la estructura cerebral se basan en datos transversales. Sin embargo, las investigaciones longitudinales sugieren que la relación entre el grosor de la corteza cerebral y la capacidad general cambia a lo largo del tiempo, con más correlaciones negativas en la primera infancia que se convierten en correlaciones positivas. La edad y el desarrollo de los lóbulos frontales son sumamente importantes en el desarrollo de la capacidad general (Shaw et al., 2006).

Además del tamaño y volumen del cerebro, la activación cerebral también se relaciona con el funcionamiento cognitivo; no obstante, los estudios sobre activación cerebral ofrecen resultados más heterogéneos: algunos indican una relación positiva entre la activación y el rendimiento en tareas con saturaciones en *g* (Lee et al., 2006), y otros señalan que una activación cerebral mayor no necesariamente se corresponde con un funcionamiento cognitivo mejor (Schmithorst y Holland, 2006). Al igual que muestran los resultados de los estudios sobre el desarrollo del cerebro y su estructura, la asociación entre la activación de las redes cerebrales y la capacidad general puede ser distinta en los dos sexos (Schmithorst y Holland, 2006).

Aunque las estructuras son importantes por sí mismas, la interconectividad puede tener un papel más relevante en el funcionamiento cognitivo general. Las redes existentes entre los lóbulos parietales y frontales son particularmente importantes, pero otras regiones cerebrales, incluidas la orbitofrontal, temporal y parahipocampal, contribuyen al rendimiento en los tests (Colom et al., 2009; Colom et al., 2010). Las estructuras subcorticales de los ganglios basales influyen en el procesamiento de la información entre las regiones frontal y parietal (Koziol y Budding, 2009). Dado que se cree que la capacidad general está relacionada con una red de regiones cerebrales que interactúan, las conexiones de la materia blanca que transfieren información de forma completa

y eficiente son un componente importante de la cognición (Deary, Penke y Johnson, 2010). La capacidad cognitiva general se relaciona con una amplia red de vías interconectadas y recursivas que procesan, refinan y manipulan información para producir conductas resolutivas de problemas complejos humanos.

Neurodesarrollo y capacidades cognitivas específicas

Comprender la relación entre el funcionamiento cerebral y las capacidades cognitivas específicas proporciona información sobre los factores del neurodesarrollo que pueden influir en la expresión cognitiva y conductual de los trastornos del desarrollo, neurológicos, psiquiátricos y médicos. Las puntuaciones de los índices primarios de la WISC-V miden capacidades cognitivas específicas: comprensión verbal, procesamiento visoespacial, razonamiento fluido, memoria de trabajo y velocidad de procesamiento. Aunque se utilizan tanto las estructuras globales del cerebro como las locales, cada uno de estos dominios cognitivos se relaciona con regiones cerebrales.

La comprensión verbal refleja la interacción de las destrezas lingüísticas básicas con la resolución y el razonamiento de problemas complejos. La WISC-V evalúa las destrezas lingüísticas en el contexto del razonamiento, la recuperación del conocimiento semántico, la expresión verbal y la abstracción. El desarrollo de la destreza lingüística varía considerablemente de un niño a otro. La fonología (sonidos), la semántica (significado de las palabras y vocabulario) y la sintaxis (gramática) interactúan dinámicamente e influyen en otros aspectos del desarrollo del lenguaje (Kuhl et al., 2008). Muchas destrezas lingüísticas básicas, como el procesamiento fonético y sintáctico, se desarrollan notablemente durante la primaria infancia y la edad preescolar, pero su desarrollo posterior es más limitado. En cambio, el desarrollo del vocabulario se acelera durante la primera infancia y la edad preescolar, pero muestra un crecimiento continuo a lo largo de toda la vida (Kuhl, 2010). Las estructuras cerebrales implicadas en el desarrollo del lenguaje varían en relación con las destrezas lingüísticas particulares en desarrollo, pero incluyen áreas de los lóbulos temporales y frontales (Dehaene-Lambertz et al., 2006). El desarrollo del lenguaje puede estar asociado con una importante especialización funcional del cerebro (es decir, la lateralización) en sujetos con destrezas lingüísticas normalmente desarrolladas, cuyo desarrollo del lenguaje está más estrechamente relacionado con las estructuras del hemisferio izquierdo (Dehaene-Lambertz, Hertz-Pannier, Dubois y Dehaene, 2008; Gauger, Lombardino y Leonard, 1997). Esta lateralización se observa en la primera infancia (Dubois et al., 2009) y se incrementa con la edad en los niños con un desarrollo normal (Knaus et al., 2009). Cabe destacar que la investigación indica que los niños de familias con un nivel socioeconómico más bajo muestran una menor lateralización izquierda, lo que puede resultar en un desarrollo del lenguaje menos eficiente (Raizada, Richards, Meltzoff y Kuhl, 2008).

Cuando los niños alcanzan la edad escolar, han desarrollado las destrezas lingüísticas básicas (p. ej., procesamiento semántico y sintáctico); y cuando aprenden a leer, empiezan a vincular dichas destrezas con el sistema visual a lo largo de vías neurales similares a las del lenguaje oral (Dehaene, 2010). El proceso del aprendizaje de la lectura cambia el funcionamiento de las regiones cerebrales implicadas en el lenguaje oral (Monzalvo y Dehaene-Lambertz, 2013). Aunque las vías relacionadas con la comprensión de la escucha se establecen hacia los 4 años, el incremento de la activación de los lóbulos frontales se asocia positivamente con la comprensión de historias en niños más mayores, lo que sugiere que la comprensión de material más complejo supone una mayor implicación de los lóbulos frontales (Berl et al., 2010). Los lóbulos frontales y las regiones prefrontales son especialmente importantes para el procesamiento semántico y las relaciones entre las palabras (p. ej., prueba *Semejanzas*), siendo las relaciones más abstractas las que requieren mayor implicación de los lóbulos frontales (Bookheimer, 2002; Bunge, Wendelken, Badre y Wagner, 2005; Goldberg, Perfetti, Fiez y Schneider, 2007; Green, Kraemer, Fugelsang, Gray y Dunbar, 2010; Kotz, Cappa, von Cramon y Friederici, 2002). Las regiones frontales son importantes para las capacidades cristalizadas (Colom, Burgaleta, et al., 2013); no obstante, en tareas de razonamiento verbal que requieren razonamiento semántico, también se activa el sistema temporal del hemisferio izquierdo (Goel, Buchel, Frith y Dolan, 2000). La especificidad del razonamiento verbal puede radicar en la región temporal izquierda, mientras que el razonamiento, más general, se asocia con las redes frontales.

El procesamiento visoespacial, a semejanza del procesamiento verbal, implica varias regiones del cerebro, en función de los requisitos de procesamiento de los estímulos visuales. Existe algún indicio de que el hemisferio derecho está más implicado en el procesamiento de las relaciones visoespaciales (Hugdahl, 2011). Los procesos visoespaciales se relacionan con las estructuras de los ganglios basales derechos y con el tálamo derecho, el córtex parietal derecho inferior y superior (Barbey et al., 2012; Burgaleta et al., 2013), el volumen parahipocampal derecho (Colom, Stein, et al., 2013) y los córtex fusiforme y temporal (Colom, Burgaleta, et al., 2013). Las tareas de razonamiento espacial se asocian con la activación de los lóbulos frontales bilaterales y el lóbulo parietal derecho (Prabhakaran et al., 1997). Mientras que el hemisferio derecho es fundamental para el procesamiento visoespacial, la importancia relativa de determinadas regiones cerebrales se ve influida por otras destrezas necesarias.

La capacidad de razonamiento fluido parece activar distintas regiones del cerebro. Estructuralmente, correlaciona con el grosor del giro frontal medial (Colom, Burgaleta, et al., 2013); el tamaño de las estructuras de los ganglios basales derechos; el núcleo caudado, el putamen y el núcleo accumbens (Burgaleta et al., 2013); y el volumen cerebelar (Posthuma et al., 2003). La activación cerebral regional durante el razonamiento fluido se observa en los lóbulos frontal, parietal, temporal y occipital del hemisferio izquierdo y en el córtex frontal derecho (Prabhakaran, Smith, Desmond, Glover y Gabrieli, 1997). La capacidad de discernir patrones o conexiones significativos, como los de las tareas de razonamiento de matrices, se asocia con una mayor activación del córtex prefrontal derecho (Christoff et al., 2001; Krawczyk, McClelland y Donovan, 2011; Wendelken, Nakhabenko, Donohue, Carter y Bunge, 2008). El razonamiento fluido se relaciona con el funcionamiento ejecutivo, el córtex prefrontal y frontal y las conexiones entre el córtex frontal y otras regiones (Langeslag et al., 2013). Las tareas de razonamiento fluido requieren la implicación de las redes frontales-parietales junto con varios procesos corticales y subcorticales. En ellas se observa una tendencia hacia una mayor activación en el hemisferio derecho que en el izquierdo; no obstante, esto puede deberse al hecho de que en muchas de las tareas de razonamiento fluido se utilizan estímulos visuales en vez de verbales.

La evaluación de los procesos de razonamiento aportan información sobre cómo responde el cerebro ante distintos tipos de tareas de resolución de problemas (p. ej., verbal, espacial, fluido). Las tareas de razonamiento, tanto inductivo como deductivo, activan el córtex prefrontal lateral izquierdo y los córtex frontal-dorsal, parietal y occipital bilaterales; sin embargo, la inducción y la deducción activan regiones diferentes, particularmente en el córtex frontal (Goel y Dolan, 2004; Knauff, Mulack, Kassubek, Salih y Greenlee, 2002; Parsons y Osherson, 2001). Las diferencias en la activación de una región u otra se relaciona más con el tipo de razonamiento requerido y con la familiaridad y significación contextual de los estímulos que con la modalidad (p. ej., visual frente a verbal) de los estímulos (Goel, 2007). Las investigaciones sugieren que las tareas de razonamiento, por lo general, activan redes cerebrales frontales y a menudo implican los lóbulos parietales, lo cual es consistente con las relaciones, descritas anteriormente, entre el funcionamiento intelectual general y las redes del cerebro.

La memoria de trabajo interviene a menudo para respaldar las capacidades de razonamiento. Las tareas de memoria de trabajo activan consistentemente las regiones de los lóbulos frontales, además de las regiones posteriores, incluidas las regiones corticales parietal, premotora y cingulada (Dumontheil y Klingberg, 2012; Owen, McMillan, Laird y Bullmore, 2005; Palva, Monto, Kulashekhar y Palva, 2010). Por otra parte, el desarrollo de las vías entre las regiones parietal y frontal durante la infancia y la adolescencia correlacionan con la mejora del funcionamiento de la memoria de trabajo (Klingberg, 2006). Las funciones de mantenimiento y manipulación de la memoria de trabajo, se asocian con distintas redes de las regiones prefrontal y parietal (Champod y Petrides, 2010). Desde el final de la infancia hasta la adolescencia, se activa el córtex prefrontal ventral al realizar el mantenimiento de la información en la memoria de trabajo. No obstante, cuando se requiere la manipulación mental de la información, en los adolescentes y los adultos más jóvenes se observa intervención del córtex prefrontal dorsolateral derecho; mientras que en los niños y los preadolescentes permanece la activación de las vías frontales ventrales, menos eficientes (Crone, Wendelken, Donohue, van Leijenhorst y Bunge, 2006; Scherf, Sweeney y Luna, 2006). La activación prefrontal dorsolateral, más que la activación ventral, refleja una

capacidad mayor para la manipulación cognitiva, aunque está por confirmar la contribución relativa de cada región. La investigación respalda la importancia de la interconectividad de las regiones frontoparietales, y el entrenamiento de la memoria de trabajo incrementa la activación de estas redes (Klingberg, 2010; Kundu, Sutterer, Emrich y Postle, 2013; Takeuchi et al., 2010).

La lateralización de la activación del córtex prefrontal se produce o bien bilateralmente o bien con preferencia de la derecha respecto a la izquierda en las tareas espaciales más que en las tareas con mayor mediación verbal; sin embargo, no está claro si las diferencias en la activación regional guardan relación con los requisitos de la tarea o con el tipo de estímulo (Braver et al., 2001; D'Esposito et al., 1998). Las regiones del cerebelo y las conexiones corticocerebelosas también incrementan la activación como respuesta a la mayor carga de memoria de trabajo (Kirschen, Chen, Schraedley-Desmond y Desmond, 2005; Marvel y Desmond, 2010). Los requisitos de la tarea, la carga de memoria de trabajo y la modalidad influyen en las regiones cerebrales que se activan durante la activación de la memoria de trabajo.

Otro componente de la memoria de trabajo son las funciones de control ejecutivo. El procesamiento eficiente de la información requiere ignorar la información irrelevante o superflua durante la resolución de un problema. Los bucles de los ganglios basales del córtex prefrontal filtran la información irrelevante para la tarea antes de que se inicien las operaciones de la memoria de trabajo (McNab y Klingberg, 2008). El hecho de que la activación y la ejecución sean coincidentes en las tareas de memoria de trabajo y en las de razonamiento se explica por el control de la interferencia, el cual se asocia a un mejor rendimiento en varios dominios cognitivos (Bunge y Wright, 2007; Burgess, Gray, Conway y Braver, 2011). La región cortical prefrontal es crítica para mantener la atención y ejercer el control cognitivo durante el procesamiento de la información; por ello, es una región primordial cuando se llevan a cabo tareas que requieren memoria de trabajo y razonamiento (Kane y Engle, 2002).

La capacidad de procesar información rápida y eficientemente se asocia con el volumen total del cerebro, del neocórtex y de la materia gris (Betjemann et al., 2010; Posthuma et al., 2003); con la integridad de la materia blanca en el cuerpo calloso y en los lóbulos parietales, temporales y frontales (Aukema et al., 2009; Turken et al., 2008); y con el tamaño del tálamo y del cuerpo calloso (Till et al., 2011). Los sujetos con una nivel alto de inteligencia reaccionan más rápidamente a los estímulos visuales y auditivos, y necesitan menos recursos cerebrales para procesar eficientemente, que los sujetos con una inteligencia media o baja (Deary et al., 2010). Ejercitar la velocidad de procesamiento produce un efecto en la estructura de la materia gris localizada en el lóbulo temporal y en sus conexiones regionales (Takeuchi et al., 2011). La relación entre la integridad de la materia blanca y la velocidad de procesamiento se ha estudiado principalmente en los adultos mayores, y se ha observado una asociación entre los cambios en la materia blanca, bien documentados, y el deterioro de la velocidad de procesamiento (Gunning-Dixon, Brickman, Cheng y Alexopoulos, 2009). Sin embargo, la relación entre el desarrollo de la materia blanca y su integridad no se ha estudiado tan bien en los niños y los adolescentes.

Los sistemas neurales que intervienen en la ejecución de funciones cognitivas específicas y en el funcionamiento cognitivo general se solapan considerablemente. En una amplia muestra de niños y adolescentes de entre 6 y 18 años, Karama y sus colaboradores (2011) no encontraron correlación entre el grosor de la corteza cerebral y medidas específicas de razonamiento espacial/fluido o razonamiento verbal cuando se controlaban los efectos de la capacidad general. En la memoria de trabajo y en el funcionamiento intelectual general intervienen redes similares (Colom et al., 2010), y estas destrezas tienden a mostrar una alta correlación (Colom, Rebollo, Palacios, Juan-Espinosa y Kyllonen, 2004). Dado el alto grado de correlación entre la capacidad general y la memoria de trabajo, es importante establecer asociaciones distintas con la estructura cerebral. Colom, Jung y Haier (2007) observaron que, incluso tras controlar la varianza coincidente entre las medidas, la capacidad general y la memoria de trabajo correlacionan con el volumen de materia gris en las regiones parietal y frontal, lo que indica que las redes parietales y prefrontales contribuyen, tanto en común como de forma independiente, en estas capacidades. En futuras investigaciones se seguirá estudiando sobre la existencia de una red específica para las capacidades cognitivas frente a una red más generalizada que sirve para todas las tareas cognitivas.

Modelos e investigación sobre la memoria de trabajo

Por lo general, la memoria de trabajo se define como la capacidad de mantener información en la mente mientras se usa para, simultáneamente, llevar a cabo y completar algunos tipos de tareas cognitivas (Cowan y Alloway, 2009). La memoria de trabajo es un aspecto clave del funcionamiento intelectual, con fuertes vínculos con el razonamiento fluido y otros procesos cognitivos de nivel superior (Burgess y Braver, 2010; Chuderski, 2013; Dang, Braeken, Colom, Ferrer y Liu, 2014; Giofrè, Mammarella y Cornoldi, 2013; Hornung, Brunner, Reuter y Martin, 2011); también interviene en una amplia variedad de problemas académicos y trastornos clínicos que afectan a los niños en edad escolar (p. ej., Borella, Carretti y Pelegrina, 2010; Fitzpatrick y Pagani, 2012; Hutchinson, Bavin, Efron y Sciberras, 2012). Consecuentemente, uno de los objetivos desde el inicio del desarrollo de la WISC-V fue mejorar las medidas de memoria de trabajo.

Es deseable que en algunos modelos de memoria de trabajo haya tareas con demandas cognitivas simultáneas y competitivas. Los niños en edad escolar generalmente pueden comprender las instrucciones de las tareas duales y de los paradigmas de secuenciación que crean demandas concurrentes en las medidas de memoria de trabajo. La tareas de memoria de trabajo de la WISC-V, además, se basan en la interferencia proactiva, que introduce una demanda cognitiva concurrente. La interferencia proactiva tiene lugar cuando un ítem visto previamente interfiere con la memoria del ítem presente. En la WISC-V se utiliza el mismo tipo de estímulo (dibujos, números, letras) sistemáticamente en los distintos ítems de cada prueba. Diversos estudios han demostrado que la incorporación de la interferencia proactiva en un conjunto de ítems de memoria de trabajo incrementa la dificultad y la carga de memoria de trabajo de los ítems posteriores (Lipinski, Simmering, Johnson y Spencer, 2010; Makovski y Jiang, 2008; Szmalec, Verbruggen, Vandierendonck y Kemps, 2011).

Aunque queda fuera del alcance de este apartado presentar un análisis completo de los modelos de memoria de trabajo actuales, seguidamente se ofrece una comparación de dos modelos ampliamente aceptados dado que son relevantes para las pruebas de memoria de trabajo de la WISC-V. Por una parte, el modelo multicomponente establece dos sistemas de almacenamiento específicos según el dominio, que almacenan, repiten o manipulan temporalmente la información: el bucle fonológico para la información verbal y la agenda visoespacial para la información visual y espacial (Baddeley, 2000, 2002, 2012). Por otra parte, el modelo de procesos anidados (Cowan, 1988, 1999; Towse y Cowan, 2005) propone que la información de la memoria de trabajo se desarrolla cuando se activa un subconjunto de memorias a largo plazo, y un subconjunto de estas memorias activadas se convierte en el foco de atención cuando es activado por los estímulos del entorno. Estos dos modelos (multicomponente y de procesos anidados) se diferencian en varios aspectos. Por ejemplo, el modelo multicomponente propone distinciones en función del dominio (p. ej., verbal y visoespacial), mientras que el modelo de procesos anidados no las establece. A pesar de estas diferencias, los defensores de ambos enfoques reconocen que las tareas de memoria de trabajo visual pueden implicar algunos procesos de memoria mediados verbalmente; este hecho demuestra la excesiva simplificación de concebir las memorias como exclusivas de un dominio u otro (Towse y Cowan, 2005).

No obstante, los estudios llevados a cabo con niños respaldan el modelo multicomponente de Baddeley. Parece que los componentes de almacenamiento específicos según el dominio son distintos en los niños de entre 6 años y 16 años 11 meses (rango de edad de la WISC-V) (Gathercole, Pickering, Ambridge y Wearing, 2004; Michalczyk, Malstädt, Worgt, Könen y Hasselhorn, 2013) y que muestran una sensibilidad diferencial a varias afecciones clínicas (Meyer, Salimpoor, Wu, Geary y Menon, 2010; Schuchardt, Bockmann, Bornemann y Maehler, 2013). Por todo ello, en la WISC-V se han incluido pruebas de memoria de trabajo auditiva y de memoria de trabajo visual para que puedan proporcionar información algo distinta y potencialmente una mayor amplitud del constructo.

Mejora de la adecuación al desarrollo

Al evaluar a niños en edad escolar, hay que tener en cuenta muchos aspectos fundamentales del desarrollo. Por ello, en esta nueva versión de la WISC se le concedió una importancia primordial a la mejora de la adecuación al desarrollo; en particular, respecto a las consignas y el redactado de los ítems, los criterios de puntuación y las bonificaciones por tiempo.

Consignas y redactado de los ítems

Un aspecto importante en la evaluación de los niños es su comprensión de las consignas de las tareas. Todas las consignas de las pruebas se simplificaron en cada una de las fases de desarrollo iniciales de la WISC-V. Las consignas se probaron repetidamente con niños de 4:6 a 5:11 años (más pequeños que los niños del rango menor de edad de la WISC-V), así como en niños con alguna discapacidad intelectual leve. Además, se pidió a los niños que explicaran sus estrategias y sus respuestas, con el objetivo de comprobar que los ítems implicaban los procesos cognitivos previstos. Asimismo se pidió a los examinadores su opinión sobre la claridad de las consignas.

Por otra parte, se incrementó el número de ítems de demostración, de ejemplo y de aprendizaje que se utilizan en las pruebas para explicar las consignas. Los ítems de demostración los resuelve el examinador, y le permiten explicar cómo se hace la tarea de la prueba. Los ítems de ejemplo los resuelve el niño, y le sirven para entrenarse en la tarea antes de empezar los ítems del test. A menudo también le permiten al examinador, en función del rendimiento del niño, volver a explicar en qué consiste la tarea. Para las pruebas que no cuentan con ítems de demostración ni de ejemplo (p. ej., *Vocabulario* e *Información*), si el niño no ha obtenido puntuaciones perfectas en los ítems de comienzo, que sirven como ítems de aprendizaje, el examinador dispone de estos ítems para explicar al niño cuál es la respuesta correcta.

Dado que muchos de los niños evaluados con la WISC-V tienen problemas de atención y comprensión, las consignas de esta escala son breves y concisas y, cuando es posible, se complementan con ítems de demostración y de ejemplo. En los inicios de la fase piloto se analizó la brevedad de las consignas y se realizaron varias revisiones de las mismas. Si una consigna hacía referencia a una conducta observada en pocas ocasiones o a una pregunta raramente formulada, la consigna se transformaba en una pregunta adicional. Las consignas de todas las pruebas se leyeron en voz alta y se cronometraron. Además, se contó el número de palabras de cada consigna. Para cada una de las pruebas conservadas de la WISC-IV y en las que se habían modificado las consignas, se compararon el tiempo de lectura y el número de palabras de las consignas de esa versión con los resultados correspondientes de la WISC-V para garantizar que las consignas nuevas requerían menos tiempo, estaban formadas por menos palabras y seguían siendo claras para los niños.

Por otra parte, se revisaron las pruebas cuyos ítems consisten solo en preguntas (p. ej., *Información*, *Comprensión* y *Aritmética*), teniendo en cuenta el nivel de vocabulario receptivo de los niños, y se eliminaron o simplificaron las palabras complejas siempre que fue posible.

Criterios de puntuación

Para las pruebas que requieren respuestas más elaboradas, se analizaron las respuestas que con más frecuencia daban los niños de la muestra experimental y de tipificación con el objetivo de establecer las puntuaciones apropiadas para los distintos tipos de respuestas. Asimismo, se adaptaron los criterios de puntuación teniendo en cuenta el vocabulario limitado observado en algunos niños de 6 y 7 años; y se primó el significado de la respuesta por encima de la precisión literal del contenido. Para facilitar la puntuación, se añadieron más aclaraciones a los ejemplos de respuesta que figuran en el *Manual de aplicación y corrección*.

Bonificación por tiempo

Conceder demasiada importancia a la rapidez con que se realiza una tarea puede ser especialmente problemático para los niños más pequeños, los niños que no se muestran muy predispuestos y los niños con dificultades motoras no relacionadas con problemas cognitivos. Además, todas las culturas no conceden la misma importancia a la velocidad en la ejecución de una tarea (Armour-Thomas y Gopaul-McNichol, 1997), El tiempo límite es adecuado y necesario para algunas pruebas porque permite evitar que el tiempo de aplicación sea excesivo, minimizar la frustración en el sujeto y garantizar un rango de dificultad de los ítems apropiado. Sin embargo, limitar el número de ítems con bonificación por tiempo reduce la posible confusión acerca del cronometraje de pruebas que no se han creado específicamente para medir la velocidad de procesamiento. Por ello se redujo el número de ítems de *Cubos* con bonificación por tiempo.

Aumento de la facilidad de uso

Los aspectos relacionados con la facilidad de aprendizaje, uso e interpretación de la WISC-V son diversos. Por ello, al desarrollar las instrucciones, ítems y materiales de esta nueva versión, se tuvieron en cuenta las opiniones de los examinadores expertos. En esta sección se describen las consideraciones realizadas sobre estos aspectos a lo largo del desarrollo de la WISC-V.

Seguridad de los ítems

La seguridad de los ítems es un aspecto importante para los usuarios de tests de inteligencia porque los ítems que han pasado a ser públicos se practican y se memorizan con más facilidad. Los resultados de recientes disputas sobre la difusión de los materiales de tests con derechos de autor indican que el contenido de un test se considera del dominio público cuando han transcurrido varios años desde su publicación (Demas y Brogdon, 1997). Para garantizar que el contenido de los ítems está protegido por los derechos de autor, se cambiaron todos los ítems que figuraban en la WISC-IV y que se habían publicado originalmente en la WISC, la WISC-R o la WISC-III.

Materiales

El *Manual de aplicación y corrección* contiene toda la información necesaria para aplicar y puntuar todas las pruebas y para cumplimentar las páginas de resumen, de análisis primario, de análisis secundario y de análisis de procesamiento del cuadernillo de anotación. Se distribuyeron más equitativamente los contenidos de los cuadernos de estímulos para reducir su peso y, así, facilitar su transporte y evitar que la encuadernación se estropease. Los dos cuadernos de estímulos contienen los estímulos de las pruebas principales y opcionales necesarias para obtener las puntuaciones de los índices primarios, el CI total y las puntuaciones de los índices secundarios. Asimismo, para mejorar su durabilidad y facilidad de uso, la plantilla de corrección para *Claves* es ahora de un material más rígido y duradero.

Tiempo de aplicación

Dado que los niños tienen dificultad para mantener la atención durante periodos largos, es esencial que la aplicación del test sea eficaz. Además, los profesionales piden que el tiempo de aplicación sea lo más breve posible, sin dejar de obtener la información necesaria. La reducción del tiempo de aplicación de la WISC-V permite disponer de tiempo para evaluar otros dominios importantes (p. ej., rendimiento académico, lenguaje, función ejecutiva). A lo largo del desarrollo de la WISC-V se realizó un esfuerzo considerable para lograr que el tiempo de aplicación fuera lo más breve posible y que se ofreciera una amplitud del constructo y una flexibilidad mayores, así como más puntuaciones compuestas.

Además de abreviar las consignas de las pruebas, el número de ítems aplicado se redujo al mínimo porque se disminuyó el número total de ítems y se modificaron las reglas de terminación (véase más adelante el apartado «Reglas de terminación», en este mismo capítulo). El número total de ítems de las pruebas procedentes de la WISC-IV se redujo.

El tiempo de aplicación es un aspecto fundamental al seleccionar las pruebas principales y del CI total. Un exhaustivo estudio analizó el efecto de potenciales reducciones en el número de pruebas que conforman las puntuaciones de los índices primarios y del CI total. Para tomar las decisiones finales sobre las pruebas que conforman las puntuaciones compuestas, se utilizaron diversas fuentes; entre ellas, las opiniones de destacados expertos en inteligencia, internos y externos; la evaluación de la utilidad predictiva y clínica de varias combinaciones de pruebas para obtener las puntuaciones compuestas; y las opiniones de profesionales sobre la amplitud del constructo y el tiempo de aplicación considerados adecuados.

Instrucciones de aplicación y corrección

Las instrucciones de aplicación y corrección de la WISC-V son más explícitas y sencillas que las de la WISC-IV. Por ejemplo, las reglas de invalidez, prorrateo y sustitución de las puntuaciones compuestas se simplificaron considerablemente y son más fáciles de aplicar (véase el capítulo 2 del *Manual de aplicación y corrección* para más información sobre dichas reglas). Otro ejemplo de simplificación son las reglas de repetición. En la WISC-V, la repetición de ítems o de consignas, se indica claramente cuando está permitida.

Reglas de terminación

Para todas las pruebas procedentes de la WISC-IV se redujeron significativamente las reglas de terminación. Así, por ejemplo, la regla de terminación para *Semejanzas* era, en la WISC-IV, de 5 respuestas de 0 puntos consecutivas, y en la WISC-V pasó a ser de 3 respuestas de 0 puntos consecutivas. Las reglas de terminación para las pruebas nuevas se establecieron para minimizar el número de respuestas de 0 puntos consecutivas y, al mismo tiempo, garantizar que se aplica un número adecuado de ítems para obtener una estimación precisa del rendimiento. Asimismo se estableció la misma regla de terminación (3 respuestas de 0 puntos consecutivas) para la mayoría de las pruebas principales, de este modo es más fácil recordar la regla durante la aplicación. En el capítulo 4 de este manual se ofrece información detallada sobre los estudios empíricos llevados a cabo para determinar las reglas de terminación.

Mejora de las propiedades psicométricas

Baremos y método de tipificación

Varios estudios indican que los baremos más antiguos producen puntuaciones infladas en las medidas de inteligencia (Flynn, 1999, 2007; Flynn y Weiss, 2007). Las puntuaciones de los tests deben basarse en información normativa actual y representativa de la población pertinente. En el capítulo 4 de este manual se describen con detalle las características de la muestra española. Además de los baremos actualizados, la WISC-V utiliza un método de tipificación actualizado. Este método inferencial se describe en el capítulo 4 de este manual y se utilizó para el desarrollo de los baremos de las pruebas.

Fiabilidad y validez

Se realizaron diversos estudios para demostrar la fiabilidad y validez de la escala. Los datos de retest de la muestra española se obtuvieron para tres franjas de edad (6-7, 10-11 y 14-15). Mediante estudios de correlación se obtuvieron resultados de validez convergente y discriminante con los siguientes instrumentos: WISC-IV (datos españoles), WPPSI-IV, WAIS-IV, KABC-II, *Vineland Adaptive Behavior Scales-Second Edition* (Vineland-II; Sparrow, Cicchetti y Balla, 2005), y *Behavior Assessment System for Children-Second Edition* (BASC-2 CPP; Reynolds y Kamphaus, 2004) en su forma para los padres.

Durante la tipificación estadounidense y española de la WISC-V también se recogieron datos de grupos especiales.

La validez de constructo se obtuvo de diversos estudios de análisis factorial confirmatorios y de la comparación de las puntuaciones medias obtenidas por los grupos especiales y los grupos control. En los capítulos 5 y 6, respectivamente, se describen y se presentan los resultados de los estudios de fiabilidad y validez.

Suelo y techo de las pruebas

Se prestó una atención particular al efecto suelo y techo de cada prueba para garantizar que el rango de puntuaciones fuera representativo de los distintos niveles de capacidad cognitiva, desde los más bajos (40 ≤ CIT ≤ 69) hasta los más altos (130 ≤ CIT ≤ 160). A fin de garantizar que las pruebas procedentes de versiones anteriores seguían midiendo adecuadamente todos los niveles de capacidad, se les añadieron ítems de poca y mucha dificultad si así lo requerían. Dado que los suelos de la WISC-IV eran excelentes, prácticamente no fue necesario realizar modificaciones en el nivel de dificultad baja. No obstante, se mejoró el suelo de *Dígitos* para la franja de menor edad (6:0-6:3 años). Para ampliar los techos de las pruebas, se añadieron ítems más difíciles a *Semejanzas, Vocabulario, Información, Comprensión, Cubos, Dígitos* y *Letras y números.*

Sesgo de los ítems

En la selección de los ítems de la WISC-V se usaron distintas técnicas destinadas a evitar los sesgos de los ítems. En las fases iniciales de desarrollo de esta nueva versión, un equipo de revisores internos y externos analizaron los posibles sesgos, la obsolescencia cultural, la relevancia del contenido y la utilidad clínica de todas las pruebas y los ítems que las componen. Durante la fase piloto, la de experimentación y posteriormente la de tipificación, varios expertos revisaron los ítems para detectar aquellos que podían resultar problemáticos por el sesgo y por su contenido. Por ello se han eliminado los ítems problemáticos según la opinión de los expertos y los datos empíricos de los análisis estadísticos de funcionamiento diferencial de los ítems (FDI). En todas las fases de desarrollo de la escala, expertos en evaluación de la inteligencia y en investigaciones transculturales revisaron los ítems. Además, diversos revisores de países en los que se dispone de versiones de las escalas Wechsler adaptadas y traducidas aportaron datos sobre su adecuación internacional.

Además de todas estas revisiones, se usaron los datos empíricos de los análisis FDI para servir de base en la toma de decisiones. Los resultados de los análisis FDI de Mantel-Haenszel (Holland y Thayer, 1988) y de los análisis FDI de teoría de respuesta al ítem (TRI; Hambleton, 1993) señalaron ítems potencialmente problemáticos, y estos fueron revisados de nuevo.

Opciones del nivel de significación de los valores críticos

Las opciones del nivel de significación de los valores críticos, usadas para comparar puntuaciones, se ampliaron respecto a las de la WISC-V. El número de opciones pasó de dos a cuatro: la WISC-V ofrece los niveles .01 y .10, además de los niveles .05 y .15. El nivel .01 es una opción aún más estricta que el nivel .05. Este cambio concede a los profesionales más flexibilidad en su intento de lograr un equilibrio entre la potencia estadística y el posible error de tipo I al comparar las diferencias. El nivel de significación puede ajustarse en función del número de comparaciones que hacer, y si tales comparaciones se han establecido a priori o se han realizado post hoc. En el capítulo 2 del *Manual de aplicación y corrección* y en el capítulo 7 de este manual se ofrecen más detalles sobre la adecuada selección de los niveles de significación de los valores críticos.

Incremento de la utilidad clínica

Todos los objetivos de la revisión previamente comentados guardan relación con la utilidad clínica de la escala. Por ejemplo, la mayor adecuación al desarrollo incrementa la validez de los resultados obtenidos. En los apartados siguientes se describen las características de la WISC-V que mejoran la utilidad clínica.

Ilustraciones e ítems

Es fundamental que las ilustraciones y los ítems sean actuales y sugerentes porque, de este modo, los niños se centran en las tareas en vez de distraerse con dibujos, situaciones o frases que les resultan anticuadas; y ello incrementa la utilidad clínica del test. Todas las ilustraciones de la WISC-V se han modificado para que sean actuales y más atractivas y motivadoras para los niños. Asimismo se revisaron los ítems y se sustituyeron los anticuados por ítems nuevos que reflejasen mejor las ideas y situaciones actuales.

Estructura de la escala

La estructura de la WISC-V refleja los cambios respecto a la versión anterior y fue influenciada por los modelos estructurales de la inteligencia, la teoría del neurodesarrollo, la investigación sobre la neurocognición, y la utilidad clínica. Los estudios recientemente publicados de análisis factorial de la WISC-IV y la WAIS-IV han examinado los modelos de cinco factores para ambas escalas (Benson et al., 2010; Keith et al., 2006; Weiss, Keith, Zhu y Chen, 2013a, 2013b) de forma similar a la estructura de la WISC-V. En el capítulo 6 de este manual se presentan estos estudios, así como los resultados de los estudios de análisis factorial confirmatorio de la WISC-V.

Las decisiones adoptadas sobre la estructura de la escala se fundamentan en los modelos del neurodesarrollo y de la neurocognición y en la utilidad clínica. Cabe destacar que el hecho de disponer de puntuaciones independientes para el Índice visoespacial y el Índice de razonamiento fluido incrementa la precisión de la interpretación. La incorporación de la memoria de trabajo visual mejora la utilidad clínica de la escala dada la sensibilidad diferencial de las tareas de memoria de trabajo auditiva y de memoria de trabajo visual a numerosas afecciones clínicas (Gathercole et al., 2004; Meyer et al., 2010; Michalczyk et al., 2013; Schuchardt et al., 2013).

Los dominios cognitivos representados por la WISC-V mediante las puntuaciones de los índices primarios han demostrado su utilidad en evaluaciones psicológicas de distinto tipo. El rendimiento en las pruebas de Comprensión verbal es bajo en los niños con trastornos de la lectura, trastornos con dificultad matemática, trastornos del lenguaje y trastornos del espectro autista con deterioro del lenguaje, según diversos estudios (Chilosi et al., 2009; Compton, Fuchs, Fuchs, Lambert y Hamlett, 2012; Osmon, Smerz, Braun y Plambeck, 2006). Por lo general, los niños con trastorno del espectro autista sin deterioro del lenguaje no muestran puntos débiles en ninguna área, excepto en la prueba *Comprensión*, según algunos estudios (Mayes y Calhoun, 2008; Wechsler, 2003, 2012). Los déficits en la capacidad visoespacial se asocian con un trastorno específico del aprendizaje con dificultad matemática, pero algunos estudios señalan que esta capacidad puede ser un punto fuerte relativo para los niños con trastornos del espectro autista (Mayes y Calhoun, 2008; Osmon et al., 2006; Wechsler, 2003, 2008, 2012). Los déficits del razonamiento fluido se relacionan con el trastorno de la comprensión lectora y el trastorno con dificultad matemática (Fuchs et al., 2012; Osmon et al., 2006), pero no con el TDAH ni el trastorno del espectro autista de alto funcionamiento (Oliveras-Rentas, Kenworthy, Roberson, Mattin y Wallace, 2012).

La memoria de trabajo y la velocidad de procesamiento guardan más relación con diversos trastornos clínicos. Por ejemplo, la memoria de trabajo es sensible a determinados trastornos del aprendizaje, al TDAH, a trastornos del lenguaje y a trastornos del espectro autista (Archibald y Gathercole, 2006; Borella et al., 2010; Fitzpatrick y Pagani, 2012; Hutchinson et al., 2012; Passolunghi, 2006; Pickering, 2006). La velocidad de procesamiento es sensible a los trastornos del aprendizaje, el TDAH y los trastornos del espectro autista, entre otros (Compton et al., 2012; Mayes y Calhoun, 2007, 2008). Cabe destacar que los niños con alta capacidad intelectual reflejan sus puntos fuertes en comprensión verbal, capacidad visoespacial y razonamiento fluido; sin embargo, en promedio, muestran un rendimiento más bajo en memoria de trabajo y velocidad de procesamiento (Raiford, Weiss, Rolfhus y Coalson, 2005; Rimm, Gilman y Silverman, 2008; Wechsler, 2003, 2012).

Aunque originalmente los tests de inteligencia no fueron diseñados para su uso como medidas neuropsicológicas, las puntuaciones de los índices primarios de la WISC-V representan procesos cognitivos importantes en la evaluación neuropsicológica. Los resultados de la WISC-V pueden

utilizarse para generar hipótesis sobre los déficits del procesamiento neuropsicológico, y pueden conceptualizarse e interpretarse desde el punto de vista neuropsicológico (Hale y Fiorello, 2004; Miller y Maricle, 2012). Por ejemplo, el ICV es, en parte, una medida de la expresión oral y del lenguaje receptivo. El IVE y el IVP se corresponden con los procesos cognitivos generalmente evaluados en los dominios visoespacial, y de velocidad y de eficiencia, respectivamente. El IRF puede conceptualizarse como la medición de la resolución de problemas, la flexibilidad y el razonamiento, constructos asociados a la función ejecutiva. El IMT se diseñó para medir la memoria de trabajo, un proceso generalmente evaluado como parte del aprendizaje y la memoria (Miller y Maricle, 2012). Como ya se ha mencionado, estas capacidades específicas parecen estar asociadas con la activación o el desarrollo de diferentes regiones cerebrales.

Metodología de comparación de las diferencias entre puntuaciones

El enfoque de la WISC-V hacia la comparación de las diferencias entre puntuaciones es un cambio significativo respecto al de la WISC-IV. Esta nueva metodología de la que seguidamente se ofrece un resumen, se expone con detalle en el capítulo 7 de este manual. Los puntos fuertes y débiles en los distintos índices primarios pueden evaluarse utilizando una estimación de la aptitud general (la puntuación media de los índices primarios o CI total) como puntuación de comparación. Las comparaciones con estimaciones generales del rendimiento orientan los posteriores análisis del rendimiento en dominios más específicos (comparaciones entre dos puntuaciones de los índices primarios o dos pruebas). Este enfoque evaluativo puede reducir el número de comparaciones necesarias porque va de las comparaciones generales a las específicas, de lo que resulta una menor probabilidad de errores de tipo I, es decir, falsos positivos (Grégoire, Coalson y Zhu, 2011; Longman, 2004; Naglieri y Paolitto, 2005). Las otras opciones de nivel de significación de los valores críticos también le ofrecen al profesional más control sobre los errores de tipo I, en función del motivo de consulta específico del niño.

Requisitos de lenguaje expresivo

El INV, un nuevo índice secundario, ofrece información sobre el rendimiento de los niños con problemas del lenguaje expresivo; y se obtiene a partir de pruebas que no requieren ninguna respuesta expresiva. Este índice incluye pruebas de cuatro o cinco dominios cognitivos que contribuyen al CI total (Comprensión verbal no está incluido). El INV proporciona una estimación potencialmente más adecuada de la capacidad general de los niños con retrasos del lenguaje expresivo significativos o con otros trastornos clínicos asociados a dificultades verbales expresivas. Además, ofrece una estimación útil de la capacidad general de los niños sordos o con dificultades de audición, así como de los niños que están aprendiendo español. Este índice abarca más de un dominio cognitivo y, por lo tanto, proporciona una estimación más sólida de la capacidad general que los índices primarios, como el IVE o el IRF.

Puntuaciones de los índices secundarios

Además del INV, la WISC-V ofrece otros índices secundarios: el IRC, el IMTA, el ICG y el ICC, que se describen en el capítulo 7 de este manual. Los índices secundarios proporcionan información adicional sobre el rendimiento de un niño en la WISC-V.

Puntuaciones de procesamiento

La WISC-V ofrece más puntuaciones de procesamiento que la WISC-IV. Estas puntuaciones mejoran la profundidad de la interpretación y comprensión del rendimiento, y facilitan la comprensión de las puntuaciones desde la perspectiva del enfoque de procesamiento. Las puntuaciones de procesamiento de la WISC-V se describen en el apartado «Puntuaciones de procesamiento» de este capítulo y en el capítulo 7 de este manual.

Grupos especiales

Para aumentar la utilidad clínica de la escala, durante el proceso de tipificación español y estadounidense se realizaron estudios con diversos grupos especiales. Estos estudios, que se describen en el capítulo 6 de este manual, proporcionan evidencias iniciales de la utilidad clínica de la WISC-V para los siguientes grupos especiales: alta capacidad intelectual, discapacidad intelectual, discapacidad intelectual leve, discapacidad intelectual moderada, capacidad intelectual límite, trastorno específico del aprendizaje con dificultades en la lectura (en adelante, «Trastorno específico del aprendizaje-lectura»), trastorno específico del aprendizaje con dificultades en la lectura y en la expresión escrita (en adelante, «Trastorno específico del aprendizaje-lectura y expresión escrita»), trastorno específico del aprendizaje con dificultad matemática (en adelante, «Trastorno específico del aprendizaje-matemáticas»), trastorno por déficit de atención/hiperactividad, comportamiento disruptivo, traumatismo craneoencefálico, trastornos del espectro autista con deterioro del lenguaje acompañante y trastornos del espectro autista sin deterioro del lenguaje acompañante.

Capítulo 3

Desarrollo de la versión estadounidense de la WISC-IV

El proceso iterativo de investigación previo a la publicación de la WISC-V tuvo una duración de cinco años, y cada una de sus fases conllevó mejoras en la escala. En este capítulo se describe todo el proceso, desde las directrices de investigación seguidas para desarrollar la escala, así como las fases de investigación, hasta la composición final de la escala.

Directrices de investigación

Los *Standards for Educational and Psychological Testing* (*Standards*; American Educational Research Association [AERA], American Psychological Association [APA] y National Council on Measurement in Education [NCME], 1999) han sido recursos básicos durante todo el proyecto de investigación de la WISC-V y han aportado criterios para «el valor de los tests, las buenas prácticas en la evaluación y las consecuencias del uso de los tests». En *Standards* figura abundante información sobre estos temas y otros aspectos relacionados con el desarrollo y la evaluación de los tests.

Principales fases de la investigación

Antes de establecer la estructura final de la escala, se siguieron varias etapas experimentales, basadas en un minucioso proyecto de trabajo. La WISC-V se desarrolló en cinco fases: desarrollo conceptual, versión piloto, experimentación, tipificación, y composición y evaluación final. En los siguientes apartados se presentan brevemente estas fases, prestando especial atención a los estudios previos y a la metodología seguida.

Desarrollo conceptual

Esta primera fase se centró en definir los objetivos de la revisión de la escala, así como la metodología que se seguiría para lograr dichos objetivos. Se consultaron numerosas fuentes bibliográficas y los estudios de las encuestas realizadas a los profesionales y usuarios, así como estudios de expertos y examinadores de Pearson a fin de establecer una primera serie de objetivos. Posteriormente se formó el equipo del proyecto, compuesto por el equipo investigador y de desarrollo, el director del producto, el director del proyecto, los consultores de Pearson y un equipo asesor. Asimismo participaron en el proyecto diversos miembros de Pearson especialistas en edición, producción, investigación de campo (recolección de datos) y psicometría. Y, finalmente, se contó con la inestimable colaboración de expertos, internos y externos, en psicometría, en las escalas de inteligencia de Wechsler, así como en otras áreas de investigación relacionadas y en el desarrollo de tests.

Equipo asesor

El equipo investigador contó, a lo largo del proyecto, con la colaboración de reconocidos expertos en psicología clínica infantil, neuropsicología y trastornos del aprendizaje (véase el anexo C, en el que figura el panel de expertos). Tras cada una de las principales fases (p. ej., piloto, experimentación o tipificación) se revisó el proyecto de trabajo junto con el equipo asesor, teniendo en cuenta los resultados psicométricos obtenidos y las opiniones del equipo del proyecto, expertos, revisores de sesgo y examinadores.

Investigación profesional

En los inicios del proceso de desarrollo de la WISC-V se llevó a cabo una encuesta en línea (N = 690) con profesionales expertos en evaluación infantil y con profesionales que utilizaban la WISC-IV. A partir de los resultados de esta encuesta, de las revisiones del test, de las opiniones de los expertos y de la consulta de numerosas fuentes bibliográficas se elaboró el plan de trabajo y se desarrolló la versión de la escala para la fase piloto.

Estudios semiestructurados de expertos y examinadores

En todas las etapas de recogida de datos se realizaron estudios semiestructurados en los que participaron expertos en evaluación y capacidades cognitivas. Asimismo se llevaron a cabo encuestas en línea que permitieron a los expertos y examinadores valorar las versiones de la escala y realizar comentarios en términos de adecuación al desarrollo, utilidad clínica y facilidad de uso. Los resultados de estos estudios se resumieron y analizaron con el equipo del proyecto. Además, a lo largo de las fases de desarrollo de la escala, se realizaron diversas modificaciones al plan de trabajo investigador y a las distintas versiones de la escala a partir de las opiniones de los expertos y los examinadores y los datos del análisis cualitativo y cuantitativo.

Estudios piloto

El principal objetivo de esta fase fue crear la versión de la escala que se usaría en la fase de experimentación. Se realizó una serie de tres estudios minipiloto (N = 17, 5 y 20) y tres estudios piloto (N = 431, 397 y 120). En cada uno de ellos se utilizó una versión de la escala formada por grupos de pruebas ya existentes en la WISC-IV y pruebas experimentales de las que se estaba valorando su inclusión en la escala. Esta fase se centró en el análisis del contenido y la relevancia de los ítems, la adecuación de los techos y suelos de las pruebas, la claridad de las consignas tanto para el examinador como para el niño, los procedimientos de aplicación, la identificación de los procesos de respuesta, los criterios de puntuación y el sesgo de los ítems.

Experimentación

En la fase de experimentación se utilizó una versión de la escala formada por 22 pruebas. Los datos se obtuvieron de una muestra estratificada de 356 niños representativa de la población de Estados Unidos. La muestra estaba dividida en los nueve grupos de edad siguientes: 6:0-6:11, 7:0-7:11, 8:0-8:11, 9:0-9:11, 10:0-10:11, 11:0-11:11, 12:0-13:11, 14:0-14:11 y 15:0-16:11. La muestra se diseñó de acuerdo a los datos de población descritos en la American Community Survey 2012 de la Oficina del Censo de los Estados Unidos (Ruggles et al., 2010 [los datos del censo son del 2012]). Los distintos grupos de edad se estratificaron según las variables siguientes: sexo, nivel educativo de los padres y zona geográfica. Dentro de cada grupo de edad, la muestra de experimentación se aproximó mucho a los datos del censo para las variables demográficas.

Esta muestra, amplia y representativa, permitió volver a examinar algunas de las cuestiones planteadas en la fase piloto y resolver cuestiones nuevas. El orden de los ítems se refinó teniendo en cuenta estimaciones más precisas de la dificultad relativa de los ítems, y se llevaron a acabo análisis factoriales exploratorios y confirmatorios para determinar la estructura factorial de la escala. Además, se recopilaron datos de dos grupos especiales (niños con un diagnóstico de alta capacidad

intelectual [*N* = 27] y niños con discapacidad intelectual [*N* = 29]) para analizar la adecuación de los suelos y techos de las pruebas y evaluar la utilidad clínica inicial de la escala.

Tipificación

Tomando en cuenta los resultados de los estudios piloto y de experimentación, se construyó la versión de tipificación de la WISC-V. Dado que los resultados obtenidos con la versión de experimentación mostraron que las pruebas eran estructural y psicométricamente sólidas, para la versión de tipificación se realizaron mínimas modificaciones (p. ej., ítems, orden de los ítems, estímulos).

La fase de tipificación se centró en la recopilación y obtención de baremos y la comprobación de la fiabilidad, validez y utilidad clínica de la escala. Los datos se obtuvieron de una muestra estratificada de 2200 niños de edades comprendidas entre los 6 años y los 16 años 11 meses, así como de muestras de niños pertenecientes a 13 grupos especiales. Para demostrar la validez de la escala, se aplicó la WISC-V y otras medidas (p. ej., WISC-IV) a otras muestras de niños. La información demográfica detallada de las muestra de tipificación se presenta más adelante, en este mismo capítulo. Las características de las muestras que participaron en los estudios de validez concurrente y de grupos especiales figuran en el capítulo 6 de este manual.

El equipo investigador analizó los resultados psicométricos de los estudios de tipificación junto con los informes realizados por expertos y examinadores. A partir de todos los datos recopilados a lo largo del proceso de investigación, se estableció la estructura final de la escala para componer y evaluar la versión definitiva.

Composición final y evaluación de la WISC-V

Consistencia de los ítems, las consignas y los estímulos

A fin de garantizar que se conservaban los mejores ítems tras la selección efectuada, las versiones experimentales de la WISC-V incluían más ítems de los que serían necesarios para la mayoría de las pruebas de la versión definitiva. A lo largo de las diversas fases de desarrollo de la escala, se evaluaron los ítems con el objeto de conservarlos, modificarlos o eliminarlos. Por lo tanto, todos los ítems habían sido evaluados en diversos momentos antes de la tipificación de la escala. La selección final de los ítems se realizó tras analizar los datos obtenidos de las muestras de tipificación y para eliminar los ítems redundantes y los ítems que eran demasiado fáciles o demasiado difíciles. Siguiendo este criterio, se eliminaron como promedio dos ítems de cada prueba, y, en la medida de lo posible, se mantuvo el orden de los ítems dentro de las pruebas. Los ítems establecidos como puntos de comienzo en la versión de tipificación se mantuvieron como tales en la versión final.

Asimismo, no se modificaron las consignas para el niño, ni los ítems del cuaderno de estímulos. Aunque se modificó ligeramente la formulación de las instrucciones para el examinador para que fueran más claras, los cambios realizados no suponen ninguna alteración de los procedimientos de aplicación estandarizados de la escala.

Consistencia del orden de aplicación

El desarrollo de la estructura final de la escala incluyó también la determinación del tipo de prueba. El orden de aplicación de las pruebas no se modificó.

Adaptación española de la WISC-V

El objetivo de la adaptación española de la WISC-V ha sido proporcionar a los profesionales españoles una herramienta adaptada a sus necesidades prácticas y clínicas, y unos baremos actualizados, acordes con los cambios poblacionales españoles.

La adaptación española de la WISC-V se inició en mayo del 2014. A lo largo de estos meses no se ha realizado una simple traducción, sino una adaptación completa de la escala. Se ha llevado a cabo un desarrollo técnico y psicométrico completo que permite asegurar la validez y fiabilidad de la escala en la cultura española. Como es bien sabido, las diferencias culturales, lingüísticas y educativas entre países generan a menudo representaciones conceptuales diferentes de un mismo término. Por ello, la adaptación de la WISC-V, especialmente en el caso de las pruebas verbales, ha comportado una transposición cultural. El objetivo de una transposición cultural es encontrar las equivalencias entre las culturas que permitan adaptar la prueba conservando su esencia. Con este objetivo, en el proceso de adaptación española, y a partir del análisis del contenido original, se tuvieron en cuenta diversos criterios para seleccionar los ítems, redactar las consignas de aplicación y establecer los criterios de puntuación.

En este capítulo se presentan, en primer lugar, los procedimientos empleados para seleccionar los ítems de la versión experimental, las características de la muestra experimental, y los análisis que permitieron determinar el orden que debía ocupar cada ítem dentro de cada prueba. En segundo lugar, se presenta la información referente al proceso de tipificación, incluidas las modificaciones de los ítems respecto a la WISC-IV, las características de la muestra de tipificación y su representatividad. Por último, se exponen los métodos adoptados para la creación de los baremos y para la obtención de las puntuaciones compuestas, entre otros.

Fase experimental

Las pruebas que componen la WISC-V pueden clasificarse en función de múltiples criterios. Teniendo en cuenta su contenido, podemos diferenciar entre pruebas mayoritariamente verbales y pruebas con escaso contenido verbal. En este sentido, los ítems que tradicionalmente requieren ser adaptados son los de las pruebas verbales, puesto que con frecuencia, al traducir un ítem de un idioma a otro, puede perderse la esencia o el significado original que pretendía.

La adaptación de cualquier instrumento de evaluación psicológica se inicia con el análisis minucioso del contenido de la versión original en la que se basa. La versión experimental española de la WISC-V se apoyó en la versión original estadounidense, pero teniendo en cuenta la versión anterior de la escala para garantizar cierta continuidad (WISC-IV; Wechsler, 2011 [ed. orig., 2005]). Los objetivos de la fase experimental fueron traducir, revisar y adaptar los ítems de las pruebas verbales, crear nuevos ítems equivalentes a los originales, probar psicométricamente el funcionamiento de los ítems y establecer el orden que debían ocupar en la versión de tipificación.

Selección de los ítems experimentales

En la versión experimental de la WISC-V se incluyeron tanto los ítems verbales como los ítems gráficos de las pruebas con contenido verbal: *Información, Semejanzas, Vocabulario, Comprensión* y *Aritmética.*

El proceso de elaboración de la versión experimental se llevó a cabo en varios pasos. En primer lugar, el equipo de I+D junto con una experta traductora y filóloga tradujeron al castellano todos los ítems de las pruebas verbales de la versión original estadounidense. En segundo lugar, se examinaron en detalle tanto la traducción de los ítems verbales como las imágenes que ilustraban los ítems gráficos. Se decidió conservar, modificar o eliminar cada ítem en función de la idoneidad de su contenido y la adecuación a la cultura española. En el caso de los ítems conservados o modificados, se tuvo en cuenta que la traducción y/o la imagen ilustrada recogieran el significado y el objetivo primario de la versión original, así como que mantuvieran su equivalencia en cuanto a la dificultad. Finalmente, para reemplazar los ítems eliminados, se crearon ítems nuevos para cada prueba verbal.

La tabla 4.1 presenta un resumen de los ítems conservados, eliminados y nuevos en la versión experimental española, respecto a la versión original estadounidense. En la tabla puede observarse que la versión experimental contó con más ítems que los incluidos en la versión definitiva. Este procedimiento permitió, tras analizar los resultados de la muestra experimental, seleccionar los ítems con mejores propiedades psicométricas.

Tabla 4.1. Ítems conservados, eliminados y nuevos respecto a la versión estadounidense para la versión experimental española

	Versión estadounidense WISC-V	Versión experimental española WISC-V				
Prueba	Total ítems	Conservados sin modificaciones	Conservados con ligeras modificaciones	Eliminados	Nuevos	Total ítems experimentados
Información	31	28	1	2	10	39
Semejanzas	23	20	1	2	7	28
Vocabulario	29	25	1	3	8	34
Comprensión	19	16	0	3	8	24
Aritmética	34	34	0	0	0	34

Nota: Las cifras de la tabla representan la cantidad de ítems.

Al crear nuevos ítems para la versión experimental, se tuvieron en cuenta los siguientes criterios:

- Ítems con nivel de dificultad equivalente a los ítems de la versión original estadounidense.
- Ítems con respuestas fácilmente clasificables en 0, 1 o 2 puntos, según la prueba.
- Ítems sin significado ambiguo o diferente en otras zonas hispanohablantes.
- Ítems con una formulación clara y precisa.
- Ítems sin connotaciones religiosas o peyorativas.
- Ítems sin sesgo de género.
- Ítems apropiados para todas las edades.

Además, se tuvo en cuenta que los ítems cubrieran una amplia variedad de conocimientos y temáticas. Así, en *Información* se procuró que la proporción de ítems de las áreas de historia y geografía, ciencias naturales y cuerpo humano, entre otras, fuera equivalente. Con el objeto de evitar cualquier sobrerrepresentación en el formato de los ítems, en *Vocabulario* también se controló la frecuencia de aparición de verbos, sustantivos y adjetivos. En *Comprensión* y *Semejanzas*, todos los ítems se formularon o reformularon de manera precisa para evitar respuestas muy concretas o personalizadas que no responden al ítem. Y en *Aritmética* se tuvo en cuenta que las situaciones que planteaban los ítems resultaran familiares para los niños de todos los grupos de edad.

Finalmente, la versión experimental de la WISC-V estuvo compuesta por 39 ítems para *Información,* 28 para *Semejanzas,* 34 para *Vocabulario,* 24 para *Comprensión* y 34 para *Aritmética.* Para evaluar estos ítems, se diseñó una muestra experimental estratificada y representativa de los niños de la población española y se seleccionó a los examinadores que se encargarían de la recogida de datos.

Selección y formación de los examinadores

Con la colaboración de los colegios oficiales de psicólogos de toda España, se contactó con psicólogos infantiles que tenían amplia experiencia en evaluación de la inteligencia y de las aptitudes cognitivas y que procedían de diversos campos de la psicología (clínico, educativo, forense, etc.).

Los 27 psicólogos seleccionados, procedentes de distintas zonas geográficas de España, participaron en un curso de formación presencial donde se expusieron los objetivos y requerimientos de la fase experimental, los procedimientos específicos de aplicación de la WISC-V y los cambios respecto a la versión anterior de la escala.

La formación permitió homogeneizar los procedimientos de aplicación entre examinadores y aclarar las dudas referentes a las instrucciones de aplicación. Se pidió a todos ellos que, durante cada aplicación, transcribiesen palabra por palabra las respuestas de los niños y que indicaran sus reacciones frente a los ítems y al material. La corrección y la puntuación de las respuestas fueron realizadas por el departamento de I+D de Pearson Clinical & Talent Assessment.

Aplicación de la versión experimental

Para la aplicación de la versión experimental, se proporcionó a los examinadores todo el material necesario y una plantilla personalizada de los niños que cada uno debía reclutar, atendiendo a los criterios de estratificación. El material experimental incluía un manual donde se especificaban las instrucciones y consignas de aplicación y los criterios de inclusión de la muestra; el cuadernillo de anotación para el registro de todas las respuestas del niño y sus características sociodemográficas; y el cuaderno de estímulos que contenía los ítems gráficos de *Vocabulario* y *Aritmética.*

En la fase experimental de la WISC-V no se aplicaron los límites de tiempo y reglas de comienzo, retorno y terminación de las pruebas. Se pidió a los examinadores que aplicaran todos los ítems con la intención de valorar, posteriormente, la dificultad y la posición que debía ocupar cada ítem en cada prueba.

La aplicación de la forma experimental también permitió conocer las respuestas que con más frecuencia daban los niños de la muestra, y a partir de ellas se elaboraron los criterios de puntuación para la versión de tipificación.

Control y seguimiento

Para controlar y garantizar la calidad de los datos recogidos, se puso en marcha un sistema de evaluación «luz verde». Este proceso consistió en corregir y analizar minuciosamente el primer protocolo enviado por cada uno de los examinadores, para detectar cualquier tipo de error en la aplicación de las pruebas o en el registro de las respuestas. Si el protocolo cumplía los requisitos de calidad propuestos inicialmente por el departamento de I+D de Pearson, se daba al examinador «luz verde» para que prosiguiera sus aplicaciones. En caso contrario, se contactaba con el examinador para explicarle en detalle los errores u omisiones observados en su primer protocolo, y se le proporcionaban las indicaciones necesarias que debía tener en cuenta para sus próximas aplicaciones. El segundo protocolo enviado por el examinador era de nuevo evaluado cuidadosamente para comprobar la incorporación de las recomendaciones.

Para garantizar el cumplimiento de los plazos y la buena calidad de los datos recogidos, se pidió a los examinadores que enviaran al menos un protocolo semanalmente.

Muestra experimental

La recogida de datos de la fase experimental de la WISC-V se realizó entre junio del 2014 y septiembre del 2014; y se recogieron datos de una muestra de 223 niños, estratificada por edad, sexo, nivel educativo de los padres y zona geográfica, siguiendo las proporciones del último Censo de Población y Viviendas español (Instituto Nacional de Estadística; INE, 2011). Todos los padres de los niños que participaron en la fase experimental firmaron un documento de consentimiento informado.

Edad. La forma experimental se aplicó a once grupos de edad: 6:0-6:11 (n = 22), 7:0-7:11 (n = 20), 8:0-8:11 (n = 23), 9:0-9:11 (n = 19), 10:0-10:11 (n = 21), 11:0-11:11 (n = 20), 12:0-12:11 (n = 22), 13:0-13:11 (n = 20), 14:0-14:11 (n = 20), 15:0-15:11 (n = 18) y 16:0-16:11 (n = 18).

Sexo. La distribución por sexo fue equivalente en todos los grupos de edad. La muestra total estuvo compuesta por 110 niños (49.3 %) y 113 niñas (50.7 %).

Nivel educativo de los padres. Se clasificó el nivel educativo de los padres o tutores legales de los niños en cuatro categorías. Si el niño residía con uno de los dos padres, se asignaba el nivel educativo del padre con el que residía. En el caso de que ambos padres residieran con el niño y tuviesen distinto nivel educativo, se tuvo en cuenta el nivel educativo superior. La descripción de las categorías y el porcentaje de niños de la muestra experimental incluidos en cada categoría fueron los siguientes:

- Sin estudios: estudios primarios sin finalizar o inferior (4.5 %).
- Primer grado: graduado escolar (8.º de EGB) o 2.º de ESO o bachillerato elemental completados (9.4 %).
- Segundo grado: ESO o bachilllerato superior o formación profesional o COU completados (60.5 %).
- Tercer grado: diplomatura o licenciatura o titulaciones superiores completadas (25.6 %).

Zona geográfica. Se determinaron cuatro zonas geográficas para que la muestra experimental fuese representativa de todo el territorio peninsular e insular. La descripción de las cuatro zonas y el porcentaje de niños de la muestra experimental evaluados en cada zona fueron los siguientes:

- Norte: Asturias, Cantabria, Castilla y León, Galicia, La Rioja, Navarra y País Vasco (15.2 %).
- Sur: Andalucía, Islas Canarias y Murcia (32.3 %).
- Este: Cataluña, Comunidad Valenciana e Islas Baleares (29.2 %).
- Centro: Aragón, Castilla-La Mancha, Extremadura y Madrid (23.3 %).

Tipo de población. El tipo de población donde residían los niños no se consideró una variable de estratificación en esta fase. Sin embargo, teniendo en cuenta su valor descriptivo, a continuación se presenta la distribución de la muestra experimental en tres categorías. La descripción de cada categoría y el porcentaje de niños de la muestra experimental evaluados en cada tipo de población fueron los siguientes:

- Urbana: población igual o superior a 50 000 habitantes (45.8 %).
- Intermedia: población entre 10 000 y 49 999 habitantes (36.0 %).
- Rural: población inferior a 10 000 habitantes (18.2 %).

Análisis de los datos de la fase experimental

Los datos recogidos a partir de la muestra de la fase experimental fueron sometidos a análisis estadísticos. Solo los ítems que cumplieron con las cualidades psicométricas requeridas fueron seleccionados para la versión definitiva de la escala. En cada ítem se analizaron los siguientes aspectos:

- **Índice de dificultad.** El índice de dificultad de un ítem se define como el cociente entre el número de niños que han acertado el ítem y el número total de niños que han intentado resolver el ítem. El valor mínimo del índice de dificultad es 0 (ningún niño acierta el ítem) y el valor máximo es 1 (todos los niños que lo intentan lo aciertan). Si el índice de dificultad se acerca a 0, indica que el ítem ha resultado muy difícil; si se acerca a 1, indica que ha resultado muy fácil; y si se acerca a 0.5, indica una dificultad media. El índice de dificultad permite fijar el orden de presentación de los ítems para cada prueba.
- **El parámetro de dificultad según el modelo de Rasch.** El índice de dificultad tiene el inconveniente de ser dependiente de la muestra utilizada para calcularlo. Si los niños tienen un nivel de aptitud elevado, el ítem puede parecer fácil; si los niños tienen un nivel de aptitud bajo, el mismo ítem puede parecer difícil. La teoría de la respuesta al ítem (TRI) constituye un nuevo enfoque en psicometría que permite superar esta limitación calculando índices de dificultad invariantes respecto a la muestra. El modelo de Rasch es el índice más conocido de la TRI y permitió estimar el parámetro de dificultad de los ítems de la fase experimental.
- **Índice de discriminación.** Un buen ítem tiene que diferenciar, de la forma más clara posible, a los niños en función de su nivel de aptitud. Por ejemplo, los niños con un nivel de aptitud inferior a la media responderían incorrectamente a un ítem de dificultad mediana (0.5), mientras que los niños con un nivel de aptitud superior a la media responderían correctamente a ese mismo ítem. Si no es el caso, entonces ese ítem demuestra ser poco discriminativo. Para cada prueba, las correlaciones entre la puntuación en el ítem y la puntuación directa total se utilizaron como indicadores de la cualidad discriminante del ítem.

Los análisis estadísticos realizados sobre la muestra de la fase experimental permitieron seleccionar los ítems que formarían parte de la versión de tipificación española de la WISC-V y asegurar sus cualidades métricas. A partir de esta versión se llevó a cabo el estudio de tipificación.

Fase de tipificación

Finalizada la fase experimental, un panel de expertos en evaluación y adaptación de instrumentos examinó las pruebas que por su escaso contenido verbal no habían sido probadas en la fase experimental, y dictaminó que el resto de las pruebas no mostraban sesgos aparentes. Esto es así porque, durante la fase de desarrollo original de la WISC-V, el equipo de I+D estadounidense consultó con el resto de los equipos de investigación europeos la adecuación del contenido de estos ítems en cada uno de sus países. La intención no era otra que la de elaborar una herramienta que fuera, en la medida de lo posible, culturalmente adecuada a la mayoría de los países, o que al menos requiriera mínimos cambios.

Modificaciones respecto a la WISC-IV

Respecto a su predecesora, en la versión española de la WISC-V se aplicaron las siguientes modificaciones: se eliminaron las pruebas *Adivinanzas*, *Figuras incompletas* y *Conceptos*; se añadieron tres pruebas nuevas *Balanzas*, *Puzles visuales* y *Span de dibujos*; se actualizaron los ítems; y se simplificaron las consignas y los procedimientos de aplicación. La tabla 4.2 resume los principales cambios realizados en la adaptación española de la WISC-V, respecto a la adaptación española de la WISC-IV.

Tabla 4.2. Cambios en la adaptación española de la WISC-V respecto a la WISC-IV

	Adaptación española WISC-IV	Adaptación española WISC-V			
Prueba	**Total ítems**	**Conservados***	**Eliminados**	**Nuevos**	**Total ítems**
Cubos	14	5	9	8	13
Semejanzas	23	12	11	11	23
Matrices	35	12	23	20	32
Dígitos orden directo	16 intentos	1	15	17	18 intentos
Dígitos orden inverso	16 intentos	2	14	16	18 intentos
Dígitos orden creciente	—	—	—	18	18 intentos
Claves (parte A)	59	0	59	75	75
Claves (parte B)	119	0	119	117	117
Vocabulario	36	11	25	18	29
Balanzas	—	—	—	34	34
Puzles visuales	—	—	—	29	29
Span de dibujos	—	—	—	26	26
Búsqueda de símbolos (parte A)	45	0	45	42	42
Búsqueda de símbolos (parte B)	60	0	60	60	60
Información	33	10	23	21	31
Letras y números	30 intentos	4	26	26	30 intentos
Cancelación/Animales	2	0	2	2	2
Comprensión	21	7	14	12	19
Aritmética	34	24	10	10	34

Nota: Las cifras de la tabla representan la cantidad de ítems.

* Algunos de los ítems conservados fueron ligeramente modificados.

Diseño de la muestra de tipificación

Para diseñar una muestra lo más representativa posible de la población española, se siguieron las proporciones y los datos del Censo de Población y Viviendas español (INE, 2011), que era el que aportaba todos los datos referentes al cruce de las variables de estratificación. La muestra de tipificación se planeó sobre una muestra de 1100 niños, estratificada según las variables demográficas siguientes: edad, sexo, nivel educativo y zona geográfica.

Edad. La muestra se diseñó teniendo en cuenta 11 grupos de edad, de 100 niños cada uno: 6:0-6:11, 7:0-7:11, 8:0-8:11, 9:0-9:11, 10:0-10:11, 11:0-11:11, 12:0-12:11, 13:0-13:11, 14:0-14:11, 15:0-15:11 y 16:0-16:11.

Sexo. Porcentaje equivalente de niños y niñas.

Nivel educativo de los padres. Se determinó el nivel educativo de los padres o tutores legales de los niños en función de tres categorías. Los resultados de los análisis llevados a cabo durante la fase experimental no revelaron diferencias significativas entre las categorías *Sin estudios* y *Primer grado* por lo que ambas clasificaciones se agruparon en una única categoría denominada *Primer grado*. Si el niño residía con uno de los dos padres, se asignaba el nivel educativo del padre con el que residía. En el caso de que ambos padres residieran con el niño y tuviesen distinto nivel educativo, se tuvo en cuenta el nivel educativo superior. Las tres categorías utilizadas para estratificar la muestra fueron:

- Primer grado: estudios primarios o inferior, graduado escolar (8.º de EGB) o 2.º de ESO o bachillerato elemental completados.
- Segundo grado: ESO o bachilllerato superior o formación profesional o COU completados.
- Tercer grado: diplomatura o licenciatura o titulaciones superiores completadas.

Zona geográfica. A fin de que los baremos estuvieran construidos a partir de una muestra muy diversa, procedente de distintas zonas geográficas de nuestro país, se planificó la recogida de datos en todo el territorio peninsular e insular. Se determinaron cuatro zonas geográficas:

- Norte: Asturias, Cantabria, Castilla y León, Galicia, La Rioja, Navarra y País Vasco.
- Sur: Andalucía, Islas Canarias y Murcia.
- Este: Cataluña, Comunidad Valenciana e Islas Baleares.
- Centro: Aragón, Castilla-La Mancha, Extremadura y Madrid.

Tipo de población. También se consideró como una variable de estratificación el tipo de población donde residían los niños. Siguiendo la clasificación utilizada por el censo, se definieron tres categorías:

- Urbana: población igual o superior a 50 000 habitantes.
- Intermedia: población entre 10 000 y 49 999 habitantes.
- Rural: población inferior a 10 000 habitantes.

Paralelamente al proyecto de tipificación, se diseñó la muestra para los estudios de fiabilidad test-retest y de validez de la WISC-V con su predecesora, la WISC-IV. Se recogieron dos muestras estratificadas por edad, sexo y nivel educativo de los padres. Las características de estas muestras y los resultados de estos estudios pueden consultarse en los capítulos 5 y 6 de este manual, respectivamente.

El plan de muestreo presentado en este apartado permitió elaborar las tablas de estratificación personalizadas para cada examinador.

Selección y formación de los examinadores

Siguiendo los mismos criterios que en la fase experimental, se seleccionó a 85 examinadores para la fase de tipificación.

Se realizó un curso de formación para todos los examinadores. Durante la formación se expusieron los objetivos y los requerimientos de la fase de tipificación, así como una explicación detallada acerca del procedimiento específico de aplicación de la WISC-V en esta fase y los cambios respecto a la versión anterior.

La formación permitió homogeneizar los procedimientos de aplicación entre examinadores y aclarar las dudas referentes a las instrucciones de aplicación. Se pidió a todos los examinadores que durante las aplicaciones registraran todas las respuestas del niño, y observaciones relevantes, a fin de facilitar la labor de puntuación realizada por el equipo de corrección de Pearson Clinical & Talent Assessment.

Aplicación de la versión de tipificación

La fase de tipificación se inició en septiembre del 2014 y concluyó en mayo del 2015.

Para la aplicación de la forma de tipificación se proporcionó a todos los examinadores todo el material necesario y una plantilla personalizada de los niños que cada uno debía reclutar, atendiendo a los criterios de estratificación. El material de tipificación incluía un manual donde se especificaban las instrucciones, las consignas de aplicación y los criterios de inclusión de la muestra; un cuadernillo de anotación para el registro de las respuestas del niño y sus características sociodemográficas; dos cuadernos de estímulos con los ítems gráficos de las pruebas; dos cuadernillos de respuesta para las pruebas *Claves*, *Búsqueda de símbolos* y *Cancelación*; y 9 cubos bicolores necesarios para aplicar *Cubos*. Estos materiales fueron iguales a los definitivos.

En esta fase, y a diferencia de la fase experimental, se incluyeron las limitaciones de tiempo, las reglas de comienzo, retorno y terminación. Se pidió a los examinadores que aplicaran la escala siguiendo de manera estricta las reglas y los procedimientos definitivos, a fin de obtener resultados válidos para la creación de los baremos.

Puntuación de los protocolos

Para controlar y garantizar la calidad de los datos recogidos, se puso en marcha el mismo sistema de evaluación «luz verde» adoptado en la fase experimental.

Para la corrección de las pruebas verbales de la fase de tipificación se partió de las respuestas más frecuentes que los niños dieron a los ítems en la fase experimental. A fin de crear los criterios de puntuación para las pruebas verbales, y diferenciar bien las respuestas en 0, 1 y 2 puntos, se analizaron las respuestas más frecuentes que los niños dieron a los ítems en la fase experimental. Esta labor se realizó con la ayuda de una filóloga experta y en función de las definiciones recogidas en distintos diccionarios de lengua castellana. Los criterios de puntuación para el resto de las pruebas no requirieron ajustes importantes puesto que son en su mayoría muy objetivas. La puntuación de los protocolos fue realizada por el equipo de correctores del departamento de I+D de Pearson Clinical & Talent Assessment, previamente formado y entrenado para esta tarea. Aunque el sistema de corrección de las pruebas de la WISC-V no implica una dificultad elevada, el control de la fiabilidad de la puntuación es fundamental. Por ello, se efectúo una doble corrección, por parte de dos correctores independientes sobre 150 protocolos y para todas las pruebas. Los resultados del análisis de acuerdo entre correctores pueden consultarse en el capítulo 5 de este manual.

Muestra de tipificación real

La muestra de tipificación estuvo compuesta por 1008 niños (frente a los 1100 planeados), de edades de 6:0 a 16:11 años, repartidos en los 11 grupos de edad previstos en el diseño de la muestra. La muestra estuvo compuesta por 508 niños (50.4%) y 500 niñas (49.6%).

En la tabla 4.3 se muestra el tamaño de la muestra por sexo y para cada grupo de edad.

Tabla 4.3. Tamaño de la muestra de tipificación por sexo y por grupo de edad

	Tamaño de la muestra		
Grupo de edad	Varones	Mujeres	Total
6	52	46	98
7	55	48	103
8	52	48	100
9	42	46	88
10	51	44	95
11	47	50	97
12	48	47	95
13	38	43	81
14	42	47	89
15	42	41	83
16	39	40	79

En el diseño de la muestra se puso una atención especial en que participaran niños pertenecientes a todas las comunidades autónomas de España, objetivo que se consiguió ampliamente; casi todas las comunidades españolas están representadas en la muestra de tipificación. La figura 4.1 muestra la distribución de los profesionales que recogieron los datos de la muestra de tipificación y, por tanto, las zonas de procedencia de los participantes de la muestra.

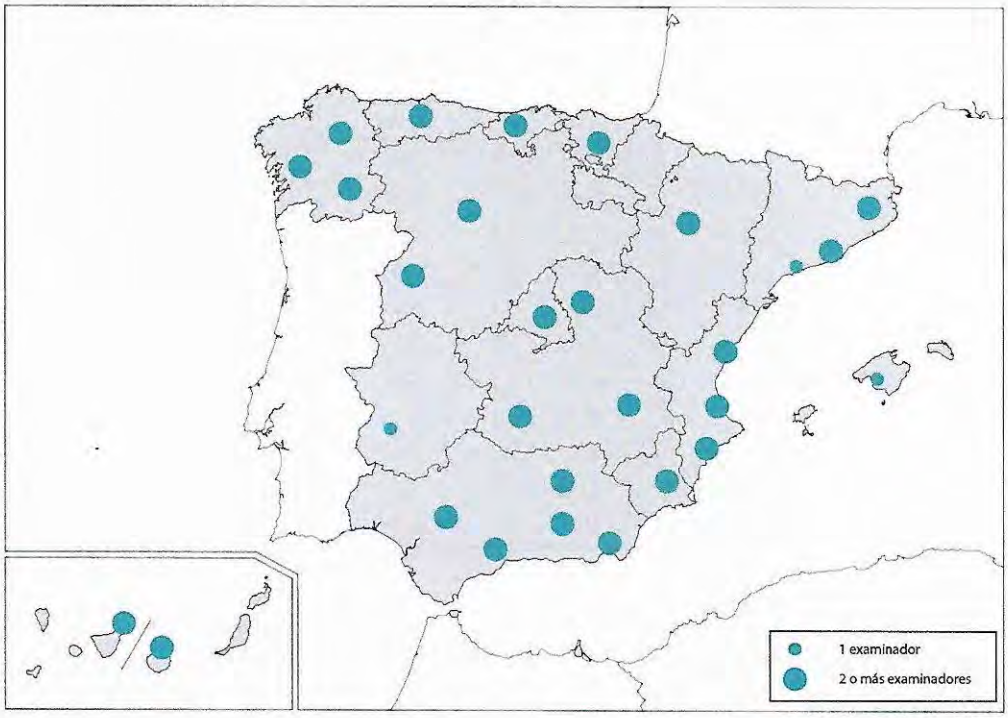


Figura 4.1. Distribución de los examinadores de la fase de tipificación

Los criterios de estratificación de las variables sexo, nivel educativo, tipo de población y zona geográfica se respetaron de forma escrupulosa por parte de los examinadores para conseguir que la representatividad de la muestra fuese muy elevada.

Representatividad de la muestra

Las tablas 4.4-4.7 muestran los porcentajes teóricos de población, basados en el censo, y los porcentajes reales conseguidos en la muestra de tipificación. Los porcentajes se presentan para las variables nivel educativo, tipo de población, zona geográfica y sexo, y en función de cada grupo de edad. Los porcentajes que se muestran a continuación están redondeados al decimal más próximo.

Tabla 4.4. Porcentajes teóricos (censo) y reales (muestra de tipificación), según edad y nivel educativo

	Nivel educativo					
	Primer grado		Segundo grado		Tercer grado	
Edad	Teórico	Real	Teórico	Real	Teórico	Real
6	8.2	10.2	61.9	56.1	29.9	33.7
7	8.4	7.8	62.9	57.3	28.7	35.0
8	9.3	9.0	63.1	62.0	27.6	29.0
9	9.6	10.3	64.3	58.0	26.1	31.8
10	10.5	11.6	64.2	58.9	25.3	29.5
11	11.6	11.3	64.4	62.9	24.0	25.8
12	12.4	10.6	64.7	61.1	22.9	28.4
13	13.3	12.4	64.0	63.0	22.8	24.7
14	13.8	10.1	64.3	65.2	22.0	24.7
15	14.7	10.8	63.8	67.5	21.5	21.7
16	15.3	5.1	64.0	65.8	20.7	29.1

Tabla 4.5. Porcentajes teóricos (censo) y reales (muestra de tipificación), según edad y tipo de población

	Tipo de población					
	Urbana		Intermedia		Rural	
Edad	Teórico	Real	Teórico	Real	Teórico	Real
6	51.1	52.0	29.6	23.5	19.4	24.5
7	50.4	56.3	29.6	29.1	20.1	14.6
8	51.1	51.0	29.0	27.0	19.9	22.0
9	51.0	53.4	28.9	25.0	20.1	21.6
10	50.8	50.5	29.4	24.2	19.8	25.3
11	50.6	59.8	28.7	20.6	20.7	19.6
12	50.4	57.4	29.1	21.3	20.5	21.3
13	50.1	54.3	29.1	28.4	20.9	17.3
14	50.2	51.7	28.9	28.1	20.8	20.2
15	50.3	50.6	28.9	27.7	20.8	21.7
16	49.7	60.3	29.2	23.1	21.1	16.7

Tabla 4.6. Porcentajes teóricos (censo) y reales (muestra de tipificación), según edad y sexo

	Sexo			
	Varón		Mujer	
Edad	Teórico	Real	Teórico	Real
6	50.0	53.1	50.0	46.9
7	50.0	53.4	50.0	46.6
8	50.0	52.0	50.0	48.0
9	50.0	47.7	50.0	52.3
10	50.0	53.7	50.0	46.3
11	50.0	48.5	50.0	51.5
12	50.0	50.5	50.0	49.5
13	50.0	46.9	50.0	53.1
14	50.0	47.2	50.0	52.8
15	50.0	50.6	50.0	49.4
16	50.0	49.4	50.0	50.6

Tabla 4.7. Porcentajes teóricos (censo) y reales (muestra de tipificación), según edad y zona geográfica

	Zona geográfica							
	Norte		Sur		Este		Centro	
Edad	Teórico	Real	Teórico	Real	Teórico	Real	Teórico	Real
6	18.0	18.4	27.9	27.6	30.5	30.6	23.5	23.5
7	18.1	16.5	28.0	30.1	30.1	26.2	23.7	27.2
8	18.3	19.0	28.2	29.0	29.9	29.0	23.6	23.0
9	18.3	14.8	28.7	31.8	29.6	29.5	23.4	23.9
10	18.2	14.7	28.6	34.7	29.7	27.4	23.5	23.2
11	18.3	16.5	28.9	29.9	29.3	26.8	23.6	26.8
12	18.6	15.8	29.0	34.7	29.3	25.3	23.1	24.2
13	18.5	18.5	28.9	32.1	29.0	30.9	23.5	18.5
14	18.3	14.6	29.3	31.5	29.1	30.3	23.3	23.6
15	18.4	15.7	29.2	34.9	29.0	31.3	23.3	18.1
16	18.3	13.9	29.6	32.9	28.9	34.2	23.3	19.0

Tal y como puede observarse en las tablas, 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.

Los gráficos que siguen a continuación (figura 4.2) reflejan la distribución de las características demográficas de la muestra de tipificación respecto a la distribución total de la población procedente del censo.

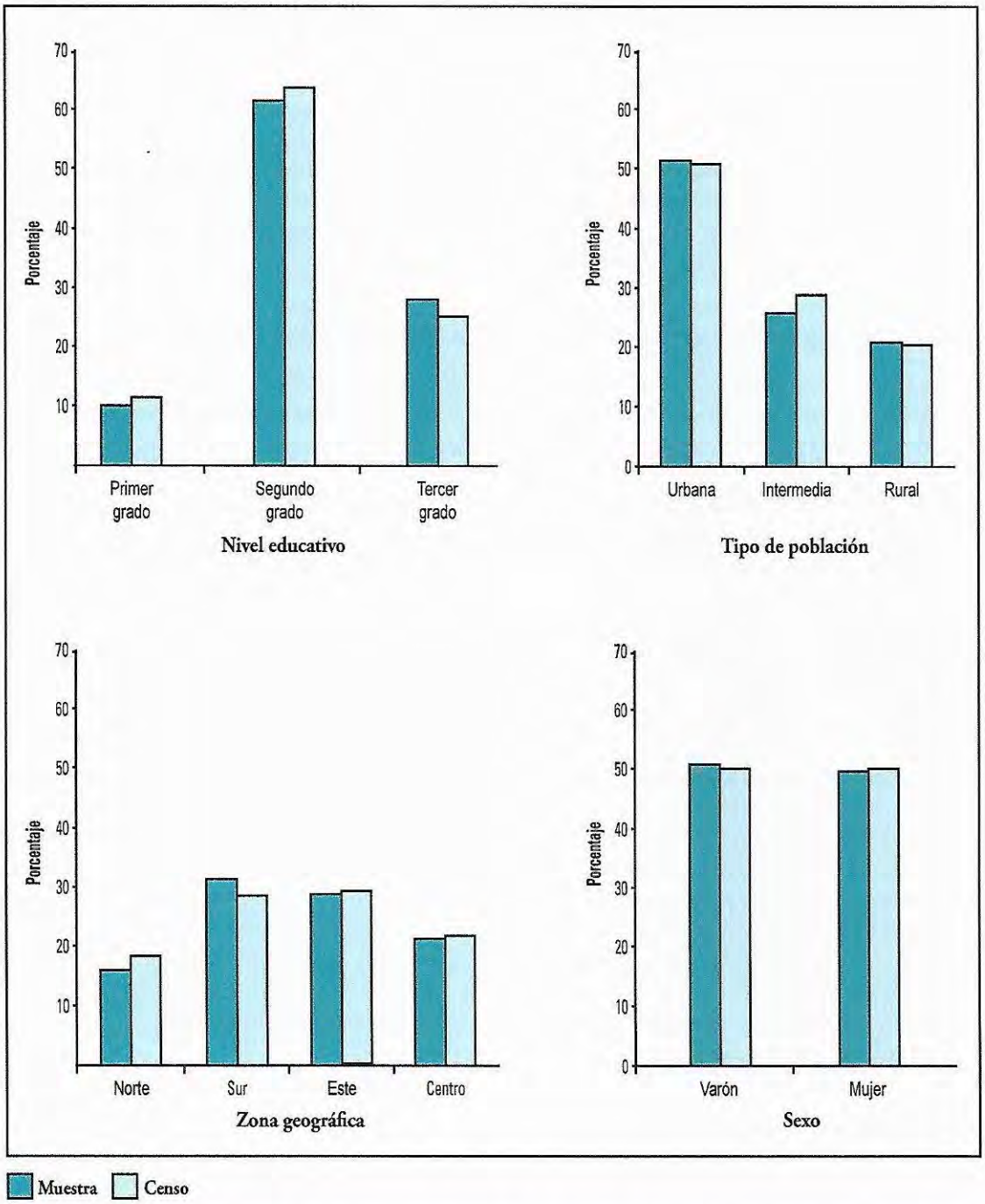


Figura 4.2. Comparación entre los datos del censo y los datos de la muestra de tipificación según las características demográficas

Como puede observarse en los gráficos, los porcentajes son extremadamente similares. Las diferencias más grandes entre lo teórico y lo real se observaron para las variables nivel educativo (tercer grado) y tipo de población (intermedia). Sin embargo, estas diferencias fueron mínimas.

La extrema semejanza entre los porcentajes del censo y los de la muestra de tipificación se consiguió porque cada examinador contó con una plantilla única y personalizada donde figuraban los perfiles de cada uno de los niños que debía reclutar. Aunque la elaboración de estos perfiles supuso una tarea de reclutamiento ardua, gracias al diseño personalizado y controlado en función de las características del censo y al esfuerzo de los examinadores, se consiguió una muestra de tipificación notablemente representativa de la población española.

Así, los baremos construidos a partir de los datos de la muestra de tipificación, los estudios de validez y de fiabilidad permiten una interpretación precisa y correcta de los resultados obtenidos en la WISC-V.

Creación de los baremos

En esta sección se resume el procedimiento seguido para la creación de los baremos, que se inició con la preparación de los datos (establecimiento de los puntos de comienzo y reglas de terminación de las pruebas) y finalizó con la obtención de las puntuaciones típicas (de las pruebas, de procesamiento y compuestas) y las puntuaciones de edad equivalente.

Establecimiento de los puntos de comienzo y de las reglas de terminación

Los puntos de comienzo y de terminación, fijados a partir de los datos de la versión estadounidense, se incorporaron a la escala para evitar los sentimientos de frustración causados por la presencia de ítems demasiado fáciles o difíciles. Los puntos de comienzo permiten limitar el uso de las reglas de retorno y garantizar que la mayoría de los niños podrán responder correctamente a los primeros ítems de cada prueba. Por su parte, las reglas de terminación cumplen un doble objetivo: limitar el número de errores consecutivos cometidos por el niño y garantizar que se ha aplicado el número potencialmente suficiente de ítems para los que se espera un acierto.

En la versión de tipificación, los ítems de cada prueba se ordenaron siguiendo un nivel de dificultad progresivo. Los puntos de comienzo se fijaron para reducir el número de ítems aplicados, garantizando, al mismo tiempo, que al menos el 95 % de las puntuaciones directas de las pruebas y al menos el 98 % de las puntuaciones escalares de las pruebas no cambiaban tras la aplicación de los nuevos puntos de comienzo. En la versión final se mantuvieron los mismos puntos de comienzo.

Las reglas de terminación de la versión de tipificación se establecieron generosamente para permitir que los niños tuvieran la oportunidad de realizar los ítems para los que se esperaba un acierto, aunque se limitara igualmente la cantidad de ítems presentados. Finalmente, se ajustaron las reglas de terminación para cada prueba a partir de los estudios empíricos sobre los datos de tipificación, siempre y cuando se cumplieran una serie de criterios. En primer lugar, el coeficiente de correlación por rangos de las puntuaciones directas registradas antes y después del ajuste era igual o superior a 0.98. En segundo lugar, menos del 5 % de las puntuaciones directas habían cambiado después del ajuste; y en los casos en que se habían producido cambios, estos eran iguales o menores a 2 puntos. Y, finalmente, el patrón observado en los cambios de puntuaciones de los ítems era aleatorio en la muestra de tipificación.

Establecimiento de las bonificaciones por tiempo

Los análisis de los datos de tipificación estadounidenses para las pruebas *Cubos* y *Claves* (solo parte A) revelaron que, si se concedían puntos extra por rapidez y precisión en la ejecución, se reducía el efecto techo y aumentaba la discriminación entre los niños con altas capacidades. Para estas pruebas se evaluó la distribución de los tiempos de ejecución de los niños que obtuvieron la puntuación directa máxima (niños que completaban correctamente la construcción o todos los símbolos dentro del tiempo límite), y se estableció el sistema de puntos de bonificación para cada prueba de forma que se concedieran más puntos cuanto menor fuera el tiempo de ejecución.

Obtención de las puntuaciones escalares

Los datos de tipificación española se obtuvieron utilizando un método de tipificación inferencial (Wilkins, Rolfhus, Weiss y Zhu, 2005; Zhu y Chen, 2011). Para cada uno de los grupos de edad de la muestra de tipificación se calcularon varios parámetros de las puntuaciones (medias, desviaciones típicas y asimetrías) y se representaron gráficamente en todas las edades. Los datos de los parámetros se ajustaron a varias regresiones polinómicas, que abarcaban desde polinomios lineales hasta polinomios de cuarto grado. Las mejores curvas para cada parámetro se seleccionaron a partir de la consistencia con las expectativas teóricas subyacentes y con el patrón de las curvas de crecimiento observado en la WISC-V y en otras muestras de tipificación comparables (p. ej., WISC-IV). Estas curvas seleccionadas se usaron para obtener estimaciones de los parámetros de la población, que luego permitieron generar distribuciones teóricas para los grupos de edad normativos, ofreciendo porcentajes para cada puntuación directa. Mediante el método del percentil medio del intervalo, estos porcentajes se convirtieron en puntuaciones escalares. Seguidamente, se examinó la progresión de las puntuaciones escalares dentro de cada grupo de edad y de un grupo de edad a otro, y las pequeñas irregularidades se corrigieron mediante suavizado.

Las puntuaciones de las pruebas y las puntuaciones de procesamiento se convirtieron en puntuaciones escalares con una media de 10, una desviación típica (*Dt*) de 3 y un rango de 1-19. Las puntuaciones escalares correspondientes a las puntuaciones directas de las pruebas se presentan en la tabla A.1 del *Manual de aplicación y corrección* en intervalos de 4 meses.

Obtención de las puntuaciones compuestas

Las puntuaciones compuestas (p. ej., ICV, CIT, IRC) se basan en la suma correspondiente de las puntuaciones escalares ajustadas a la edad. Para elaborar las tablas de puntuaciones compuestas, se calculó la suma de las puntuaciones escalares de cada puntuación compuesta para cada niño de la muestra de tipificación, utilizando las puntuaciones escalares de las pruebas principales que contribuyen a esa puntuación compuesta.

Las medias y las desviaciones típicas de la suma de las puntuaciones escalares de cada puntuación compuesta figuran, por franja de edad, en las tablas del anexo A de este manual. Estos datos son muy similares de una edad a otra, y el análisis de varianza no reveló diferencias significativas de las sumas promedio de las puntuaciones escalares para cada índice en función de la edad. Asimismo, los resultados del test de Bartlett para la homogeneidad de varianzas revelaron que las varianzas de las sumas de las puntuaciones escalares no diferían significativamente de una edad a otra. Además, el examen de las distribuciones de las sumas de las puntuaciones escalares mostró que estas sumas se aproximaban a la distribución normal. En consecuencia, los grupos de edad fueron combinados para elaborar las tablas de puntuaciones compuestas equivalentes.

Para cada escala se utilizó la distribución de la suma de las puntuaciones escalares para obtener los percentiles en el medio del intervalo correspondiente, que se convirtieron en puntuaciones típicas con una media 100 y una desviación típica 15, y un rango de 45-155 para las escalas Comprensión verbal, Visoespacial, Razonamiento fluido, Memoria de trabajo, Velocidad de procesamiento, Razonamiento cuantitativo y Memoria de trabajo auditiva; y un rango de 40-160 para la Escala total y las escalas No verbal, Capacidad general y Competencia cognitiva. Las distribuciones de las puntuaciones típicas resultantes se suavizaron visualmente para eliminar posibles irregularidades y garantizar que la distribución era aproximadamente normal, tratando en todo momento de mantener las medias y desviaciones típicas de las escalas próximas a 100 y 15, respectivamente (Allen y Yen, 2002; Cureton y Tukey, 1951). Las puntuaciones compuestas correspondientes a las sumas de las puntuaciones escalares de las escalas Comprensión verbal, Visoespacial, Razonamiento fluido, Velocidad de procesamiento y Escala Total se presentan en las tablas A.2-A.7 del *Manual de aplicación y corrección*; y las puntuaciones compuestas correspondientes a las sumas de las puntuaciones escalares de las escalas Razonamiento cuantitativo, Memoria de trabajo auditivo, No verbal, Capacidad general y Competencia cognitiva figuran en las tablas C.1-C.5 de dicho manual.

Obtención de las tasas base

Para las medidas con distribuciones no normales y con un gran sesgo, no es posible obtener puntuaciones escalares. Estas medidas, que incluyen las puntuaciones de span, se presentan como tasas base (porcentajes acumulados). Las tasas base son el porcentaje de niños de la muestra de tipificación que tienen una edad determinada y que muestran un comportamiento en concreto, no los percentiles basados en la transformación *z* que permiten determinar dónde se sitúan las puntuaciones respecto a la distribución normal. Las puntuaciones de tasas base no se suavizaron, por lo que se conservan las irregularidades entre los grupos de edad. Estos valores son tasas base puras de la ocurrencia de una conducta determinada en la muestra de tipificación.

Cálculo de los valores críticos para las diferencias entre puntuaciones

Puntos fuertes y débiles

Las puntuaciones de los índices primarios pueden compararse con la media de los índices primarios (MIP) o con el CI total. Los valores críticos que se utilizan para comparar los índices primarios y la MIP se calcularon con la fórmula de Davis (1959):

$$z\sqrt{\left(\sum_{j}^{i} ETM_a^2\right)/K^2 + \left[\frac{K-2}{K}\right]\left(ETM_i^2\right)}$$

donde *z* es el valor de la curva normal asociado con el nivel de significación bilateral requerido, calculado con la corrección propuesta por Dunn-Šidák para las comparaciones múltiples (Seidler, Vondáček y Saxl, 2000); *K* es el número de índices que se comparan, $\sum_{j}^{i} ETM_a^2$ es la suma del cuadrado de los *ETM* de los índices que se comparan; y ETM_i^2 es el cuadrado del *ETM* del índice que se analiza.

Dado que algunas pruebas principales también contribuyen al CI total, el error típico de la diferencia (ET_{dif}) usado para obtener los valores críticos para comparar índices primarios y el CI total se ha calculado con la fórmula siguiente:

$$ET_{dif} = \sqrt{ETM_{xi}^2 + ETM_{xj}^2 - 2\sqrt{\frac{\sum ETM_i^2}{\sum ETM_j^2}}\left(ETM_{xi}\right)\left(ETM_{xj}\right)}$$

donde ETM_{xi} y ETM_{xj} son los errores típicos de medida del índice y del CI total, $\sum ETM_i^2$ es la suma del cuadrado de los errores típicos de medida de todas las pruebas compartidas por el índice y el CI total, y $\sum ETM_j^2$ es la suma del cuadrado de los errores típicos de medida de todas las pruebas que conforman el CI total.

Al multiplicar el ET_{dif} por un valor *z* apropiado, obtenido con la corrección de Dunn-Šidák para las comparaciones múltiples, se obtiene el valor crítico, es decir, la diferencia requerida para la significación estadística en cualquier nivel de confianza. La misma metodología se ha usado para calcular el ET_{dif} que figura en la tabla C.6 del *Manual de aplicación y corrección*, para las diferencias entre el ICG y el CI total y entre el IMT y el IMTA.

La tabla B.3 del *Manual de aplicación y corrección* presenta los valores críticos para comparar cada una de las pruebas principales con la media de las puntuaciones escalares de las pruebas principales (MPE-P) o con la media de las puntuaciones escalares de las pruebas que conforman el CI total (MPE-CI). Para comparar las pruebas principales con la MPE-P, o las pruebas que conforman el CI total con la MPE-CI, se calcularon los valores críticos utilizando la fórmula usada para comparar cada uno de los índices primarios con la MIP.

Los valores críticos para comparar las pruebas principales que no contribuyen al CI total (*Puzles visuales, Span de dibujos* y *Búsqueda de símbolos*), ni a la MPE-CI, se calcularon mediante la fórmula siguiente sin ajuste de la varianza compartida:

$$z\sqrt{\left(\sum_{j}^{i} ETM_a^2\right) / K^2 + ETM_i^2}$$

donde z el valor de la curva normal asociado con el nivel de significación bilateral requerido, calculado con la corrección propuesta por Dunn-Šidák para las comparaciones múltiples; $\sum_{j}^{i} ETM_a^2$ es la suma del cuadrado de los *ETM* de las pruebas que conforman el CI total; ETM_i^2 es el cuadrado de los *ETM* de la prueba que se analiza; y K es el número de pruebas que contribuyen al CI total.

Comparación entre índices, pruebas o puntuaciones compuestas

La comparación entre índices, pruebas o puntuaciones compuestas describe la diferencia entre dos índices, dos pruebas, una puntuación de una prueba y una puntuación de procesamiento, o dos puntuaciones de procesamiento. Con excepción de las comparaciones ICG-CIT e IMT-IMTA, que figuran en la tabla C.6 del *Manual de aplicación y corrección*, los valores críticos para la comparación entre índices, pruebas o puntuaciones compuestas se calcularon con la fórmula siguiente

$$z\sqrt{ETM_a^2 + ETM_b^2}$$

donde z es el valor de la curva normal asociado con el nivel de significación bilateral requerido, y ETM_a y ETM_b son los errores típicos de medida de las dos puntuaciones. Los valores críticos para las comparaciones ICG-CIT e IMT-IMTA, dado que son puntuaciones que comparten el contenido de las pruebas, se calcularon utilizando la fórmula usada para comparar los índices primarios y el CI total.

Cálculo de las puntuaciones de edad equivalente

La edad equivalente de una puntuación directa indica la edad en la que dicha puntuación es igual a la media. A fin de obtener las edades equivalentes de una prueba, se determinó la puntuación directa correspondiente a una puntuación escalar de 10 para cada uno de los 33 grupos de edad tipificados (los 11 grupos de edad tipificados en intervalos de 4 meses), y dicha puntuación directa se fijó como punto medio del rango de edad de cada uno de los grupos (p. ej., 6:2 para el grupo de edad 6:0-6:3). Si la misma puntuación directa se correspondía con dos grupos de edad sucesivos (p. ej., 6:0-6:3 y 6:4-6:7), se asignaba al grupo de menor edad (p. ej., 6:2). Si la misma puntuación directa se correspondía con tres grupos de edad sucesivos (p. ej., 6:0-6:3, 6:4-6:7 y 6:8-6:11), se asignaba al grupo de edad intermedia (p. ej., 6:6). Si el mismo rango de puntuaciones directas se correspondía con una puntuación escalar de 10 en grupos adyacentes (p. ej., puntuaciones directas de 18 y 19 que se corresponden a una puntuación escalar de 10 en grupos de edad adyacentes), se dividían los valores del rango a partes iguales entre los grupos de edad adyacentes.

En la tabla A.9 del *Manual de aplicación y corrección* se presentan las edades equivalentes a las puntuaciones directas de cada prueba. Las pruebas *Claves* y *Búsqueda de símbolos* están formadas por dos partes que se aplican en función de la edad cronológica del niño. Se calcularon las puntuaciones de edad equivalente para cada parte de forma independiente. Las puntuaciones de edad equivalente correspondientes a cada una de las partes figuran sombreadas en dicha tabla.

Ha de tenerse en cuenta que las puntuaciones de edad equivalente a menudo se malinterpretan y tienen limitaciones psicométricas; por ello, deben utilizarse con cautela. Las puntuaciones de edad equivalente no ofrecen información comparativa de un mismo grupo de edad; esta información se obtiene mediante las puntuaciones típicas y los percentiles. Las puntuaciones de edad equivalente indican rendimientos medios de los distintos grupos de edad y no proporcionan

información sobre la distribución de las puntuaciones en un determinado grupo. Una puntuación de edad equivalente de un niño puede sugerir la presencia de un déficit cuando, de hecho, el niño se sitúa en el rango medio de funcionamiento para su edad. En el capítulo 7 de este manual se ofrece más información sobre las limitaciones y el uso adecuado de las puntuaciones de edad equivalente.

Fiabilidad

Las propiedades psicométricas de un test proporcionan información a los profesionales sobre la exactitud de las puntuaciones obtenidas. Para realizar un adecuado análisis e interpretación de los resultados logrados tras la aplicación de la escala, es fundamental que el profesional conozca la fiabilidad y estabilidad, los errores típicos de medida, los intervalos de confianza, las diferencias significativas entre puntuaciones (valores críticos) y la frecuencia con la que aparecen esas diferencias (tasas base).

Es un hecho bien documentado que los contrastes estadísticos de hipótesis presentan serias limitaciones (Balluerka, Gómez e Hidalgo, 2005; Cohen, 1990, 1994; Dunst y Hamby, 2012; Hubbard y Lindsay, 2008; Kirk, 2001; Morris y Fritz, 2013). Se recomienda que los intervalos de confianza y el tamaño del efecto se presenten junto con las puntuaciones obtenidas y la significación estadística (AERA, 2006; Morris y Fritz, 2013). Por ello, en este manual se presentan los intervalos de confianza y el tamaño del efecto, así como las pruebas de significación estadística y el valor de *p*, a fin de demostrar la fiabilidad y la validez de la escala.

En la presentación de los resultados psicométricos se han seguido una serie de criterios. Por una parte, todos los análisis usan los valores tradicionales para el nivel de significación (alfa = .05). Por otra parte, en este manual, el término *diferencia típica* se refiere a la *d* de Cohen. Los valores del tamaño del efecto se interpretan siguiendo las sugerencias de Cohen (1988, 1992): un tamaño del efecto de .20 se considera pequeño; de .50, moderado; y de .80 grande. Aunque estas referencias no describen por completo todos los aspectos de la interpretación del tamaño del efecto, en atención a la simplicidad y a la consistencia con anteriores versiones de las escalas de inteligencia de Wechsler, los valores de *d* comprendidos entre .20 y .49 se consideran pequeños, los comprendidos entre .50 y .79 se consideran moderados, y los valores de .80 o superiores se consideran grandes.

Si bien en este manual se siguen los criterios habituales respecto a la significación estadística y a la interpretación del tamaño del efecto, siempre deben tenerse en cuenta los valores observados. Por ejemplo, aunque un valor de $p = .07$ no resulte estadísticamente significativo, se debe valorar de distinta manera si el tamaño del efecto que tiene asociado es de .60, que si es de .10 (p. ej., una medida potencialmente importante clínicamente que no es significativa estadísticamente dado el tamaño de la muestra). Al contrario, si un valor de $p = .04$ resulta estadísticamente significativo, pero tiene asociado un tamaño del efecto de .10, el efecto puede ser demasiado pequeño para ser relevante en la práctica.

Las sugerencias de Cohen han de entenderse como directrices generales, y, en función de la situación, la importancia asociada a un tamaño del efecto determinado podrá ser muy distinta a la sugerida por Cohen. En ciertas situaciones clínicas, un tamaño del efecto pequeño podría representar un importante hallazgo. Se anima al profesional a evaluar los detalles de un determinado resultado cuando interprete la significación y el tamaño del efecto. Véase Cohen (1988) para más información sobre la interpretación de los tamaños del efecto.

Fiabilidad y errores de medida

La fiabilidad de un test se refiere a la exactitud, consistencia y estabilidad en diferentes situaciones de las puntuaciones (Anastasi y Urbina, 1997; Sattler, 2008a). La teoría clásica de los tests considera que la puntuación obtenida en un test es una aproximación a la puntuación verdadera hipotética de un niño, es decir, a la puntuación que el niño obtendría si el test fuera totalmente fiable. La diferencia entre la puntuación verdadera hipotética y la puntuación obtenida en la prueba es lo que se denomina *error de medida*. Un test fiable tiene errores de medida relativamente pequeños y proporciona resultados consistentes tanto si se aplica una sola vez como si se aplica en diferentes ocasiones. La fiabilidad de un test siempre debería tenerse en cuenta en la interpretación de las puntuaciones obtenidas por el niño y de las diferencias encontradas entre las puntuaciones obtenidas por ese mismo niño en distintas ocasiones (p. ej., test-retest).

Consistencia interna

Los análisis de consistencia interna se realizaron con la muestra de tipificación española (N=1008). Los coeficientes de fiabilidad se calcularon utilizando el método de las dos mitades. El coeficiente de fiabilidad basado en el método de las dos mitades es la correlación entre las puntuaciones de las dos mitades corregida mediante la fórmula de Spearman-Brown para la prueba completa (Crocker y Algina, 1986; Haertel, 2006; Li, Rosenthal y Rubin, 1996).

Este método no se considera adecuado para valorar la fiabilidad de las pruebas *Claves*, *Búsqueda de símbolos* y *Cancelación*; por ello, para estas pruebas, se usaron los coeficientes de estabilidad test-retest. Estos coeficientes de estabilidad se basaban en las puntuaciones obtenidas en el análisis test-retest, como se describe más adelante en este capítulo. El coeficiente de estabilidad es la correlación entre las puntuaciones de la primera y segunda aplicación, corregido de acuerdo con la variabilidad de la muestra de tipificación (Allen y Yen, 2002; Haertel, 2006; Magnusson, 1967).

Los coeficientes de fiabilidad de consistencia interna para la puntuación compuesta se calcularon con la fórmula recomendada por Guilford (1954), Haertel (2006) y Nunnally y Bernstein (1994). En la tabla 5.1. figuran los coeficientes de fiabilidad de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas, para toda la muestra y para cada grupo de edad. Los coeficientes de fiabilidad promedio se calcularon utilizando la trasformación *z* de Fisher (Silver y Dunlap, 1987; Strube, 1988).

Tabla 5.1. Coeficientes de fiabilidad de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas (datos españoles)

Prueba/ punt. de procesamiento/ punt. compuesta	Edad											Total promedio
	6	7	8	9	10	11	12	13	14	15	16	r_{xx} [a]
Semejanzas	.88	.89	.81	.90	.87	.88	.84	.85	.80	.79	.78	.85
Vocabulario	.70	.88	.77	.86	.80	.82	.81	.84	.77	.75	.80	.81
Información	.78	.73	.76	.78	.82	.87	.85	.77	.87	.84	.85	.82
Comprensión	.77	.76	.67	.80	.73	.79	.74	.77	.60	.71	.75	.74
Cubos	.87	.79	.81	.85	.87	.78	.72	.77	.85	.81	.81	.82
Puzles visuales	.83	.88	.92	.86	.89	.89	.86	.84	.82	.88	.90	.87
Matrices	.91	.89	.87	.92	.82	.88	.83	.84	.82	.87	.80	.86
Balanzas	.93	.93	.93	.95	.95	.94	.94	.87	.93	.89	.92	.93
Aritmética	.84	.85	.87	.84	.86	.92	.93	.85	.83	.89	.88	.87
Dígitos	.90	.91	.90	.91	.91	.91	.90	.91	.89	.92	.88	.90
Span de dibujos	.78	.86	.90	.80	.83	.80	.88	.83	.83	.89	.86	.85
Letras y números	.92	.92	.85	.78	.82	.84	.81	.74	.74	.87	.84	.84
Claves	.78	.78	.79	.79	.81	.81	.82	.82	.86	.86	.86	.82
Búsqueda de símbolos	.83	.83	.80	.80	.79	.79	.67	.67	.87	.87	.87	.81
Cancelación	.80	.80	.83	.83	.84	.84	.81	.81	.81	.81	.81	.82
Cs	.87	.79	.81	.84	.87	.75	.72	.76	.83	.78	.80	.81
Cp	.86	.92	.94	.89	.94	.84	.81	.76	.79	.84	.78	.87
Dd	.67	.64	.76	.69	.80	.79	.80	.78	.68	.89	.73	.76
Di	.87	.80	.88	.79	.76	.83	.78	.81	.78	.78	.82	.81
Dc	.75	.86	.85	.86	.89	.85	.82	.82	.83	.86	.77	.84
CAa	.80	.80	.82	.82	.81	.81	.80	.80	.81	.81	.81	.81
CAe	.80	.80	.82	.82	.82	.82	.82	.82	.82	.82	.82	.82
ICV	.86	.93	.87	.93	.89	.91	.90	.91	.87	.84	.86	.89
IVE	.90	.89	.92	.91	.93	.90	.87	.88	.89	.90	.91	.90
IRF	.95	.94	.92	.95	.92	.94	.92	.90	.92	.92	.91	.93
IMT	.88	.92	.93	.90	.91	.90	.93	.91	.91	.93	.92	.91
IVP	.88	.88	.86	.86	.88	.87	.85	.84	.91	.92	.90	.88
CIT	.95	.96	.95	.96	.96	.96	.95	.95	.95	.95	.96	.95
IRC	.91	.92	.93	.91	.94	.96	.96	.90	.92	.92	.94	.93
IMTA	.94	.95	.93	.90	.92	.93	.91	.90	.89	.94	.91	.92
INV	.94	.95	.95	.95	.95	.95	.95	.94	.95	.95	.96	.95
ICG	.94	.95	.94	.96	.95	.95	.94	.94	.94	.93	.94	.94
ICC	.91	.93	.92	.92	.92	.92	.92	.91	.93	.94	.93	.92

[a] Los coeficientes de fiabilidad promedio se calcularon con la transformación *z* de Fisher.

Como muestran los datos de la tabla 5.1 los coeficientes de fiabilidad promedio de las pruebas, para la muestra de tipificación española, se sitúan entre adecuados y excelentes (véase Prieto y Muñiz [2000] para valorar los coeficientes de consistencia interna). En general, la fiabilidad de las pruebas de la WISC-V conservan la fiabilidad de las pruebas de la WISC-IV. Los coeficientes de fiabilidad de *Balanzas* y *Puzles visuales* son superiores a los de *Figuras incompletas*, una prueba de Razonamiento perceptivo eliminada en esta versión y sustituida por las dos pruebas nuevas mencionadas. En cuanto a la prueba nueva *Span de dibujos*, aunque no hay una prueba de memoria de trabajo visual comparable en la WISC-IV, su coeficiente de fiabilidad es similar al del resto de las pruebas de la WISC-IV. Los coeficientes de fiabilidad promedio de las puntuaciones compuestas de los índices primarios de la WISC-V oscilan entre .88 y .93 y generalmente son más altos que los de cada una de las pruebas que los conforman. Esta diferencia se explica porque cada prueba representa solo una pequeña parte de todo el funcionamiento intelectual de un niño, mientras que las puntuaciones compuestas (p. ej., índices primarios) resumen el rendimiento de un niño en un ámbito más amplio de aptitudes.

El coeficiente de fiabilidad promedio del IVP, aun siendo excelente, es ligeramente inferior al del resto de las puntuaciones compuestas dado que se basa en las fiabilidades test-retest, que tienden a ser más bajas que las fiabilidades de las dos mitades. La fiabilidad del ICV de la WISC-V es ligeramente inferior a la del ICV de la WISC-IV porque al ICV de la WISC-V contribuyen dos pruebas mientras que al ICV de la WISC-IV contribuyen tres pruebas. A pesar de estar compuestos solo por dos pruebas, el IVE conserva un nivel de fiabilidad similar al del IRP de la WISC-IV, y el IRF muestra un nivel de fiabilidad ligeramente superior al IRP de la WISC-IV debido a la fiabilidad alta estimada para *Matrices* y *Balanzas*. Los coeficientes de fiabilidad promedio de los índices primarios son similares (ICV e IVE) o ligeramente superiores (IRF, IMT e IVP) a las puntuaciones compuestas correspondientes de la WISC-IV.

El coeficiente de fiabilidad del CIT, pese a estar compuesto por siete pruebas, es excelente y tiene el mismo valor que el de la WISC-IV. Los coeficientes de fiabilidad de las puntuaciones de los índices secundarios también son excelentes. El INV y el ICG muestran una fiabilidad similar a la del CIT.

Los resultados de fiabilidad en poblaciones especiales respaldan la generalizabilidad de la escala. Para la WISC-V se obtuvieron resultados de la consistencia interna de las pruebas en grupos especiales de población española, tanto de niños con alta capacidad intelectual como de niños con discapacidad intelectual. También se presentan los resultados del estudio estadounidense, para los siguientes grupos especiales: alta capacidad intelectual, discapacidad intelectual leve, discapacidad intelectual moderada, capacidad intelectual límite, trastorno específico del aprendizaje-lectura, trastorno específico del aprendizaje-lectura y expresión escrita, trastorno específico del aprendizaje-matemáticas, trastorno por déficit de atención/hiperactividad, comportamiento disruptivo, traumatismo craneoencefálico, trastornos del espectro autista con deterioro del lenguaje y trastornos del espectro autista sin deterioro del lenguaje. En el capítulo 6 de este manual se ofrece la información demográfica detallada para estos grupos especiales, y en el anexo B de este manual consta una completa descripción de los criterios de inclusión para cada grupo.

En las tablas 5.2 y 5.3 se presentan los coeficientes de fiabilidad de consistencia interna de las pruebas y de las puntuaciones de procesamiento para los grupos especiales españoles y estadounidenses, respectivamente. Los coeficientes de fiabilidad se calcularon utilizando la misma metodología descrita para los análisis que se presentan en la tabla 5.1. Para las pruebas de Velocidad de procesamiento (*Claves*, *Búsqueda de símbolos* y *Cancelación*) no se presentan los coeficientes de fiabilidad de los grupos especiales puesto que estos grupos no participaron en el estudio test-retest.

Tabla 5.2. Coeficientes de fiabilidad de las pruebas y de las puntuaciones de procesamiento para los grupos especiales (datos españoles)

Prueba/ punt. de procesamiento	Alta capacidad intelectual		Discapacidad intelectual		Promedio
	N	*r*	*N*	*r*	r_{xx}
Semejanzas	36	.86	20	.97	.93
Vocabulario	36	.92	20	.95	.94
Información	36	.93	20	.93	.93
Comprensión	36	.93	20	.89	.91
Cubos	36	.83	20	.96	.92
Puzles visuales	36	.91	20	.93	.92
Matrices	36	.90	20	.96	.94
Balanzas	36	.95	20	.95	.95
Aritmética	36	.94	20	.95	.95
Dígitos	36	.94	20	.96	.95
Span de dibujos	36	.90	20	.93	.92
Letras y números	36	.82	19	.97	.93
Cs	36	.80	20	.96	.91
Cp	34	.83	13	.92	.88
Dd	36	.86	20	.72	.80
Di	36	.83	20	.96	.92
Dc	36	.89	20	.85	.87

Tabla 5.3. Coeficientes de fiabilidad de las pruebas y de las puntuaciones de procesamiento para los grupos especiales (datos estadounidenses)

Prueba/ puntuación de procesamiento	Alta capacidad intelectual		Discapacidad intelectual leve		Discapacidad intelectual moderada		Capacidad intelectual límite		Trastorno específico del aprendizaje-lectura		Trastorno específico del aprendizaje-lectura y expresión escrita		Trastorno específico del aprendizaje-matemáticas		Trastorno por déficit de atención/ hiperactividad		Comportamiento disruptivo		Traumatismo craneo-encefálico		Trastornos del espectro autista con deterioro del lenguaje		Trastornos del espectro autista sin deterioro del lenguaje		Promedio
	N	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*	r_{xx} [a]
Semejanzas	95	.93	74	.94	37	.97	20	.92	30	.89	22	.95	27	.86	48	.91	21	.96	17	.95	29	.95	31	.90	.93
Vocabulario	95	.95	74	.92	37	.93	19	.95	30	.94	22	.94	28	.94	48	.93	21	.92	19	.97	29	.96	32	.96	.94
Información	95	.90	74	.94	37	.91	20	.95	30	.91	22	.86	27	.94	48	.93	21	.96	20	.95	29	.95	32	.95	.93
Comprensión	95	.91	74	.94	37	.94	20	.91	30	.90	22	.89	28	.92	48	.91	21	.93	20	.91	30	.93	32	.86	.92
Cubos	95	.90	74	.89	37	.87	20	.89	30	.91	22	.87	28	.88	48	.92	21	.88	19	.85	30	.93	32	.89	.89
Puzles visuales	95	.88	74	.94	37	.97	20	.91	30	.93	22	.90	28	.95	48	.94	21	.88	20	.95	29	.96	32	.90	.93
Matrices	95	.88	74	.95	37	.94	20	.83	30	.83	22	.67	28	.96	48	.92	21	.92	20	.91	30	.96	32	.91	.91
Balanzas	95	.95	74	.95	36	.97	20	.97	30	.95	22	.95	28	.97	48	.97	21	.98	20	.96	30	.98	32	.93	.97
Aritmética	95	.94	74	.97	37	.97	20	.99	30	.88	22	.95	28	.94	48	.95	21	.96	20	.97	30	.97	32	.95	.96
Dígitos	95	.94	73	.95	34	.95	20	.98	30	.83	22	.93	28	.93	47	.94	21	.96	20	.96	30	.99	32	.97	.96
Span de dibujos	95	.89	74	.92	36	.91	19	.91	30	.86	22	.86	28	.87	48	.89	21	.94	20	.89	30	.94	32	.90	.90
Letras y números	93	.86	73	.95	33	.93	20	.92	30	.81	22	.95	27	.78	47	.94	21	.98	20	.96	29	.93	32	.93	.93
Cs	95	.90	74	.89	37	.87	20	.89	30	.91	22	.85	28	.87	48	.92	21	.87	17	.81	30	.94	32	.89	.89
Cp	95	.90	74	.93	36	.96	20	.96	30	.92	22	.93	28	.79	47	.91	21	.94	17	.91	30	.97	32	.89	.93
Dd	95	.88	74	.87	37	.84	20	.88	30	.54	22	.86	28	.90	48	.82	21	.82	20	.90	30	.89	32	.88	.86
Di	95	.86	74	.91	37	.94	20	.93	30	.48	22	.84	28	.84	48	.82	21	.96	20	.92	30	.91	32	.79	.88
Dc	95	.82	73	.89	34	.95	20	.93	30	.81	22	.89	28	.89	47	.84	21	.87	20	.97	30	.97	32	.88	.91

Nota: La tabla no incluye las puntuaciones de procesamiento de *Claves, Búsqueda de símbolos* y *Cancelación* porque no se recogieron datos del test-retest para los grupos especiales.

[a] Los coeficientes de fiabilidad promedio se calcularon con la transformación *z* de Fisher.

Col·legi Oficial de Psicologia de Catalunya
Biblioteca

Errores típicos de medida e intervalos de confianza

El error típico de medida (ETM) proporciona una estimación de la cantidad de error que contiene una puntuación obtenida por un niño en un test. Dado que el *ETM* es inversamente proporcional a la fiabilidad, cuanto mayor es la fiabilidad del instrumento, menor es el *ETM*, lo que aumenta la precisión de los resultados obtenidos en el test. El *ETM* normalmente se expresa en términos de unidades de puntuación típica; es decir, el *ETM* es la desviación típica de la distribución del error de medida, y se calcula a partir de la siguiente fórmula:

$$ETM = Dt\sqrt{1 - r_{xx}}$$

donde *Dt* es la unidad de desviación típica del test en la escala de medida empleada y r_{xx} es el coeficiente de fiabilidad que tiene esa medida. No se deben comparar los *ETM* de las puntuaciones de las pruebas con los de las puntuaciones compuestas porque se basan en diferentes desviaciones típicas. Dado que la desviación típica de las puntuaciones de las pruebas es 3 y la de las puntuaciones compuestas es 15, los *ETM* de las puntuaciones de las pruebas suelen ser menores que los de las puntuaciones compuestas. En la tabla 5.4. figuran los *ETM* de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas para cada grupo de edad, así como el promedio de todas las edades.

Tabla 5.4. Errores típicos de medida de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas (datos españoles)

Prueba/ punt. de procesamiento/ punt. compuesta	Edad											Total promedio
	6	7	8	9	10	11	12	13	14	15	16	*ETM*
Semejanzas	1.04	0.99	1.31	0.95	1.08	1.04	1.20	1.16	1.34	1.37	1.41	1.18
Vocabulario	1.64	1.04	1.44	1.12	1.34	1.27	1.31	1.20	1.44	1.50	1.34	1.34
Información	1.41	1.56	1.47	1.41	1.27	1.08	1.16	1.44	1.08	1.20	1.16	1.30
Comprensión	1.44	1.47	1.72	1.34	1.56	1.37	1.53	1.44	1.90	1.62	1.50	1.54
Cubos	1.08	1.37	1.31	1.16	1.08	1.41	1.59	1.44	1.16	1.31	1.31	1.30
Puzles visuales	1.24	1.04	0.85	1.12	0.99	0.99	1.12	1.20	1.27	1.04	0.95	1.08
Matrices	0.90	0.99	1.08	0.85	1.27	1.04	1.24	1.20	1.27	1.08	1.34	1.13
Balanzas	0.79	0.79	0.79	0.67	0.67	0.73	0.73	1.08	0.79	0.99	0.85	0.82
Aritmética	1.20	1.16	1.08	1.20	1.12	0.85	0.79	1.16	1.24	0.99	1.04	1.08
Dígitos	0.95	0.90	0.95	0.90	0.90	0.90	0.95	0.90	0.99	0.85	1.04	0.93
Span de dibujos	1.41	1.12	0.95	1.34	1.24	1.34	1.04	1.24	1.24	0.99	1.12	1.19
Letras y números	0.85	0.85	1.16	1.41	1.27	1.20	1.31	1.53	1.53	1.08	1.20	1.24
Claves	1.41	1.41	1.37	1.37	1.31	1.31	1.27	1.27	1.12	1.12	1.12	1.28
Búsqueda de símbolos	1.24	1.24	1.34	1.34	1.37	1.37	1.72	1.72	1.08	1.08	1.08	1.34
Cancelación	1.34	1.34	1.24	1.24	1.20	1.20	1.31	1.31	1.31	1.31	1.31	1.28
Cs	1.08	1.37	1.31	1.20	1.08	1.50	1.59	1.47	1.24	1.41	1.34	1.34
Cp	1.12	0.85	0.73	0.99	0.73	1.20	1.31	1.47	1.37	1.20	1.41	1.15
Dd	1.72	1.80	1.47	1.67	1.34	1.37	1.34	1.41	1.70	0.99	1.56	1.50
Di	1.08	1.34	1.04	1.37	1.47	1.24	1.41	1.31	1.41	1.41	1.27	1.31
Dc	1.50	1.12	1.16	1.12	0.99	1.16	1.27	1.27	1.24	1.12	1.44	1.23
CAa	1.34	1.34	1.27	1.27	1.31	1.31	1.34	1.34	1.31	1.31	1.31	1.31
CAe	1.34	1.34	1.27	1.27	1.27	1.27	1.27	1.27	1.27	1.27	1.27	1.28
ICV	5.61	3.97	5.41	3.97	4.97	4.50	4.74	4.50	5.41	6.00	5.61	5.02
IVE	4.74	4.97	4.24	4.50	3.97	4.74	5.41	5.20	4.97	4.74	4.50	4.74
IRF	3.35	3.67	4.24	3.35	4.24	3.67	4.24	4.74	4.24	4.24	4.50	4.07
IMT	5.20	4.24	3.97	4.74	4.50	4.74	3.97	4.50	4.50	3.97	4.24	4.43
IVP	5.20	5.20	5.61	5.61	5.20	5.41	5.81	6.00	4.50	4.24	4.74	5.26
CIT	3.35	3.00	3.35	3.00	3.00	3.00	3.35	3.35	3.35	3.35	3.00	3.20
IRC	4.50	4.24	3.97	4.50	3.67	3.00	3.00	4.74	4.24	4.24	3.67	4.02
IMTA	3.67	3.35	3.97	4.74	4.24	3.97	4.50	4.74	4.97	3.67	4.50	4.24
INV	3.67	3.35	3.35	3.35	3.35	3.35	3.35	3.67	3.35	3.35	3.00	3.38
ICG	3.67	3.35	3.67	3.00	3.35	3.35	3.67	3.67	3.67	3.97	3.67	3.56
ICC	4.50	3.97	4.24	4.24	4.24	4.24	4.24	4.50	3.97	3.67	3.97	4.17

Nota: Los *ETM* se presentan en métrica de puntuaciones escalares para las pruebas y las puntuaciones de procesamiento, y en métrica de puntuaciones compuestas para los índices y el CI total. Para calcular estos *ETM*, se utilizaron los coeficientes de fiabilidad que figuran en la tabla 5.1 y las desviaciones típicas (a saber, 3 para las puntuaciones escalares y 15 para las puntuaciones compuestas).

[a] Los ETM promedios se calcularon promediando la suma de los ETM elevados al cuadrado en cada grupo de edad, y obteniendo la raíz cuadrada del resultado.

El *ETM* permite estimar los intervalos de confianza en los que se espera encontrar las verdaderas puntuaciones del niño. Los intervalos de confianza nos indican que el error de medida es inherente a la medición, y que la puntuación observada de un test es solo una estimación de la puntuación verdadera, es decir, de la capacidad real del niño. Los intervalos de confianza pueden usarse para presentar una puntuación dentro de un rango en el que probablemente se sitúa la puntuación verdadera del niño.

Para calcular los intervalos de confianza, pueden utilizarse distintos métodos, y cada uno de ellos produce valores ligeramente diferentes. En todas las fórmulas para calcular los intervalos de confianza se utiliza el *ETM*. En algunas ocasiones, los profesionales prefieren calcular los intervalos de confianza de las puntuaciones compuestas utilizando solo la puntuación observada y el *ETM*. Los intervalos de confianza basados en la puntuación observada y el *ETM* se calculan mediante la siguiente fórmula:

$$p\% \text{ intervalo de confianza} = \text{puntuación observada} \pm z_p(ETM)$$

donde p es el nivel de confianza, como 90 o 95 %, y donde z_p es el valor z asociado al nivel de confianza, que puede obtenerse en una tabla de áreas bajo la curva normal.

Para calcular los intervalos de confianza centrados en las puntuaciones observadas, el profesional, en función de su preferencia, puede usar el *ETM* para el grupo de edad del niño o para toda la muestra. En los ejemplos de este apartado se han utilizado los *ETM* basados en el grupo de edad del niño. Por ejemplo, si un niño de 6 años obtiene un CI total de 108, el profesional puede estar seguro en un 95 % de que el verdadero CI total del niño se sitúa en el rango de 101-115 (porque el intervalo de confianza del 95 % es 108 ± 1.96 *ETM*, donde el *ETM* es 3.35), y en un 90 % de que el CI total del sujeto se encuentra en el rango de 102-114 (108 ± 1.65 *ETM*).

Los intervalos de confianza de los rangos de puntuaciones compuestas que figuran en las tablas A.2-A.7 y C.1-C.5 del *Manual de aplicación y corrección* se obtuvieron mediante las puntuaciones verdaderas estimadas (*PVE*) y el error típico de estimación (*ETE*), de acuerdo con el método propuesto por Dudek (1979) y Glutting, McDermott y Stanley (1987). La puntuación verdadera estimada se obtiene mediante la fórmula:

$$\text{PVE} = 100 + r_{xx}(X - 100)$$

donde X es la puntuación compuesta observada, y r_{xx} es el coeficiente de fiabilidad de la puntuación compuesta. El *ETE* se deriva con la fórmula propuesta por Stanley (1971):

$$ETE = Dt(r_{xx})\sqrt{1 - r_{xx}}$$

donde *Dt* es la desviación típica teórica de la puntuación compuesta, y r_{xx} es el coeficiente de fiabilidad de la puntuación compuesta.

Este método centra el intervalo de confianza en la puntuación verdadera estimada, en lugar de hacerlo en la puntuación observada, lo que da lugar a un intervalo de confianza asimétrico respecto a la puntuación observada. Esta asimetría se produce porque la puntuación verdadera estimada está más cerca de la media de la escala que la puntuación observada. Un intervalo de confianza basado en la puntuación verdadera estimada y en el *ETE* es una corrección de la regresión de la puntuación verdadera respecto a la media. Por ejemplo, si un niño de 6 años obtiene un CI total de 113, su puntuación verdadera estimada será de 112.4, el intervalo de confianza del 95 % del CI total verdadero del niño se situará en el rango de 106-119 (porque el intervalo de confianza del 95 % es 112.4 ± 1.96 *ETE*, donde el *ETE* es 3.19), y el intervalo de confianza del 90 % del CI total verdadero del niño se encontrará en el rango de 107-118 (112.4 ± 1.65 *ETE*).

Los intervalos de confianza que figuran en las tablas A.2-A.7 y C.1-C.5 del *Manual de aplicación y corrección* se calcularon utilizando los promedios del coeficiente de fiabilidad para toda la muestra, en vez de para cada grupo de edad, porque los coeficientes de fiabilidad de las puntuaciones

compuestas y los *ETE* son muy similares entre los grupos de edad. Del mismo modo, dado que los *ETE* de los grupos de edad no son muy diferentes, los intervalos de confianza centrados en la puntuación obtenida y calculados utilizando el promedio para toda la muestra o el valor basado en la edad serían muy similares. Asimismo, dado que los coeficientes de fiabilidad de las puntuaciones compuestas son relativamente elevados, los intervalos de confianza calculados con el *ETM* centrado en la puntuación obtenida y los intervalos de confianza calculados con el *ETE* centrado en la puntuación verdadera estimada serán similares. Cuando se observan diferencias, estas se dan en los extremos más alto y más bajo de la distribución, dado que la mayor regresión a la media afecta a los extremos de la distribución de puntuaciones. En algunos casos, los profesionales pueden querer calcular los intervalos de confianza centrados en la puntuación observada, usar niveles de confianza distintos a los indicados en este manual o utilizar las fiabilidades del grupo de edad del niño. Los intervalos de confianza calculados usando el *ETM*, el *ETE* o los valores del grupo de edad se interpretan del mismo modo.

Estabilidad temporal: test-retest

La prueba de estabilidad test-retest para las pruebas, las puntuaciones de procesamiento y las puntuaciones compuestas se obtuvo a partir de una muestra española de 67 niños, de 6 a 15 años. A los participantes se les aplicó la WISC-V en dos ocasiones, con un intervalo temporal de entre 8 y 97 días y un intervalo medio de 33 días. La tabla 5.5. presenta los datos demográficos del estudio test-retest y del estudio de fiabilidad por acuerdo entre correctores, que se describe más adelante. En ella figuran la edad media y los porcentajes de representación de la muestra por sexo, nivel educativo, zona geográfica y tipo de población.

Tabla 5.5. Datos demográficos del estudio de estabilidad test-retest y del estudio de acuerdo entre correctores (datos españoles)

	Estabilidad test-retest	Acuerdo entre correctores
N	67	150
Edad		
Media	10.7	11.2
Dt	3.4	3.2
Sexo		
Mujer	43.3	54.7
Varón	56.7	45.3
Nivel educativo de los padres		
Primer grado	9.0	8.0
Segundo grado	64.1	70.0
Tercer grado	26.9	22.0
Zona geográfica		
Norte	13.4	12.7
Sur	44.8	34.0
Este	19.4	31.3
Centro	22.4	22.0
Tipo de población		
Urbana	53.7	54.0
Intermedia	22.4	22.7
Rural	23.9	23.3

La fiabilidad test-retest se estimó para todas las edades empleando la correlación de Pearson. La media y la desviación típica de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas se presentan en la tabla 5.6. Los coeficientes de estabilidad test-retest promedio para todas las edades se calcularon utilizando la transformación *z* de Fisher. Las tablas también recogen las diferencias típicas (tamaño del efecto) entre la primera y la segunda aplicación y los coeficientes de correlación corregidos de acuerdo con la variabilidad de la muestra de tipificación española (Allen y Yen, 2002; Magnusson, 1967). La diferencia típica se calculó usando la diferencia de la puntuación media obtenida en las dos aplicaciones dividida por la desviación típica común (Cohen, 1988).

Tabla 5.6. Coeficientes de estabilidad de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas (datos españoles)

Todas las edades	Primera aplicación		Segunda aplicación					
Prueba/ punt. de procesamiento/ punt. compuesta	Media	*Dt*	Media	*Dt*	*n*	r_{12}[a]	*r* corregido[b]	Diferencia típica[c]
Semejanzas	10.1	2.6	11.1	2.5	66	.58	.69	0.39
Vocabulario	10.4	2.6	11.0	2.4	67	.76	.82	0.24
Información	10.3	2.4	11.3	2.5	67	.65	.78	0.41
Comprensión	11.6	3.0	12.2	3.2	62	.70	.70	0.18
Cubos	10.4	3.1	11.1	3.1	67	.79	.78	0.23
Puzles visuales	9.3	3.1	10.9	3.1	67	.76	.74	0.52
Matrices	9.8	3.1	10.8	3.1	67	.79	.78	0.32
Balanzas	9.9	2.9	10.7	2.8	67	.65	.67	0.28
Aritmética	10.3	2.9	11.1	3.0	67	.72	.74	0.27
Dígitos	10.6	2.8	11.6	2.8	67	.67	.71	0.36
Span de dibujos	10.3	2.6	11.5	3.1	67	.61	.71	0.42
Letras y números	10.7	2.5	11.6	3.0	66	.73	.81	0.33
Claves	10.0	3.1	11.7	3.1	67	.79	.78	0.55
Búsqueda de símbolos	10.4	2.8	12.4	3.1	67	.77	.80	0.68
Cancelación	10.1	3.2	11.0	3.5	67	.86	.84	0.27
Cs	10.4	3.2	11.0	3.1	67	.77	.74	0.19
Cp	11.1	2.3	11.3	2.7	29	.68	.81	0.08
Dd	10.5	2.8	11.3	2.9	67	.63	.68	0.28
Di	10.1	2.6	11.3	2.8	66	.59	.69	0.44
Dc	10.7	2.8	11.2	2.5	65	.62	.67	0.19
CAa	9.9	3.1	11.0	3.2	67	.79	.78	0.35
CAe	10.2	3.2	10.9	3.4	67	.74	.71	0.21
ICV	101.2	11.5	105.5	11.4	67	.75	.85	0.38
IVE	99.1	15.5	105.6	15.9	67	.87	.87	0.41
IRF	99.1	15.3	104.4	14.5	67	.75	.74	0.36
IMT	102.6	13.4	108.6	14.1	67	.75	.80	0.44
IVP	101.2	14.6	111.6	16.4	67	.82	.83	0.67
CIT	101.0	12.9	108.2	12.1	67	.85	.89	0.58
IRC	100.5	14.4	105.2	14.2	67	.77	.79	0.33
IMTA	103.7	13.1	108.8	14.6	66	.78	.84	0.37
INV	99.4	14.2	108.1	14.3	67	.87	.88	0.61
ICG	100.6	13.1	106.2	11.9	67	.85	.89	0.45
ICC	102.2	13.6	112.1	15.3	67	.86	.89	0.68

[a] Los coeficientes de estabilidad promedio entre todas las edades se calcularon con la transformación *z* de Fisher.

[b] Se corrigieron las correlaciones de acuerdo con la variabilidad de la muestra de tipificación española (Allen y Yen, 2002; Magnusson, 1967).

[c] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Como se desprende de los datos expuestos en la tabla 5.6, las puntuaciones de la WISC-V muestran, por lo general, una adecuada estabilidad temporal (véase Prieto y Muñiz [2000] para valorar los coeficientes de estabilidad temporal). Los coeficientes de estabilidad promedio de las pruebas, para la muestra de tipificación española son excelentes para *Cancelación, Vocabulario, Búsqueda de símbolos* y *Letras y números*; buenos para *Información, Cubos, Matrices* y *Claves*; y adecuados para *Semejanzas, Comprensión Puzles visuales, Balanzas, Aritmética, Dígitos y Span de dibujos.*

Los coeficientes de estabilidad de las puntuaciones compuestas de los índices primarios son excelentes, excepto para el IRF, cuyo coeficiente de fiabilidad es adecuado, posiblemente porque en las dos pruebas que contribuyen a este índice puede darse una cierta tasa de aciertos por azar. El coeficiente de estabilidad del CIT es excelente. Los coeficientes de fiabilidad de las puntuaciones de los índices secundarios también son excelentes (entre .84 y .89), excepto para el IRC, cuyo coeficiente de fiabilidad es bueno (.79).

Tal y como indican los datos de la tabla 5.6, las puntuaciones medias de todas las pruebas fueron superiores en la segunda aplicación (retest) que las obtenidas en la primera aplicación (test). En general, este incremento de puntuación es menor en las pruebas de Comprensión verbal y de Memoria de trabajo que en las de Visoespacial, Razonamiento fluido y Velocidad de procesamiento. Estos resultados son consistentes con los obtenidos en otras escalas de inteligencia de Wechsler (Wechsler, 2002, 2003, 2008, 2012).

Acuerdo entre correctores

Durante la fase de tipificación española, una muestra de 150 protocolos de la WISC-V se sometieron a una doble corrección por parte de dos correctores independientes. Posteriormente se comprobó el grado de acuerdo entre ellos. Dado que la mayoría de los criterios de corrección de las pruebas eran sencillos y objetivos, el acuerdo entre correctores fue muy alto, de .98 a .99.

En la tabla 5.5. se recogen las características demográficas de los niños de la muestra a partir de la cual se valoró el acuerdo entre correctores, 150 niños con edades de 6 a 16 años. En ella se indican la edad media y los porcentajes de representación por sexo, nivel educativo de los padres, zona geográfica y tipo de población.

Los 150 protocolos fueron corregidos por dos correctores con un adecuado nivel de cualificación profesional. Ninguno de los correctores tenía experiencia previa con el sistema de corrección de la WISC-V.

Los coeficientes de fiabilidad entre correctores se calcularon de acuerdo con los métodos de correlación intraclase (McGraw y Wong, 1996; Shrout y Fleiss, 1979). En este análisis se utilizaron las puntuaciones totales basadas en ítems corregidos subjetivamente. Los coeficientes de fiabilidad entre correctores fueron los siguientes: .98 para *Semejanzas*, .98 para *Vocabulario*, .99 para *Información* y .98 para *Comprensión*. Estos resultados muestran que, aunque la corrección de estas pruebas requiera una valoración más subjetiva, pueden ser puntuadas de modo fiable, incluso por correctores sin experiencia previa con los criterios de corrección de la WISC-V.

Resumen

En este capítulo se han presentado los estudios de fiabilidad realizados sobre la WISC-V. La consistencia interna 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. Todo ello respalda sólidamente la precisión de las puntuaciones de la WISC-V, una condición previa para que la interpretación sea válida.

Capítulo 6

Validez

La validez es el aspecto más importante desde el punto de vista de la construcción de un test (AERA, APA y NCME, 1999; Sattler, 2008a). Tradicionalmente, los investigadores y constructores de tests se han referido a tres principales tipos de validez: validez de contenido, validez en relación con un criterio y validez de constructo. Así, se consideraba que un test tenía validez de contenido si mostraba adecuadamente los aspectos relevantes del constructo que se estaba midiendo. Un test poseía validez en relación con un criterio si las puntuaciones mostraban relación con un criterio específico externo, como el rendimiento en alguna otra medida o la pertenencia a un grupo. Por último, la validez de constructo de un test se conseguía si realmente se medía el significado del constructo que se pretendía medir con el test. La evidencia de validez del constructo puede proceder de diferentes fuentes, como el análisis factorial, la revisión por parte de expertos, los estudios multirrasgo-multimétodo y las investigaciones clínicas.

Las definiciones actuales de validez son de carácter más unitario. Como se afirma en *Standards for educational and psychological testing*, «en la evolución del concepto de [...] validez no se habla ya de diferentes tipos de validez, sino de diferentes fuentes de evidencia de validez, todas ellas encaminadas a proporcionar una información relevante para lograr una interpretación específica de las puntuaciones del test» (AERA, APA y NCME, 1999, p. 5). La validez, por tanto, se refiere al grado en el que existe evidencia que apoye la interpretación de las puntuaciones del test con un fin determinado. Por consiguiente, en el examen de la validez de un test debe intervenir el juicio evaluativo del propio usuario del test.

Aunque los constructores de los tests son responsables de proporcionar evidencias iniciales de validez, el profesional es quien debe evaluar si estas evidencias justifican el uso que el profesional desea realizar del test con un fin determinado. El examen exhaustivo de los datos sobre la validez de una escala ha de incluir el de los trabajos de investigación más importantes acerca de las versiones anteriores de la escala, así como el análisis de los estudios que aplican la última versión de la escala en diferentes contextos, en poblaciones distintas y con propósitos diversos. De esta forma, se espera que el uso de la WISC-V proporcione en el futuro más información sobre la validez de la escala. Si se desea más información sobre los datos que apoyan la validez de la predecesora de la WISC-V, la WISC-IV, véanse Flanagan y Harrison (2012), Flanagan y Kaufman (2009), Prifitera, Saklofske y Weiss (2008) y Wechsler (2003).

Este capítulo presenta diferentes fuentes de evidencias de la validez de la WISC-V que apoyan el uso de la escala como medida de la aptitud intelectual, incluidos los datos sobre el contenido, los procesos de respuesta, la estructura interna de la escala, su relación con otras variables y las consecuencias de la aplicación de tests.

Evidencia basada en el contenido

El estudio de la relación entre el contenido de un test y el constructo que se pretende medir constituye una fuente de información acerca de la validez del test. Los datos que prueban la validez de contenido de un test no se basan generalmente en análisis estadísticos o en pruebas empíricas, sino que se fundamentan en el grado en que los ítems del test se relacionan y representan adecuadamente el rasgo o función que se pretende medir. El contenido de un test también implica la redacción y el formato de los ítems, así como los procedimientos para la aplicación y corrección del test.

El objetivo de las revisiones del contenido de la WISC-V fue asegurar que todos los ítems y pruebas abordaran adecuadamente los distintos dominios del funcionamiento intelectual que la escala pretende evaluar. Los ítems y las pruebas cubren un amplio rango de procesos cognitivos, incluidos el razonamiento verbal y analógico, la formación de conceptos, el procesamiento secuencial y simultáneo, el razonamiento visoespacial, la comprensión auditiva, el razonamiento cuantitativo, la flexibilidad cognitiva, la memoria de trabajo, la construcción y la velocidad de procesamiento psicomotora (Flanagan, Alfonso y Ortiz, 2012; Flanagan y Kaufman, 2009; Sattler, 2008a). Para mejorar la amplitud y la relevancia del contenido, se llevaron a cabo revisiones exhaustivas de la literatura y se consultó con expertos para examinar el contenido de la WISC-V y evaluar así la pertinencia de incorporar nuevos ítems y pruebas. Los miembros del equipo de expertos y asesores (que figuran en el anexo C de este manual) se seleccionaron atendiendo a su experiencia en psicología clínica, neuropsicología y/o trastornos del aprendizaje. El capítulo 2 de este manual describe con detalle el contenido y la estructura de la WISC-V.

Evidencia basada en los procesos de respuesta

En la WISC-V, la evidencia de validez basada en los procesos de respuesta debería confirmar que el niño usa el proceso cognitivo esperado al realizar las tareas de una prueba determinada. Este tipo de evidencia puede proceder de fuentes teóricas o de análisis psicométricos.

La mayoría de las pruebas de la WISC-V cuentan con una fuerte base teórica y empírica de validez, basada en los procesos de respuesta. Así, las pruebas que se han conservado de versiones anteriores cuentan con el apoyo de numerosas investigaciones sobre los procesos de respuesta que intervienen en la ejecución de una tarea (Carroll, 1993; Flanagan y Kaufman, 2009; Sattler, 2008a). A fin de aportar datos sobre la validez de los procesos de respuesta que participan en las pruebas nuevas, se revisó la literatura científica, se consultó con expertos y se realizaron exámenes empíricos. Los detalles sobre las bases teóricas de todas las pruebas, así como la información acerca de los procesos de respuesta relacionados, se ofrecen en el capítulo 2 de este manual.

Los análisis empíricos y cualitativos que se realizaron durante todo el proceso de desarrollo de la escala permitieron obtener evidencias sobre la validez de los procesos de respuesta. Por ejemplo, se examinaron las frecuencias de respuesta a los ítems con respuesta múltiple, para identificar qué respuestas incorrectas se daban con más frecuencia. Las respuestas incorrectas más frecuentes se examinaron para determinar la posibilidad de considerarlas como respuestas aceptables o posibles, y se pidió a los niños que explicaran las razones que los habían llevado a dar tales respuestas. Además, para examinar las consignas e ítems nuevos de *Matrices*, se pidió a los niños que explicaran las estrategias de solución de problemas que habían seguido para resolver dichos ítems; y, en función de sus respuestas, se modificaron las consignas o el contenido de los ítems (estímulos, distractores) o se eliminó el ítem.

Las preguntas directas al niño también son útiles para poner de manifiesto diferencias en los procesos de respuesta, en especial en las pruebas nuevas. Así, al aplicar los ítems de *Puzles visuales*, se les pidió a los niños que explicaran cómo entendían los puzles. Sus respuestas llevaron a ampliar la consigna del ítem de ejemplo para aclarar que en un puzle no se puede superponer las piezas. Para la prueba *Balanzas* también se obtuvo información sobre cómo los niños entendían el funcionamiento de este utensilio de medición. En atención a las respuestas se añadieron ítems de ejemplo para garantizar que los niños entendieran cómo se equilibra una balanza. Asimismo se les pidió si entendían que las figuras de distintos colores y formas podían tener pesos diferentes; la información obtenida de los niños dio lugar a la modificación de la consigna del ítem de ejemplo para aclarar este punto. Atendiendo a las respuestas de los niños, se observó que en *Span de dibujos* era necesario añadir dos ítems de ejemplo para explicar el proceso de reordenamiento de los dibujos y cómo seleccionar ordenadamente la(s) respuesta(s) de entre las opciones de respuesta. Al igual que sucede con otras evidencias de validez, se espera que las futuras investigaciones ofrezcan más datos sobre los procesos de respuesta evocados por los ítems y las pruebas de la WISC-V.

Evidencia basada en la estructura interna

El estudio de la estructura interna de un test «puede indicar el grado en el que las relaciones entre los ítems del test y sus componentes se ajustan al constructo en el que se basa la interpretación de la puntuación del test» (AERA, APA y NCME, 1999, p. 13). Los estudios de investigación que han utilizado las escalas de inteligencia de Wechsler han aportado datos concluyentes sobre su validez respecto a la estructura interna (Wechsler, 2003, 2008, 2012).

El tipo de evidencias aportadas ha evolucionado paralelamente a los avances experimentados en la teoría y en la medición de la inteligencia, así como en los campos de la psicometría, la neuropsicología y el desarrollo cognitivo. Las primeras versiones de las escalas de Wechsler ofrecían dos índices: CI verbal (CIV) y CI manipulativo (CIM). En una serie de artículos (1957a, 1957b, 1959), Cohen presentó los resultados del análisis factorial de la escala Wechsler-Bellevue, la *Wechsler Adult Intelligence Scale* (WAIS; Wechsler, 1955) y la WISC, y sugirió la presencia de un tercer factor débil (memoria de trabajo) dentro de la escala manipulativa. Así, para confirmar la existencia de este tercer factor, se incluyó la prueba *Búsqueda de símbolos* en la WISC-III, pero se puso de manifiesto la presencia de un cuarto factor: la velocidad de procesamiento. Del mismo modo, la inclusión de las pruebas *Letras y números* y *Búsqueda de símbolos* en la *Escala de inteligencia de Wechsler para adultos-III* (WAIS-III; Wechsler, 1997, 2010) también condujo a cuatro factores. Consecuentemente, se desarrollaron cuatro índices para la WISC-III y cuatro índices para la WAIS-III con el objetivo de representar aptitudes cognitivas específicas: Comprensión verbal, Organización perceptiva, Independencia de la distracción (o Memoria de trabajo en la WAIS-III) y Velocidad de proceso; manteniendo al mismo tiempo las puntuaciones tradicionales para el CI verbal, el CI manipulativo y el CI total. Conforme los nuevos índices ganaban aceptación y un mayor apoyo empírico, la utilidad clínica de los mismos empezó a ser incuestionable. Con la incorporación de *Claves* y *Búsqueda de símbolos* en la *Escala de inteligencia de Wechsler para preescolar y primaria-III* (WPPSI-III, Wechsler, 2002, 2009b) para los niños de 4:0 a 7:3 años, se puso de manifiesto la presencia de un factor de velocidad de procesamiento junto con los factores verbal y manipulativo, de lo que resultó un modelo de tres factores para las edades mencionadas.

La WISC-IV y la WAIS-IV añadieron otras medidas de razonamiento fluido, memoria de trabajo y velocidad de procesamiento. Para ambas escalas se mantuvo la estructura de cuatro factores, se cambió la denominación *Índice de organización perceptiva* por *Índice de razonamiento perceptivo* para reflejar el mayor énfasis del factor en el razonamiento fluido, y se eliminaron el CIV y el CIM. Estas modificaciones son consistentes con la teoría de la inteligencia contemporánea, los estudios de análisis factorial y las investigaciones clínicas. Además, para la WISC-IV, se cambió la denominación *Índice de independencia de la distracción* por *Índice de memoria de trabajo*. Diversos estudios independientes han confirmado la estructura de cuatro factores de la WISC-IV y de la WAIS-IV (p. ej., Bodin, Pardini, Burns y Stevens, 2009; Ward, Bergman y Hebert, 2012; Watkins, 2010; Watkins, Wilson, Kotz, Carbone y Babula, 2006).

Desde la publicación de la WISC-IV y de la WAIS-IV, los investigadores han identificado una estructura factorial alternativa, común a ambas escalas. Los análisis factoriales de varios estudios sugieren la presencia de cinco factores en dichas escalas (Benson et al., 2010; Keith et al., 2006; Weiss et al., 2013a, 2013b). Los resultados de estos estudios muestran que los factores Comprensión verbal y Velocidad de procesamiento son iguales a los de los tests publicados, y que las pruebas de los otros dos factores (Razonamiento perceptivo y Memoria de trabajo) se distribuyen en tres factores, que representan las aptitudes visoespaciales, de razonamiento fluido y de memoria de trabajo. En la WISC-IV, las pruebas *Cubos* y *Figuras incompletas* contribuyen al factor Visoespacial; *Conceptos, Matrices* y *Aritmética* contribuyen al factor Razonamiento fluido; y *Dígitos* y *Letras y números* siguen contribuyendo al factor Memoria de trabajo. En la WAIS-IV, el modelo de cinco factores se pone de manifiesto para las edades de 16 a 69 años cuando todas las pruebas se incluyen en el análisis (incluidas *Balanzas, Letras y números* y *Cancelación*). Por lo general, las pruebas *Cubos, Puzles visuales* y *Figuras incompletas* contribuyen al factor Visoespacial; *Matrices, Balanzas* y *Aritmética* contribuyen al factor Razonamiento fluido; y *Dígitos* y *Letras y números* contribuyen al factor Memoria de trabajo.

Este modelo de cinco factores se refleja para las edades de 4:0 a 7:7 años en la WPPSI-IV (Wechsler, 2012), que incluye pruebas nuevas de memoria de trabajo y pruebas de Velocidad de procesamiento más adecuadas al desarrollo. El factor Visoespacial se mide con las pruebas *Cubos* y *Rompecabezas*; el Razonamiento fluido con *Matrices* y *Conceptos*; la Memoria de trabajo con *Reconocimiento* y *Localización*; y la Velocidad de procesamiento con *Búsqueda de animales, Cancelación* y *Clave de figuras*. Para las edades de 2:6 a 3:11 años, el modelo de dos factores (Verbal y Manipulativo) se amplió a tres factores (Comprensión verbal, Visoespacial y Memoria de trabajo).

La estructura interna de la WISC-V se analizó teniendo en cuenta los datos aportados por los análisis factoriales y la incorporación de tres pruebas nuevas, con el objetivo de que cada una de ellas refuerce la medición de un factor: *Puzles visuales* (prueba opcional de Visospacial), *Balanzas* (pruebas principal de Razonamiento fluido) y *Span de dibujos* (prueba opcional de Memoria de trabajo). Se esperaba que los resultados del análisis factorial confirmaran la presencia de los cinco factores descritos.

El análisis de la estructura interna de la WISC-V conllevó el examen de las relaciones entre los ítems, las pruebas y las puntuaciones compuestas para demostrar la validez de la escala. En el capítulo 5 de este manual se describe la consistencia interna de los ítems de cada una de las pruebas.

Estudios de intercorrelación

Campbell y Fiske (1959) presentaron una metodología teórica para interpretar los patrones de correlaciones en una matriz multirrasgo-multimétodo con el objetivo de demostrar evidencias de *validez convergente* y *validez discriminante*. Este estudio, constantemente citado, es una de los más importantes sobre la validez de los tests (Eid y Nussbeck, 2009). Su metodología se basaba en el estudio del patrón de correlaciones de una matriz para la que se esperan correlaciones relativamente altas (validez convergente) para algunos pares de variables y correlaciones relativamente bajas (validez discriminante) para otros pares. Los datos que apoyan las hipótesis formuladas a priori sobre el patrón de relaciones pueden interpretarse como una evidencia de validez de constructo.

Intercorrelaciones de pruebas, puntuaciones de procesamiento y puntuaciones compuestas

Se formularon varias hipótesis respecto a los estudios de intercorrelación. En primer lugar, se esperaba que todas las puntuaciones mostrasen correlaciones de bajas a moderadas, hipótesis basada en la suposición de que todas las pruebas miden un factor general de inteligencia (es decir, *g*). En segundo lugar, se esperaba que las pruebas de un dominio cognitivo específico (p. ej., Comprensión verbal) tuviesen correlaciones más altas entre ellas que con las pruebas de otros dominios.

En tercer lugar, los resultados obtenidos en estudios previos indican que algunas pruebas están más relacionadas que otras con el factor *g*. Por ejemplo, *Semejanzas, Vocabulario, Información, Cubos, Matrices* y *Aritmética* mostraron tener saturaciones altas en *g* en versiones anteriores de la WISC (Wechsler, 1991, 2003). A partir de estos datos se realizaron varias suposiciones. Así, se esperaba que las pruebas con saturaciones altas en *g*, independientemente del dominio del que forman parte, tuvieran correlaciones relativamente altas entre ellas. Del mismo modo, se esperaba que la correlación entre dos pruebas que presentan saturaciones elevadas en *g* y, además, forman parte del mismo dominio cognitivo (p. ej., Comprensión verbal) fuera más alta que la correlación entre dos pruebas que, aunque muestren saturaciones elevadas en *g*, forman parte de distintos dominios (p. ej., Comprensión verbal y Visoespacial).

En cuarto lugar, y a partir de estudios anteriores (Wechsler, 2003, 2008, 2012), se esperaba que una prueba, *Aritmética*, mostrara correlaciones secundarias relativamente altas con pruebas de otro dominio, concretamente de Memoria de trabajo.

En quinto lugar, se esperaba que las correlaciones entre las pruebas *Cubos*, *Dígitos* y *Cancelación* y sus correspondientes puntuaciones de procesamiento (Cs y Cp; Dd, Di y Dc; CAa y CAe, respectivamente) fueran altas. Aunque esto se deba a veces a la relación parte-todo de las correlaciones, la información sobre los niveles relativos de estas correlaciones puede ser de utilidad para la interpretación.

En el anexo A de este manual se muestran las intercorrelaciones de las puntuaciones de las pruebas, las puntuaciones de procesamiento y las puntuaciones compuestas para cada una de las cinco franjas de edad. Y en la tabla 6.1 se presentan las intercorrelaciones medias de las puntuaciones de las pruebas, las puntuaciones de procesamiento y las puntuaciones compuestas para todas las edades, calculadas utilizando la transformación *z* de Fisher. Todas las tablas incluyen las correlaciones de las pruebas con las sumas de las puntuaciones escalares que contribuyen a cada puntuación compuesta. La correlación entre las sumas de las puntuaciones escalares que contribuyen a cada puntuación compuesta y la puntuación escalar de las pruebas se corrigió eliminando la puntuación escalar de cada prueba de la suma de puntuaciones escalares en la que participaban, como medida de control de los valores inflados. Los coeficientes sin corregir aparecen por debajo de la diagonal y los coeficientes corregidos aparecen por encima de la diagonal, en el área sombreada.

Tabla 6.1. Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas, para todas las edades (datos españoles)

Prueba/ punt. de procesamiento/ punt. compuesta	S	V	I	CO	C	PV	M	B	A	D	SD	LN	CL	BS	CA	Cs	Cp	Dd	Di	Dc	CAa	CAe	ICV	IVE	IRF	IMT	IVP	CIT	IRC	IMTA	INV	ICG	ICC
Semejanzas																							.61					.62				.62	
Vocabulario	.61																						.61					.59				.59	
Información	.60	.59																															
Comprensión	.49	.52	.45																														
Cubos	.40	.35	.38	.30																				.58				.55			.58	.52	
Puzles visuales	.41	.40	.42	.37	.58																			.58							.63		
Matrices	.44	.40	.38	.37	.44	.48																			.48			.60			.59	.58	
Balanzas	.41	.42	.41	.33	.41	.50	.48																		.48			.55	.46		.53	.57	
Aritmética	.48	.47	.54	.41	.44	.45	.48	.46																					.46				
Dígitos	.46	.42	.46	.33	.39	.39	.47	.42	.55																	.46		.59		.68			.44
Span de dibujos	.29	.31	.32	.33	.30	.33	.39	.30	.37	.46																.46					.45		.46
Letras y números	.45	.42	.44	.40	.36	.40	.43	.41	.56	.68	.47																			.68			
Claves	.24	.23	.23	.25	.29	.25	.25	.16	.28	.28	.28	.27															.55	.34			.34		.49
Búsqueda de símbolos	.26	.25	.26	.30	.30	.28	.28	.22	.34	.28	.31	.31	.55														.55						.51
Cancelación	.23	.21	.18	.23	.19	.19	.18	.17	.23	.17	.17	.17	.39	.37																			
Cs	.41	.35	.38	.30	.98	.56	.43	.41	.43	.38	.29	.36	.28	.30	.18																		
Cp	.41	.40	.38	.30	.96	.56	.46	.42	.46	.41	.29	.35	.27	.29	.15	.94																	
Dd	.39	.39	.39	.27	.28	.26	.36	.29	.39	.76	.37	.54	.19	.20	.13	.28	.32																
Di	.36	.31	.35	.24	.31	.32	.39	.32	.45	.81	.39	.54	.23	.22	.13	.32	.33	.48															
Dc	.39	.34	.39	.32	.37	.38	.41	.41	.50	.83	.38	.59	.28	.29	.17	.35	.39	.47	.50														
CAa	.20	.19	.15	.21	.16	.17	.17	.16	.20	.14	.14	.13	.35	.33	.92	.16	.13	.10	.10	.15													
CAe	.25	.23	.21	.23	.20	.20	.19	.17	.24	.20	.19	.19	.39	.36	.91	.19	.16	.16	.16	.18	.72												
ICV	.89	.90	.67	.57	.43	.45	.47	.46	.53	.49	.34	.49	.27	.29	.25	.43	.44	.43	.38	.41	.21	.26											
IVE	.46	.43	.45	.38	.88	.89	.52	.51	.50	.44	.36	.43	.30	.33	.22	.86	.85	.31	.36	.42	.19	.23	.49										
IRF	.49	.48	.46	.40	.49	.57	.86	.86	.54	.52	.40	.49	.24	.29	.20	.49	.50	.37	.41	.47	.19	.21	.54	.59									
IMT	.44	.42	.45	.38	.40	.42	.50	.42	.53	.85	.86	.67	.33	.35	.20	.39	.39	.66	.70	.70	.16	.23	.48	.46	.53								
IVP	.28	.28	.28	.31	.33	.30	.30	.22	.35	.33	.34	.33	.88	.88	.43	.33	.32	.22	.26	.33	.39	.42	.31	.36	.30	.39							
CIT	.74	.72	.63	.54	.68	.62	.72	.69	.66	.71	.48	.63	.51	.44	.32	.68	.67	.55	.57	.63	.28	.33	.82	.74	.82	.69	.54						
IRC	.52	.52	.56	.43	.50	.56	.56	.86	.85	.56	.40	.57	.26	.33	.23	.49	.51	.40	.45	.53	.21	.24	.58	.60	.82	.56	.34	.79					
IMTA	.50	.46	.49	.40	.41	.43	.49	.45	.61	.91	.51	.92	.30	.32	.18	.40	.41	.71	.74	.77	.15	.21	.54	.48	.54	.83	.36	.74	.62				
INV	.53	.51	.52	.46	.73	.77	.74	.70	.60	.58	.63	.57	.54	.47	.31	.72	.70	.42	.47	.54	.27	.32	.58	.84	.83	.71	.57	.90	.76	.63			
ICG	.77	.75	.63	.54	.70	.63	.74	.74	.62	.57	.42	.55	.31	.35	.26	.69	.69	.45	.45	.52	.23	.27	.85	.75	.86	.58	.38	.96	.80	.62	.86		
ICC	.43	.41	.44	.41	.44	.42	.48	.38	.53	.70	.71	.60	.73	.74	.38	.43	.41	.52	.57	.61	.34	.39	.47	.49	.49	.83	.84	.74	.53	.71	.76	.57	
Media	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.1	10.0	10.0	10.0	9.9	9.9	10.0	10.0	10.0	10.0	100.0	99.9	100.0	100.0	100.1	100.1	100.0	100.0	100.0	100.0	100.1
Dt	3.0	3.0	2.9	3.0	2.9	3.0	3.0	3.0	2.9	3.0	3.0	2.9	3.0	3.0	3.0	3.0	2.9	2.9	2.9	3.0	3.0	3.0	14.7	14.7	15.0	14.7	14.7	14.4	14.6	14.9	14.7	14.4	14.8

Nota: Los coeficientes sin corregir aparecen por debajo de la diagonal y los coeficientes corregidos aparecen por encima de la diagonal, en el área sombreada.

Todas las correlaciones entre las pruebas son estadísticamente significativas, y todas las pruebas correlacionan en algún grado con otra. Las puntuaciones de procesamiento escalares correlacionan en mayor medida con sus correspondientes puntuaciones de las pruebas. El patrón de intercorrelaciones es muy similar al de la WISC-IV y al de otras escalas de inteligencia de Wechsler (Wechsler, 2003, 2008, 2012).

En general, las pruebas de Comprensión verbal correlacionan entre sí en mayor medida que con el resto de las pruebas. Las correlaciones entre las pruebas de Comprensión verbal y las pruebas de Visoespacial y Razonamiento fluido son moderadas. Estos resultados son los esperados porque, al igual que las pruebas de Comprensión verbal, tanto las pruebas de Visoespacial como las de Razonamiento fluido y las de Memoria de trabajo tienen altas saturaciones en *g*. Las pruebas de Comprensión verbal muestran correlaciones más altas con las dos pruebas de Memoria de trabajo que tienen estímulos verbales (*Dígitos* y *Letras y números*) que con la prueba basada en estímulos visuales (*Span de dibujos*), pero las diferencias no son considerables, lo que sugiere que en *Span de dibujos* se da cierta mediación verbal. En general, las correlaciones entre las pruebas de Comprensión verbal y las pruebas de Velocidad de procesamiento son bajas.

Las pruebas de Visoespacial correlacionan entre sí en mayor medida que con las pruebas de otras escalas. Las correlaciones entre las pruebas de Visoespacial y las pruebas de Comprensión verbal y Razonamiento fluido son moderadas, al igual que con las pruebas de Memoria de trabajo; en cambio, con las pruebas de Velocidad de procesamiento, las correlaciones son de bajas a moderadas. Desde el punto de vista del constructo, las correlaciones bajas entre los índices y las pruebas de Visoespacial y de Velocidad de procesamiento son importantes porque las pruebas de Visoespacial de la WISC-V se modificaron para incidir menos en la velocidad, respecto a anteriores versiones.

Las pruebas de Razonamiento fluido muestran correlaciones moderadas entre sí, de forma similar a las correlaciones observadas con las pruebas de Comprensión verbal, Visoespacial y Memoria de trabajo. Este resultado es el esperado puesto que las pruebas de las cuatro escalas tienen una alta saturación en *g*. Las correlaciones entre las pruebas de Razonamiento fluido y las pruebas de Memoria de trabajo sugieren una probable intervención de la memoria de trabajo en las tareas de razonamiento fluido, así como a la implicación del córtex prefrontal en ambas aptitudes (Jolles, Kleibeuker, Rombouts y Crone, 2011; Kuwajima y Sawaguchi, 2010). Las correlaciones entre las pruebas de Razonamiento fluido y las pruebas de Velocidad de procesamiento son de bajas a moderadas.

Las pruebas de Memoria de trabajo muestran correlaciones de moderadas a altas entre ellas. Y se observan correlaciones más altas entre las pruebas de Memoria de trabajo auditiva (*Dígitos* y *Letras y números*) que con *Span de dibujos*. Las correlaciones entre las pruebas de Memoria de trabajo y las pruebas de Comprensión verbal, Visoespacial y Razonamiento fluido son moderadas. Las pruebas de Memoria de trabajo y las pruebas de Velocidad de procesamiento muestran correlaciones de bajas a moderadas.

Las pruebas de Velocidad de procesamiento muestran correlaciones moderadas entre ellas, siendo más alta la correlación entre *Búsqueda de símbolos* y *Claves* que la de estas pruebas con *Cancelación*. Entre las pruebas de Velocidad de procesamiento y las pruebas del resto de las escalas se observan correlaciones de bajas a moderadas.

La prueba *Aritmética* merece mención aparte porque muestra correlaciones moderadas con algunas o con todas las pruebas de las cinco escalas. Tal y como se ha señalado en estudios anteriores sobre las escalas de inteligencia de Wechsler (p. ej., Keith et al., 2006), esta prueba es factorialmente compleja en la WISC-V.

Los índices secundarios correlacionan en mayor medida con aquellos índices primarios con los que tiene pruebas y contenido comunes. El IRC muestra una correlación más alta con el IRF; y el IMTA, con el IMT. El INV y el ICG, que son medidas más globales que otros índices secundarios, muestran correlaciones altas con el CIT; pero el INV correlaciona en mayor medida con el IVE y el IRF; y el ICG, con el ICV y el IRF. El ICC muestra las correlaciones más altas con el IMT y el IVP.

Aunque las correlaciones varían ligeramente, estos patrones suelen repetirse en las cinco franjas de edad. Estos datos confirman la hipótesis de que la correlación entre pruebas que pretenden medir dominios cognitivos similares es mayor que la correlación que obtienen pruebas que miden dominios cognitivos diferentes, lo que supone una demostración inicial de la validez convergente y la validez discriminante de la WISC-V.

Análisis factorial confirmatorio

El objetivo del análisis factorial es identificar un modelo factorial (un conjunto de factores y la estructura de las relaciones que establecen con las pruebas y entre ellos) que explique adecuadamente las correlaciones observadas entre las pruebas. Uno de los resultados más importantes del análisis factorial es identificar cuántos factores, y de qué tipo, son necesarios para explicar cómo interrelacionan las pruebas. Por ejemplo, ¿las evidencias sugieren que en el rendimiento de todas las pruebas influyen tres, cuatro o cinco dimensiones de aptitud distintas? El análisis factorial es una herramienta que permite identificar el menor número de factores que explican los datos. ¿Existen evidencias de que entre tales dimensiones se encuentran unas aptitudes cognitivas específicas (p. ej., razonamiento fluido)? El análisis del contenido de las pruebas que guardan una estrecha relación con un factor respalda el establecimiento de inferencias sobre la naturaleza de la aptitud representada por dicho factor. De este modo, el análisis factorial relaciona las medidas de un test con los modelos de evaluación existentes.

Para lograr el objetivo descrito, es más útil incluir un gran número de pruebas en el análisis, puesto que aumenta la validez de constructo, la replicabilidad y la interpretabilidad de los factores, con lo que permite la comparación estadística de modelos alternativos. Por lo tanto, un análisis que incluya varias pruebas de cada tarea y de cada tipo de destreza podrá ofrecer una descripción más completa de los factores subyacentes. Después de haber identificado un modelo factorial óptimo en un grupo de pruebas grande, este mismo modelo puede adecuarse a un grupo de datos reducido (p. ej., las pruebas principales) a fin de confirmar que el conjunto de factores del modelo sigue explicando las correlaciones entre las pruebas.

El modelo factorial de una escala puede evaluarse mediante un tipo de modelo de ecuaciones estructurales (Jöreskog y Sörbom, 1993): el análisis factorial confirmatorio. Esta técnica está diseñada para comprobar una estructura factorial definida previamente por los investigadores a partir de estudios teóricos y empíricos. Las relaciones específicas entre las variables latentes (constructos como la aptitud de comprensión verbal) y las variables observadas (puntuaciones de las pruebas) y entre las variables latentes se especifican en el modelo. En este sentido, el análisis factorial confirmatorio es distinto del análisis factorial exploratorio, porque las pruebas se han distribuido previamente en grupos que representan diferentes factores, en vez de que los grupos se hayan creado al aplicar algoritmos estadísticos a los datos. El modelo especificado de antemano se pone a prueba para determinar si ofrece una buena justificación de las covarianzas entre las variables observadas. El resultado del análisis puede sugerir formas de mejorar el modelo para explicar mejor los datos; no obstante, estos cambios solo deben llevarse a cabo tras haber considerado con detalle su adecuación a la teoría.

Por varias razones, el análisis factorial confirmatorio es preferible al análisis factorial exploratorio cuando se dispone de una teoría explícita de la estructura factorial o cuando se han desarrollado modelos alternativos al respecto (Schmitt, 2011; Stevens, 1996). El método confirmatorio proporciona una evaluación numérica de la calidad de un modelo determinado, lo que permite al investigador comparar la efectividad de distintos modelos. Asimismo, cuando las dimensiones de aptitud hipotéticas correlacionan significativamente entre sí, el análisis factorial confirmatorio es más sólido que el análisis factorial exploratorio, cuyos resultados pueden ser muy sensibles a pequeñas diferencias en el patrón de correlaciones.

Ajuste estadístico

La efectividad de un modelo factorial se evalúa mediante índices de bondad de ajuste. La mayoría de estos indicadores se basan en el estadístico chi cuadrado. La prueba chi cuadrado se utiliza para examinar la hipótesis nula según la cual la matriz de varianza-covarianza de las pruebas basada en las cargas factoriales y las intercorrelaciones factoriales es igual que la matriz de varianza-covarianza observada (Byrne, 2001; Schumacker y Lomax, 2004; Thompson, 2000). Por lo tanto, si el modelo se ajusta adecuadamente a los datos, la hipótesis nula no se rechaza.

Sin embargo, cuando la muestra es de gran tamaño, la prueba chi cuadrado tiende a detectar diferencias entre la matriz de varianza-covarianza observada en la muestra de tipificación y la estimada y a rechazar la hipótesis nula incluso cuando las diferencias son pequeñas. Por ello, la evaluación del ajuste del modelo se basa en otras pruebas estadísticas que no son dependientes del

tamaño de la muestra (Byrne, 2001; Schumacker y Lomax, 2004; Thompson, 2000). No obstante, la prueba chi cuadrado sigue siendo necesaria dada su utilidad para examinar la significación estadística de la diferencia en el ajuste de dos modelos.

El análisis de la WISC-V presenta cinco índices adicionales de bondad de ajuste que generalmente se usan para determinar el grado en que la matriz de varianza-covarianza observada en la muestra de tipificación y la estimada son equivalentes. El índice de ajuste comparativo (*CFI*; Bentler, 1990) evalúa cómo el modelo específico reproduce la matriz de la varianza-covarianza observada en comparación con un modelo que supone que las pruebas no están intercorrelacionadas. Hu y Bentler (1999) recomiendan un valor de *CFI* de .95 o superior como el criterio de un buen ajuste. El índice de Tucker-Lewis (*TLI*; Tucker y Lewis, 1973) mide el ajuste incremental de una forma similar al *CFI*, con la diferencia de estar corregido por los grados de libertad del modelo. Para analizar el ajuste del modelo respecto a la matriz de covarianza poblacional, se ha utilizado el error de aproximación cuadrático medio (*RMSEA*) de Steiger (1990). El *RMSEA* es una estimación de cómo la varianza-covarianza estimada reproduce la matriz de la población, corregida por los grados de libertad del modelo ajustado. Browne y Cudeck (1993) sugieren que un valor de *RMSEA* de .05 o menor indica un ajuste próximo al modelo y que los valores inferiores a .08 representan un ajuste adecuado con errores de aproximación razonables en la población. Asimismo se han incluido el criterio de información de Akaike (*AIC;* Akaike, 1974) y el criterio de información bayesiano (*BIC*; Schwarz, 1978). Cuando se comparan dos modelos con estos estadísticos, los valores más bajos representan un ajuste mejor.

Modelos

A lo largo de 70 años se han realizado numerosas investigaciones con las escalas de inteligencia de Wechsler. Durante este tiempo, los avances en la teoría de la inteligencia y las mejoras logradas en las técnicas psicométricas han guiado las revisiones de las escalas Wechsler. En estudios posteriores se ha examinado la estructura de la WISC, la WAIS (Wechsler, 1955) y la WPPSI (Wechsler, 1967), y las correspondientes versiones, lo que ha dado lugar a una evolución sistemática y deliberada de estas escalas. La incorporación de pruebas nuevas (p. ej., *Claves* y *Búsqueda de símbolos*) ha llevado, a través del uso del análisis factorial, a identificar nuevos constructos fiables y distintivos (p. ej., Velocidad de procesamiento).

Como se ha señalado anteriormente, los estudios de análisis factorial de la WISC-IV y la WAIS-IV han indicado que miden cuatro dominios cognitivos (p. ej., Bodin et al. 2009; Bowden, Weiss, Holdnack y Lloyd, 2006; Ward et al., 2011; Watkins, 2010; Watkins et al., 2006; Wechsler, 2003, 2008). Esta estructura de cuatro factores ha demostrado ser muy robusta en las adaptaciones realizadas en países anglófonos (p. ej., Chen, Keith, Weiss, Zhu, y Li, 2010). No obstante, algunos autores, basándose en otras perspectivas y modelos, evaluaron la adecuación de modelos de dos o tres factores y concluyeron que dichos modelos podían igualmente explicar adecuadamente los datos (Kaufman, Lichtenberger y McLean, 2001; Ward, Ryan y Axelrod, 2000). Asimismo, varios investigadores han identificado una estructura de cinco factores alternativa y común a la WISC-IV y la WAIS-IV, que puede tener ventajas (Benson et al., 2010; Keith et al., 2006; Weiss et al., 2013a, 2013b). Los nombres de los factores varían en función del sistema taxonómico seguido por el equipo investigador y vienen a coincidir con Comprensión verbal, Visoespacial, Razonamiento fluido, Memoria de trabajo y Velocidad de procesamiento.

El desarrollo de la WISC-V se basó en el supuesto teórico de que la escala ofrece una estimación de la aptitud cognitiva general, que se manifiesta en cinco dominios cognitivos (comprensión verbal, visoespacial, razonamiento fluido, memoria de trabajo y velocidad de procesamiento). Ello se corresponde con un modelo factorial de segundo orden con cinco factores de primer orden (los cinco dominios de aptitud) y un factor de segundo orden (inteligencia general o *g*). Las tres pruebas nuevas (*Puzles visuales, Balanzas* y *Span de dibujos*) se desarrollaron con el objetivo de diferenciar aún más la medición de las aptitudes de Visoespacial, Razonamiento fluido y Memoria de trabajo, respectivamente. La distribución esperada de las pruebas nuevas dentro del modelo de cinco factores estaba clara cuando se iniciaron los análisis factoriales confirmatorios basados en la muestra de tipificación.

Procedimiento

Cuando se tuvieron todos los protocolos de la muestra de tipificación, se llevaron a cabo una serie de análisis factoriales confirmatorios con el objetivo de comparar el modelo de cinco factores hipotetizado con modelos formados por menos factores, así como para investigar cómo medía cada prueba el factor de la aptitud prevista.

En el primer modelo de la serie (modelo 1) todas las pruebas contribuyen al factor *g*, del que se supone que es el único responsable de las correlaciones entre las pruebas. Este modelo sirvió como línea base. Los otros modelos son jerárquicos: conservan la aptitud general como factor de orden superior, pero incorporan de dos a cinco factores de primer orden que representan dominios de aptitud, como la Comprensión verbal. En cada modelo sucesivo se incrementa el número de factores de primer orden o se modifica la asignación de pruebas a los factores, a partir de la teoría y las investigaciones anteriores. La tabla 6.2 presenta las pruebas asociadas a cada factor de primer orden en los modelos jerárquicos.

Tabla 6.2. Modelos del análisis factorial confirmatorio de la WISC-V (datos españoles)

	Modelo factorial																			
	2		3			4a				4b				4c				4d			
Prueba	F1	F2	F1	F2	F3	F1	F2	F3	F4	F1	F2	F3	F4	F1	F2	F3	F4	F1	F2	F3	F4
S	•		•			•				•				•				•			
V	•		•			•				•				•				•			
I	•		•			•				•				•				•			
CO	•		•			•				•				•				•			
C		•		•			•				•				•				•		
PV		•		•			•				•				•				•		
M		•		•			•					•			•				•		
B		•		•			•					•			•				•		
A	•		•					•				•			•	•		•	•	•	
D	•		•					•				•				•				•	
SD		•		•				•				•				•				•	
LN	•		•					•				•				•				•	
CL		•			•				•				•				•				•
BS		•			•				•				•				•				•
CA		•			•				•				•				•				•

	Modelo factorial																				
	5a					5b					5c					5d					5e				
Prueba	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5
S	•					•					•					•					•				
V	•					•					•					•					•				
I	•					•					•					•					•				
CO	•					•					•					•					•				
C		•					•					•					•					•			
PV		•					•					•					•					•			
M			•					•					•					•					•		
B			•					•					•					•					•		
A				•				•					•	•		•			•		•		•	•	
D				•					•					•					•					•	
SD				•					•					•					•					•	
LN				•					•					•					•					•	
CL					•					•					•					•					•
BS					•					•					•					•					•
CA					•					•					•					•					•

El modelo 2 representa la estructura tradicional de dos factores, con el factor verbal y el factor manipulativo separados. El modelo 3 permite examinar la distinción de las pruebas de Velocidad de procesamiento respecto a las tradicionalmente denominadas *pruebas manipulativas* (pruebas de Visoespacial y de Razonamiento fluido y la prueba *Span de dibujos*). Es un modelo de tres factores con un factor combinado Comprensión verbal y Memoria de trabajo auditiva; un factor combinado Visoespacial, Razonamiento fluido y Memoria de trabajo visual; y un factor Velocidad de procesamiento.

Los modelos de cuatro factores (4a-4d) definen de idéntica manera los factores Comprensión Verbal y Velocidad de procesamiento, pero difieren en cómo agrupan las ocho pruebas restantes. El modelo 4a evalúa la distinción de las pruebas de Memoria de trabajo respecto a las pruebas de Visoespacial y de Razonamiento fluido y refleja los modelos publicados de la WISC-IV y la WAIS-IV (un modelo de cuatro factores con un factor Comprensión verbal, un factor combinado Visoespacial y Razonamiento fluido, un factor Memoria de trabajo —que incluye la prueba *Aritmética*—, y un factor Velocidad de procesamiento). El modelo 4b, basado en los estudios de neurociencia cognitiva que indican que el razonamiento fluido y la memoria de trabajo comparten una función del córtex prefrontal, agrupa las pruebas de estos dos dominios en un único factor y lo contrasta con un factor Visoespacial. Los modelos 4c y 4d son variantes del modelo 4a en las que la prueba *Aritmética* puede contribuir a otro factor además de contribuir principalmente al factor Memoria de trabajo. En el modelo 4c, *Aritmética* contribuye también al factor Razonamiento fluido; y en el modelo 4d, al factor Comprensión verbal. Estas cargas fueron sugeridas por análisis factoriales de la WISC-IV realizados con anterioridad (p. ej., Keith et al., 2006; Weiss et al., 2013b).

Los modelos 5a-5e son variantes de un modelo de cinco factores que incluye Comprensión verbal, Visoespacial, Razonamiento fluido, Memoria de trabajo y Velocidad de procesamiento. Se diferencian por el patrón de carga de *Aritmética*. Estudios sobre la WISC-IV (Keith et al., 2006; Weiss et al., 2013b) y la WAIS-IV (Benson et al., 2010; Weiss et al., 2013a) indican que *Aritmética* contribuye al factor Memoria de trabajo en los modelos de cuatro factores; en cambio, en los modelos de cinco factores contribuye al factor Razonamiento fluido. Estas dos distribuciones se representan mediante los modelos 5a y 5b. En el resto de los modelos, *Aritmética* contribuye a distintos factores de forma simultánea por lo que, además de contribuir principalmente al factor Memoria de trabajo, contribuye al Razonamiento fluido (modelo 5c), a la Comprensión verbal (modelo 5d) o a ambos (modelo 5e).

Los modelos se analizaron según su nivel de complejidad, desde el más sencillo (dos factores de primer orden) hasta el más complejo (cinco factores), para poder evaluar estadísticamente la mejora en el ajuste obtenida conforme se incrementa la complejidad. El resultado de esta evaluación es un criterio importante para decidir el número de factores que debe incluir el modelo final. A menudo, un modelo más sencillo *está integrado* en un modelo más complejo, de tal modo que es un caso especial del modelo más complejo obtenido al asignar valores fijos a uno o más parámetros del modelo más complejo. Por ejemplo, el modelo 1 está integrado en el modelo 2, porque, al correlacionar los dos factores del modelo 2 en 1.0, se crearía el modelo 1. Cuando un modelo está integrado en otro, la significación estadística de la mejora en el ajuste puede examinarse comparando los valores de chi cuadrado para los dos modelos (Bentler y Bonett, 1980). Cuando los modelos no están integrados, el cambio en el ajuste se evalúa mediante una comparación más subjetiva que estadística de los valores de ajuste de los modelos.

La estrategia de investigación de la estructura factorial de la WISC-V consistió, primero, en el análisis de todas las pruebas, principales y opcionales, para todo el rango de edad. Tras identificar un modelo factorial óptimo en todas las pruebas de la WISC-V, se evaluó el ajuste de dicho modelo por separado en las cinco franjas de edad siguientes: 6-7, 8-9, 10-11, 12-13 y 14-16. Finalmente, se ajustó el modelo óptimo a las diez pruebas principales para confirmar que los factores del modelo seguían explicando las correlaciones entre este grupo de pruebas.

Resultados

La tabla 6.3 presenta los resultados del análisis factorial confirmatorio de todas las pruebas para todo el rango de edad. Las columnas de la derecha muestran la mejora en el ajuste entre el modelo de una fila determinada y un modelo más sencillo en el que está integrado. El modelo 2 se compara con el modelo 1; y el modelo 3, con el modelo 2. No se ofrecen datos sobre la comparación de los modelos de cuatro factores con el modelo 3 porque el modelo 3 no está integrado en ellos. Se realizaron, además, comparaciones entre los modelos de cuatro factores y los de cinco factores para evaluar si el cambio entre los modelos daba lugar a una mejora significativa estadísticamente.

Tabla 6.3. Índices de bondad de ajuste para el análisis factorial confirmatorio (datos españoles)

	Índices de bondad de ajuste							Incremento		
Modelo	χ^2	*gl*	*CFI*	*TLI*	*RMSEA*	*AIC*	*BIC*	$\Delta\chi^2$	Δgl	*p*
					Todas las pruebas					
Modelo 1	1138.3	90	.83	.80	.11	1198	1346			
Modelo 2	987.7	86	.85	.82	.10	1056	1223	150.6	4	<.01
Modelo 3	625.1	84	.91	.89	.08	697	874	362.6	2	<.01
Modelo 4a	303.0	82	.96	.95	.05	379	566			
Modelo 4b	368.5	82	.95	.94	.06	445	632			
Modelo 4c[a]	252.8	81	.97	.96	.05	331	523	50.2	1	<.01
Modelo 4d[b]	231.1	80	.97	.97	.04	311	508	21.7	1	<.01
Modelo 5a[a]	282.4	80	.97	.96	.05	362	560	20.6	2	<.01
Modelo 5b[a]	231.4	80	.97	.97	.04	311	509	71.6	2	<.01
Modelo 5c[c]	218.1	79	.98	.97	.04	300	502	64.3	1	<.01
Modelo 5d[c]	220.3	79	.98	.97	.04	302	504	62.1	1	<.01
Modelo 5e										
Todas las edades[d]	207.3	78	.98	.97	.04	291	498	10.8	1	<.01
Edad 6-7	123.2	78	.96	.95	.05	207	345			
Edad 8-9	99.4	78	.98	.97	.04	183	320			
Edad 10-11	145.9	78	.95	.93	.07	230	367			
Edad 12-13	102.1	78	.98	.97	.04	186	320			
Edad 14-16	107.7	78	.98	.97	.04	192	341			
					Pruebas principales					
Modelo 5e	87.5	25	.98	.97	.05	148	296			

Nota: Los valores chi cuadrado están calculados con el método de mínimos cuadrados ponderados (WLS) del paquete estadístico 9.3 de SAS.

[a]La mejora en el ajuste del modelo se basa en la comparación con el modelo 4a.

[b]La mejora en el ajuste del modelo se basa en la comparación con el modelo 4c.

[c]La mejora en el ajuste del modelo se basa en la comparación con el modelo 5a.

[d]La mejora en el ajuste del modelo se basa en la comparación con el modelo 5c.

Los resultados que figuran en la tabla 6.3 indican que se necesitan, por lo menos, cuatro factores para obtener un buen ajuste según los criterios convencionales del *CFI* y el *RMSEA*. De los modelos de cuatro factores, el modelo 4a, que combina las pruebas de Visoespacial y las de Razonamiento fluido en un factor Razonamiento perceptivo (RP), tiene un claro mejor ajuste que el modelo 4b, en el que las pruebas de Razonamiento fluido y Memoria de trabajo componen un factor. Al modificar el modelo 4a para que *Aritmética* contribuya tanto al Razonamiento perceptivo como a la Memoria de trabajo (MT; modelo 4c) se obtiene una mejora considerable en el ajuste; y, al permitir que *Aritmética* también contribuya a la Comprensión verbal (CV; modelo 4d), la mejora es aún más significativa. En el modelo 4d, las cargas factoriales son .34 (MT), .23 (RP) y .23 (CV).

Todos los modelos de cinco factores producen un ajuste excelente, significativamente mejor que el ajuste de los modelos comparables de cuatro factores. El ajuste del modelo 5a (con *Aritmética* solo en Memoria de trabajo) es significativamente mejor que el del modelo 4a (el modelo de la WISC-V, también con *Aritmética* solo en Memoria de trabajo). Cuando *Aritmética* contribuye solo al Razonamiento fluido (modelo 5b), el ajuste mejora, pero este modelo produce una saturación en *g* para el Razonamiento fluido (1.03), que es significativamente superior a 1, lo que indica que el modelo es incompleto o está mal especificado. Al permitir que *Aritmética* contribuya tanto al Razonamiento fluido como a la Memoria de trabajo (modelo 5c), se obtiene una mejora significativa respecto al modelo 5a, al igual que cuando también contribuye a la Comprensión verbal (modelo 5e). En este modelo final, las cargas factoriales de *Aritmética* son .31 (MT), .28 (RF) y .21 (CV).

El análisis del modelo 5e indicó que el ajuste podía mejorar si *Balanzas* contribuía tanto al factor Visoespacial como al Razonamiento fluido. No obstante, y pese a que se reproducía la observación realizada por Weiss et al. (2013a) en la WAIS-IV, se decidió no incorporar esta modificación al modelo de la WISC-V porque no se creyó que tuviera una base teórica clara. Por ello se seleccionó el modelo 5e para representar la estructura factorial de la WISC-V correspondiente a todas las pruebas, principales y opcionales, y para todas las edades.

En la figura 6.1 se ilustra este modelo, con las cargas factoriales. Como indican numerosos análisis recientes de escalas de inteligencia (p. ej., Kaufman y Kaufman, 2004; Keith et al., 2006; Wechsler, 2008, 2012; Weiss et al., 2013a, 2013b), el Razonamiento fluido correlaciona casi perfectamente con el factor de aptitud general de orden superior.

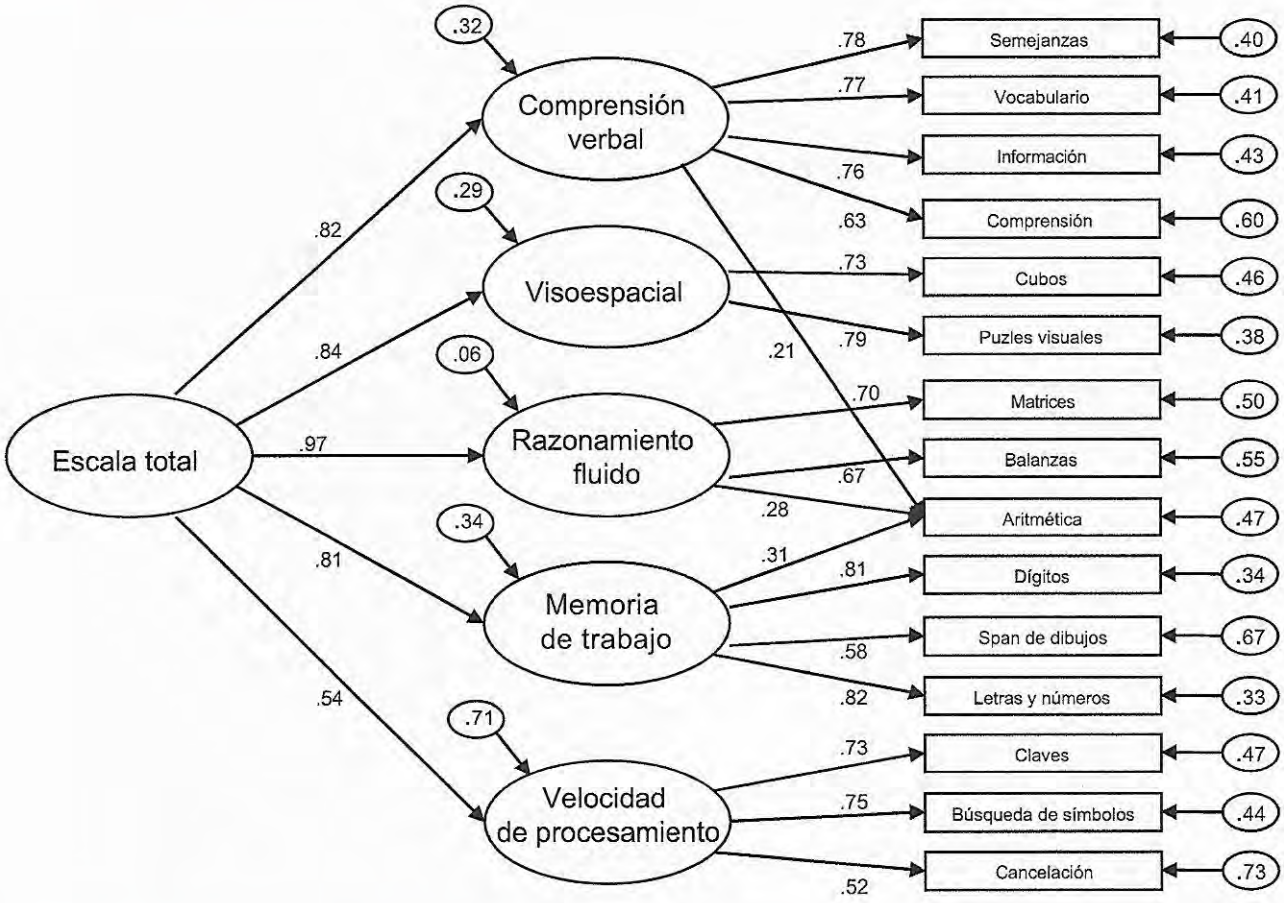


Figura 6.1. Modelo jerárquico de cinco factores para todas las pruebas (6-16 años) (datos españoles)

Además, el modelo 5e se analizó para las cinco franjas de edad siguientes: 6-7, 8-9, 10-11, 12-13 y 14-16. Los resultados (presentados en la tabla 6.3) indican que el ajuste es bueno para cada una de las franjas de edad; por tanto, este modelo refleja adecuadamente las intercorrelaciones de las pruebas en cada una de estas franjas de edad.

Asimismo, se examinó una versión no jerárquica del modelo 5e para evaluar las intercorrelaciones de los cinco factores de este modelo en todo el rango de edad. Las intercorrelaciones factoriales van de .44 (entre Razonamiento fluido y Velocidad de procesamiento) a .87 (entre Visoespacial y Razonamiento fluido). Las intercorrelaciones con el factor de Velocidad de procesamiento son inferiores a las que se dan entre los otros cuatro factores.

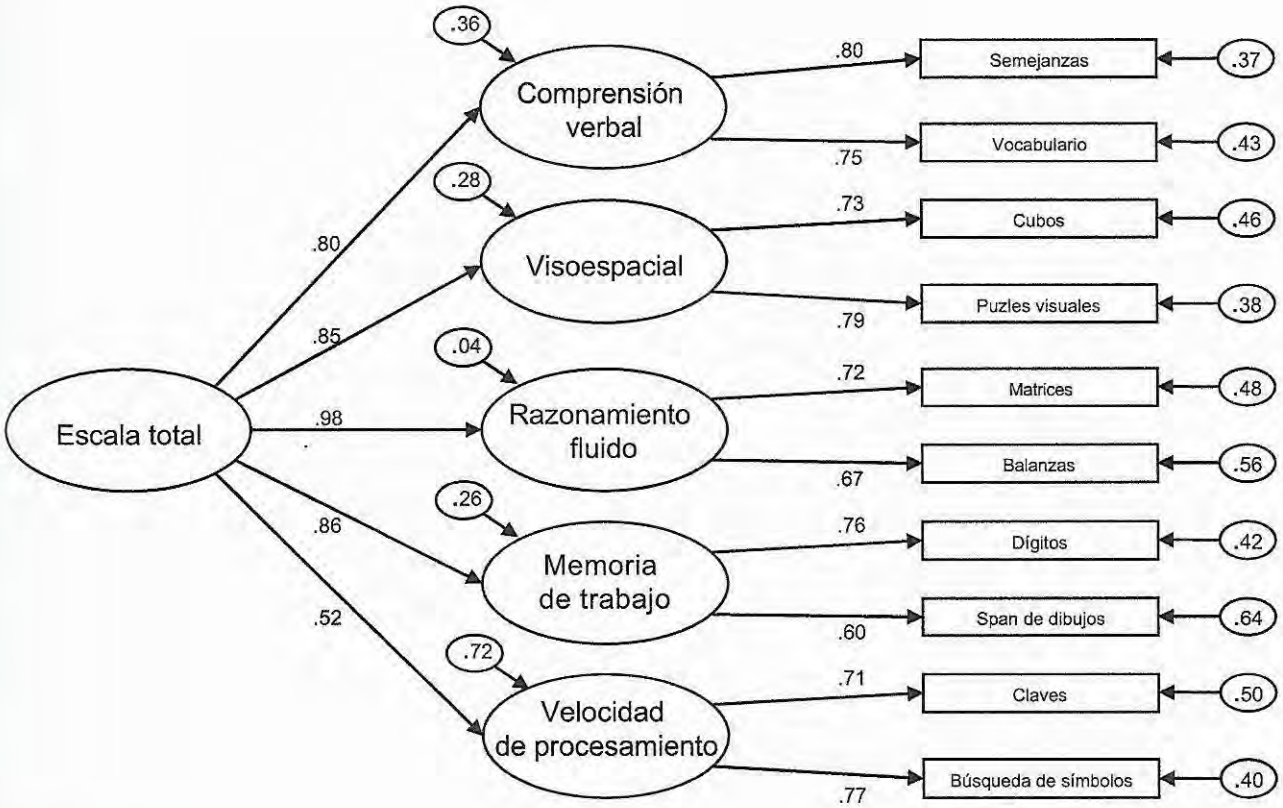


Figura 6.2. Modelo jerárquico de cinco factores para las pruebas principales (6-16 años) (datos españoles)

Las pruebas de la misma escala tienen saturaciones en *g* similares. Las saturaciones en *g* más altas se observan en las pruebas de Razonamiento fluido: *Matrices* (.68), *Balanzas* (.65) y *Aritmética* (.69). Les siguen las pruebas de Visoespacial, Comprensión verbal y Memoria de trabajo: *Puzles visuales* (.66), *Cubos* (.61), *Semejanzas* (.64), *Vocabulario* (.63), *Información* (.62), *Comprensión* (.52), *Dígitos* (.66), *Letras y números* (.66) y *Span de dibujos* (.47). Y, finalmente, las pruebas con las saturaciónes en *g* más bajas son las de Velocidad de procesamiento: *Búsqueda de símbolos* (.41), *Claves* (.39) y *Cancelación* (.28).

Análisis de las pruebas principales

El modelo seleccionado para representar la estructura de la WISC-V (modelo 5e) se ajustó a las diez pruebas principales para todo el rango de edad. En esta fase, el objetivo de la investigación no se centró en el número de factores que influyen en las pruebas, sino en cuán bien los factores identificados justifican las intercorrelaciones del reducido grupo de pruebas principales. La tabla 6.3 muestra los índices de bondad de ajuste, y la figura 6.2 ilustra las cargas factoriales de las pruebas principales. El ajuste es excelente, y las cargas factoriales son similares a las obtenidas del análisis de todas las pruebas, tanto principales como opcionales. Estos resultados respaldan la efectividad de este modelo de cinco factores para explicar las intercorrelaciones de las pruebas.

Evidencias basadas en las relaciones con otras variables

El estudio de la relación entre las puntuaciones de los tests con variables externas teóricamente relevantes proporciona una evidencia adicional acerca de la precisión de las inferencias basadas en dichas puntuaciones. En esta sección se describen las relaciones de las puntuaciones de la WISC-V con dos tipos de variables externas: por una parte, las puntuaciones de otros instrumentos diseñados para medir constructos idénticos o similares; y, por otra, las puntuaciones de niños que pertenecen a grupos especiales. Los estudios que aportan estas evidencias se llevaron a cabo durante las fases de tipificación estadounidense y española de la WISC-V.

En la primera parte de esta sección se describen las relaciones entre las puntuaciones de la WISC-V y otras medidas de aptitud intelectual o cognitiva, así como medidas de la conducta adaptativa. Históricamente, este tipo de datos ha recibido diversas denominaciones: *validez concurrente, validez convergente y discriminante, validez predictiva* y *validez de criterio.*

En la segunda parte de esta sección se describen los estudios que respaldan la validez de la WISC-V para distintos grupos clínicos de niños (p. ej., alta capacidad intelectual, discapacidad intelectual, TDAH, trastorno del espectro autista).

Relaciones con otras medidas

Cronbach y Meehl (1955) consideran que las evidencias de validez obtenidas a partir de las relaciones con otras medidas permiten situar los constructos medidos dentro de una red nomológica de variables conocidas. Campbell y Fiske (1959) también abordaron la importancia de establecer asociaciones entre constructos estrechamente relacionados (validez convergente) y entre constructos vagamente relacionados (validez discriminante) cuando se quiere determinar el constructo, o constructos, que mide un test.

Con las muestras no clínicas se analizaron las relaciones entre la WISC-V y los siguientes instrumentos de evaluación: WISC-IV, WPPSI-IV, WAIS-IV, KABC-II, Vineland-II y BASC–2 CPP. La tabla 6.4 recoge el tamaño de la muestra y los datos demográficos para cada uno de los estudios comparativos. La WISC-V y las otras medidas de aptitud intelectual (WISC-IV, WPPSI-IV, WAIS-IV y KABC-II) se aplicaron en orden contrabalanceado. En los estudios contrabalanceados, las medias, las desviaciones típicas y los coeficientes de correlación corregidos se calcularon por separado para cada parte de la muestra, según el orden de aplicación y, posteriormente, se calculó el promedio. Este método evita que los efectos de aplicación repetida reduzcan artificialmente los coeficientes de correlación. En cada estudio se presenta la diferencia típica, una medida del tamaño del efecto; es decir, la diferencia entre las puntuaciones medias de dos instrumentos dividida por la desviación típica común o, en otras palabras, la diferencia media expresada en unidades de desviación típica.

Tabla 6.4 Datos demográficos para los estudios de validez con otras medidas

	WISC–IV*	WPPSI–IV	WAIS–IV	KABC–II	VINELAND–II	BASC–2 PRS
N	80	105	112	89	61	2302
Edad						
Media	11.0	7.0	16.5	11.7	11.8	11.5
Dt	2.9	0.5	0.3	2.8	3.2	3.2
Sexo						
Mujer	56.3	55.2	57.1	50.6	49.2	51.6
Varón	43.8	44.8	42.9	49.4	50.8	48.4

Nota: Excepto para el tamaño de la muestra (*N*) y la edad, los datos se muestran en porcentajes. Si el porcentaje total no suma 100, es debido al redondeo.

* Datos españoles

Correlaciones con la WISC-IV

Se aplicó la WISC-V y la WISC-IV a 80 niños de 6 a 16 años, en orden contrabalanceado y con un intervalo de aplicación de 8 a 47 días (media = 23 días). La tabla 6.5 muestra las medias, las desviaciones típicas, las correlaciones (corregidas y sin corregir) y las diferencias típicas de las pruebas, de las puntuaciones de procesamiento y las puntuaciones compuestas.

Tabla 6.5. Correlaciones entre la WISC-V y la WISC-IV (datos españoles)

Prueba/ punt. de procesamiento/ punt. compuesta	WISC-V			WISC-IV					
	Media[a]	*Dt*[a]	*n*	Media[a]	*Dt*[a]	*n*	r_{12}[b]	r_{12}[b] corregido	Diferencia típica[c]
Semejanzas	10.5	2.6	80	11.7	2.9	80	.50	.51	0.44
Vocabulario	10.1	2.8	80	10.9	2.8	80	.62	.63	0.29
Información	10.3	2.6	80	11.1	2.9	80	.64	.70	0.29
Comprensión	9.9	2.5	78	9.7	3.0	78	.41	.48	0.07
Adivinanzas	—	—	—	11.6	3.0	79	—	—	—
Cubos	10.7	2.8	80	11.7	3.1	80	.72	.73	0.34
Puzles visuales	10.5	2.9	80	—	—	—	—	—	—
Matrices	10.8	2.7	80	11.7	2.9	80	.57	.60	0.32
Balanzas	10.0	2.7	80	—	—	—	—	—	—
Conceptos	—	—	—	11.6	3.1	79	—	—	—
Figuras incompletas	—	—	—	10.6	3.0	79	—	—	—
Aritmética	10.2	2.5	80	10.8	3.1	80	.62	.66	0.21
Dígitos	10.7	2.9	80	11.2	3.4	80	.61	.62	0.16
Span de dibujos	10.0	3.1	80	—	—	—	—	—	—
Letras y números	9.6	2.6	79	11.2	3.3	79	.53	.51	0.54
Claves	10.5	3.0	79	10.4	3.3	79	.61	.62	0.03
Búsqueda de símbolos	10.4	3.2	73	12.0	3.4	73	.51	.45	0.48
Cancelación	10.7	3.0	80	10.7	3.4	80	.61	.63	0.00
ICV	101.7	2.7	80	105.1	13.5	80	.60	.64	0.26
IVE-IRP	103.2	3.8	80	109.7	14.6	80	.72	.73	0.46
IRF-IRP	102.2	3.3	80	109.7	14.6	80	.56	.60	0.54
IMT	101.7	4.4	80	105.8	16.8	80	.61	.59	0.26
IVP	103.0	5.1	74	108.0	14.4	74	.58	.60	0.34
CIT	103.1	2.4	80	108.6	14.3	80	.79	.83	0.41
IMTA-IMT	100.7	3.3	80	105.8	16.8	80	.66	.64	0.34
ICG	102.6	2.4	78	94.8	11.9	78	.72	.79	0.64
ICC	102.2	4.5	71	109.0	15.9	71	.72	.70	0.45

Nota: El IMTA se correlacionó con el IMT de la WISC-IV. Las correlaciones se calcularon por separado para cada orden de aplicación en un diseño contrabalanceado y se corrigieron de acuerdo con la variabilidad de la muestra de tipificación española de la WISC-V (Guilford y Fruchter, 1978).

[a] Los valores de las columnas «Media» y «*Dt*» son el promedio de las medias y desviaciones típicas de los dos órdenes de aplicación.

[b] La media del coeficiente de correlación de las dos aplicaciones se obtuvo mediante la transformación *z* de Fisher.

[c] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Las medias de los índices primarios y del CIT de la WISC-V se sitúan en el rango promedio. Las diferencias típicas de las comparaciones entre las medias del IRF y el IRP y del ICG de la WISC-V y el ICG de la WISC-IV son moderadas; y las del resto de comparaciones entre los otros índices principales son pequeñas. Las medias de los índices, con excepción del ICG, son más bajas en la WISC-V que en la WISC-IV, resultados que son consistentes con el efecto Flynn. En cuanto a las pruebas, se observan diferencias típicas superiores a .40 en *Semejanzas*, *Letras y números* y *Búsqueda de símbolos*.

Las diferencias entre las medias de las puntuaciones compuestas de la WISC-IV y las correspondientes de la WISC-V siguen aproximadamente la magnitud predicha por el efecto Flynn (Flynn, 2007).

Los coeficientes de correlación corregidos de las puntuaciones compuestas de ambas escalas oscilan entre .59 para el IMT y .83 para el CI total. La correlación entre el IRF de la WISC-V y el IRP de la WISC-IV (.60) refleja la distinta naturaleza de estos índices, que solo tienen en común una prueba.

Asimismo, la correlación entre el IMT de la WISC-V y el de la WISC-IV (.59) refleja la sustitución de una tarea de memoria de trabajo auditiva (*Letras y números*) en el IMT de la WISC-IV por una tarea de memoria de trabajo visual (*Span de dibujos*) en el IMT de la WISC-V. Esta interpretación se ve respaldada por la correlación ligeramente superior (.64) observada entre el IMTA de la WISC-V, que incluye únicamente contenido auditivo, y el IMT de la WISC-IV. Además, la prueba *Dígitos* de la WISC-IV se modificó significativamente en la WISC-V al añadirle la tarea *Dígitos* en orden creciente.

Las correlaciones corregidas entre las pruebas comunes a la WISC-V y la WISC-IV son de moderadas a altas, lo que indica que miden constructos similares. La correlación más baja se observa en *Búsqueda de símbolos*, pero las otras dos pruebas de Velocidad de procesamiento, *Claves* y *Cancelación*, muestran correlaciones más altas. Las pruebas de Velocidad de procesamiento se revisaron para mejorar el control experimental. Estos resultados son consistentes con las investigaciones previas llevadas a cabo para analizar la relación entre distintas versiones de las escalas de inteligencia de Wechsler (Wechsler, 1991, 1997, 2002, 2003, 2008, 2012). La magnitud de las correlaciones entre la WISC-V y la WISC-IV sugieren que ambos instrumentos miden constructos similares.

Correlaciones con la WPPSI-IV

La WPPSI-IV es un instrumento clínico de aplicación individual diseñado para medir la aptitud cognitiva de los niños de 2:6 a 7:7 años. Para las edades de 4:0 a 7:7 años ofrece puntuaciones de pruebas y puntuaciones compuestas que representan el funcionamiento intelectual de los dominios cognitivos siguientes: Comprensión verbal (ICV), Visoespacial (IVE), Razonamiento fluido (IRF), Memoria de trabajo (IMT) y Velocidad de procesamiento (IVP); así como el CI total (CIT), el Índice de capacidad general (ICG) y el Índice de competencia cognitiva (ICC).

Se aplicó la WPPSI-IV y la WISC-V a 105 niños de 6:0 a 7:7 años, en orden contrabalanceado, con un intervalo de aplicación de 7 a 57 días (media = 22 días). La tabla 6.6 muestra las medias, las desviaciones típicas, las correlaciones (corregidas y sin corregir) y las diferencias típicas de las pruebas, las puntuaciones de procesamiento y las puntuaciones compuestas.

Tabla 6.6. Correlaciones entre la WISC-V y la WPPSI-IV (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	WISC-V			WPPSI-IV					
	Media[a]	*Dt*[a]	*n*	Media[a]	*Dt*[a]	*n*	r_{12}[b]	r_{12}[b] corregido	Diferencia típica[c]
Semejanzas	10.7	2.7	103	10.9	2.5	103	.60	.64	0.08
Vocabulario	10.3	2.7	105	10.1	2.6	105	.60	.59	0.08
Información	10.5	2.5	104	11.0	2.2	104	.51	.63	0.21
Comprensión	10.2	2.8	105	8.9	2.3	105	.55	.62	0.51
Dibujos	—	—	—	10.8	2.7	105	—	—	—
Nombres	—	—	—	10.8	2.7	105	—	—	—
Cubos	10.6	2.7	105	10.7	2.8	105	.59	.67	0.04
Puzles visuales-Rompecabezas	10.5	2.8	104	10.8	2.7	104	.24	.24	0.11
Matrices	10.9	2.9	105	11.1	2.4	105	.44	.48	0.08
Balanzas	10.6	2.4	103	—	—	—	—	—	—
Aritmética	10.4	3.0	105	—	—	—	—	—	—
Dígitos	10.5	2.8	105	—	—	—	—	—	—
Span de dibujos-Reconocimiento	10.5	2.7	105	10.1	2.6	105	.31	.37	0.15
Letras y números	10.2	2.7	105	—	—	—	—	—	—
Localización	—	—	—	11.2	2.7	105	—	—	—
Claves-Clave de figuras	10.3	2.6	104	11.2	2.8	104	.44	.51	0.33
Búsqueda de símbolos-Búsqueda de animales	10.6	2.8	103	10.4	2.7	103	.47	.53	0.07
Cancelación	10.3	2.6	104	10.7	2.6	104	.55	.60	0.15
CAa	10.0	2.5	104	10.7	2.6	104	.49	.56	0.27
CAe	10.4	2.5	104	10.7	2.5	104	.53	.60	0.12
ICV	102.7	12.9	103	104.7	11.6	103	.64	.71	0.16
IVE	103.2	12.9	104	104.5	13.9	104	.57	.60	0.10
IRF	104.2	12.4	103	106.6	12.9	103	.45	.51	0.19
IMT	102.9	13.0	105	103.9	11.5	105	.46	.53	0.08
IVP	102.5	13.3	103	103.6	13.5	103	.34	.39	0.08
CIT	103.8	12.1	101	104.9	10.9	101	.74	.83	0.10
IMTA-IMT	102.0	13.3	105	103.9	11.5	105	.38	.44	0.15
INV	103.7	12.1	102	104.9	11.5	102	.64	.74	0.10
ICG	104.1	12.0	101	105.9	11.1	101	.74	.83	0.16
ICC	103.1	13.1	103	104.3	11.0	103	.48	.57	0.10

Nota: El IMTA se correlacionó con el IMT de la WPPSI-IV. Las correlaciones se calcularon por separado para cada orden de aplicación en un diseño contrabalanceado y se corrigieron de acuerdo con la variabilidad de la muestra de tipificación estadounidense de la WISC-V (Guilford y Fruchter, 1978).

[a] Los valores de las columnas «Media» y «*Dt*» son el promedio de las medias y desviaciones típicas de los dos órdenes de aplicación.

[b] La media del coeficiente de correlación de las dos aplicaciones se obtuvo mediante la transformación *z* de Fisher.

[c] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Las medias de los índices primarios y del CIT de la WISC-V se sitúan en el rango promedio. Las puntuaciones de los índices primarios de la WPPSI-IV y el CIT son consistentemente más altas que las puntuaciones compuestas correspondientes de la WISC-V. Sin embargo, las diferencias típicas son todas inferiores a .20, como era de esperar en instrumentos tipificados con una diferencia de dos años.

En cuanto a las pruebas, se observa una diferencia típica moderada entre las medias de *Comprensión*, así como tamaños de efecto pequeños en *Información* y *Claves-Clave de figuras*. Las puntuaciones de estas tres pruebas son superiores en la WPPSI-IV, excepto las de *Comprensión*.

Los coeficientes de correlación corregidos entre las puntuaciones de los índices primarios y el CIT de la WISC-V y las correspondientes de la WPPSI-IV son de moderados a altos, excepto el del IVP, que es ligeramente más bajo. Las escalas Visoespacial de la WISC-V y de la WPPSI-IV solo tienen en común la prueba *Cubos*, y las escalas Memoria de trabajo y Velocidad de procesamiento no tienen pruebas en común. El ICC y el ICG muestran correlaciones de moderadas a altas entre ambos instrumentos. Respecto a las pruebas, los coeficientes de correlación corregidos son de moderados a altos, a excepción de *Span de dibujos-Reconocimiento*, cuyo resultado es ligeramente inferior. No se incluye aquí el par de pruebas *Puzles visuales-Rompecabezas* porque, aunque guardan relación, están formadas por tareas distintas. Por lo general, las correlaciones más altas se observan en las pruebas de Comprensión verbal.

En conjunto, estos resultados son bastante consistentes con los observados entre la WISC-IV y la WPPSI-IV (Wechsler, 2012), así como en otros estudios comparativos de distintas ediciones de la WPPSI y la WISC (Wechsler, 1991, 2002, 2003). La magnitud de las correlaciones y la media consistente de las puntuaciones compuestas sugieren que ambos instrumentos miden constructos similares.

Correlaciones con la WAIS-IV

La WAIS-IV es un instrumento clínico de aplicación individual diseñado para medir la aptitud cognitiva de los adolescentes y adultos de 16 a 90 años. Se aplicó la WISC-V y la WAIS-IV a 112 sujetos de 16 años, en orden contrabalanceado y con un intervalo de aplicación de 12 a 67 días (media = 21 días). La tabla 6.7 muestra las medias, las desviaciones típicas, las correlaciones (corregidas y sin corregir) y las diferencias típicas de las pruebas, las puntuaciones de procesamiento y las puntuaciones compuestas.

Tabla 6.7. Correlaciones entre la WISC-V y la WAIS-IV (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	WISC-V			WAIS-IV					
	Media[a]	*Dt*[a]	*n*	Media[a]	*Dt*[a]	*n*	r_{12}[b]	r_{12}[b] corregido	Diferencia típica[c]
Semejanzas	10.6	2.8	111	11.2	3.2	111	.64	.66	0.20
Vocabulario	10.2	2.6	112	10.6	2.8	112	.74	.77	0.15
Información	10.0	2.9	109	10.1	2.9	109	.73	.74	0.03
Comprensión	10.4	2.9	112	10.9	3.1	112	.52	.54	0.17
Cubos	10.4	2.8	112	10.4	3.1	112	.78	.80	0.00
Puzles visuales	10.6	2.7	112	10.3	2.6	112	.57	.63	0.11
Matrices	10.1	2.7	112	10.3	2.2	112	.43	.50	0.08
Balanzas	10.5	2.6	112	11.4	3.1	112	.66	.69	0.31
Figuras incompletas	—	—	—	10.0	2.9	112	—	—	—
Aritmética	10.3	2.7	112	10.3	2.7	112	.70	.75	0.00
Dígitos	10.5	2.7	112	9.9	2.6	112	.76	.80	0.23
Span de dibujos	10.5	3.1	111	—	—	—	—	—	—
Letras y números	10.6	3.2	112	10.0	2.6	112	.58	.65	0.21
Claves	10.9	3.1	112	9.8	2.2	112	.69	.74	0.41
Búsqueda de símbolos	10.7	2.9	112	11.0	2.8	112	.61	.63	0.11
Cancelación	9.6	3.3	111	11.2	3.9	111	.47	.41	0.44
Cs	10.3	2.6	112	10.2	2.5	112	.62	.68	0.04
Dd	10.1	2.8	112	9.6	3.0	112	.60	.62	0.17
Di	10.2	2.6	112	10.3	2.5	112	.50	.54	0.04
Dc	10.8	3.1	112	10.1	2.9	112	.55	.56	0.23
ICV	102.2	13.4	111	103.2	14.0	111	.80	.83	0.07
IVE-IRP	102.7	13.5	112	101.6	12.9	112	.78	.83	0.08
IRF-IRP	101.6	13.3	112	101.6	12.9	112	.56	.62	0.00
IMT	102.7	14.5	111	100.4	12.7	111	.72	.76	0.17
IVP	104.6	14.4	112	102.0	12.1	112	.79	.83	0.20
CIT	103.2	13.0	111	102.3	11.8	111	.84	.89	0.07
IMTA-IMT	102.9	14.4	112	100.4	12.7	112	.71	.75	0.18
ICG	102.4	13.3	111	102.7	13.2	111	.78	.82	0.02
ICC	104.5	13.7	111	101.5	11.4	111	.81	.88	0.24

Nota: El IMTA se correlacionó con el IMT de la WAIS-IV. Las correlaciones se calcularon por separado para cada orden de aplicación en un diseño contrabalanceado y se corrigieron de acuerdo con la variabilidad de la muestra de tipificación estadounidense de la WISC-V (Guilford y Fruchter, 1978).

[a] Los valores de las columnas «Media» y «*Dt*» son el promedio de las medias y desviaciones típicas de los dos órdenes de aplicación.

[b] La media del coeficiente de correlación de las dos aplicaciones se obtuvo mediante la transformación *z* de Fisher.

[c] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Las medias de los índices primarios y del CIT de la WISC-V se sitúan en el rango promedio. Las puntuaciones del IVE, el IMT y el IVP son superiores a las medias de las puntuaciones compuestas correspondientes de la WAIS-IV; el IRF es el mismo y el ICV es inferior en la WISC-V. No obstante, las diferencias típicas no son superiores a .20. La media del CIT de la WISC-V es ligeramente más alta que la de la WAIS-IV. Esta tendencia de medias de puntuaciones compuestas más altas en la WISC-V respecto a la WAIS-IV en sujetos de 16 años contrasta con la magnitud predicha por el efecto Flynn.

Por lo que respecta a las pruebas, las diferencias típicas entre las medias son moderadas en *Claves* y *Cancelación*, y pequeñas en *Semejanzas, Balanzas, Dígitos* y *Letras y números*, siendo diversa la dirección de las diferencias. En el resto de pruebas, las diferencias típicas son insignificantes.

Los coeficientes de correlación corregidos entre las puntuaciones de los índices primarios y el CIT de ambos instrumentos son de moderados a altos, observándose el valor más bajo en la

correlación IRF-IRP. El IRF de la WISC-V tiene una sola prueba (*Matrices*) en común con el IRP de la WAIS-IV, formado por tres pruebas. Las correlaciones de las pruebas son, por lo general, de moderadas a altas. La correlación más baja se observa entre las dos versiones de *Cancelación*. El diseño de *Cancelación* de la WISC-V es relativamente distinto al de la WAIS-IV, que consiste únicamente en figuras abstractas dispuestas solo de forma estructurada. Estos resultados son consistentes con las correlaciones de la WISC-IV con la WAIS-IV (Wechsler, 2008) y con la WAIS-III (Wechsler, 2003). La magnitud de estas correlaciones sugiere que los dos instrumentos miden constructos muy similares.

Correlaciones con la KABC-II

La KABC-II es un instrumento clínico de aplicación individual diseñado para medir la aptitud cognitiva de los niños y adolescentes de 3 a 18 años. Ofrece dos puntuaciones globales, el Índice fluido-cristalizado (IFC), formado por pruebas de cinco dominios (Secuenc ial/*Gsm*, Simultáneo/*Gv*, Aprendizaje/*Glr*, Planificación/*Gf* y Conocimientos/*Gc*), y el Índice de procesamiento mental (IPM), formado por pruebas de cuatro de los cinco dominios, excluyendo las pruebas de Conocimientos/*Gc*. También proporciona un Índice no verbal (INV), diseñado para medir las aptitudes de los niños con pérdida de audición, limitada competencia en habla inglesa, trastorno del lenguaje receptivo de moderado a grave, trastorno del espectro autista y otras diferencias que puedan afectar a las puntuaciones verbales. Para cada uno de los cinco dominios se dispone de puntuaciones de los índices.

Se aplicó la WISC-V y la KABC-II a 89 niños de 6 a 16 años, en orden contrabalanceado y con un intervalo de aplicación de 14 a 70 días (media = 22 días). La tabla 6.8 presenta las medias, las desviaciones típicas y las correlaciones corregidas de las pruebas y las puntuaciones compuestas de la WISC-V y de la KABC-II.

Tabla 6.8. Correlaciones entre la WISC-V y la KABC-II (datos estadounidenses)

Prueba/ punt. compuesta de la WISC-V	Pruebas de la KABC-II															
	AT	SC	NR	GC	RO	AD	EV	VK	RB	TR	BC	WO	PR	HM	RD	RI
Semejanzas	.25	.22	.20	.28	.24	.35	.55	.60	.30	.20	.11	.25	.45	.11	.33	.51
Vocabulario	.35	.26	.40	.22	.14	.37	.62	.80	.49	.35	.19	.29	.32	.16	.50	.58
Información	.32	.03	.29	.23	.23	.39	.53	.64	.28	.34	.22	.21	.37	–.08	.29	.54
Comprensión	.22	.28	.30	.20	.15	.32	.37	.56	.31	.18	.20	.31	.34	.28	.31	.46
Cubos	.18	.15	.03	.10	.36	.25	.24	.51	.25	.66	.37	.15	.62	.21	.26	.48
Puzles visuales	.26	.20	.10	.33	.17	.33	.18	.50	.30	.53	.24	.16	.30	.09	.23	.33
Matrices	.26	.15	.09	.13	.22	.15	.25	.34	.31	.51	.34	.10	.57	.24	.27	.39
Balanzas	.31	.24	.16	.16	.15	.27	.30	.38	.18	.28	.12	.13	.37	.06	.15	.37
Aritmética	.21	.16	.31	.18	.11	.38	.26	.53	.44	.45	.53	.33	.46	.21	.38	.49
Dígitos	.26	.02	.64	–.11	.22	.15	.01	.36	.28	.26	.21	.48	.26	.31	.25	.29
Span de dibujos	.28	.06	.37	.06	–.01	.15	.25	.40	.45	.20	.19	.39	.36	.37	.38	.23
Letras y números	.32	–.13	.46	–.02	.07	.13	.05	.33	.40	.18	.38	.29	.23	.24	.37	.18
Claves	–.01	.06	–.04	–.14	.13	–.04	.09	.12	–.16	.31	.12	–.07	.07	.16	–.16	.10
Búsqueda de símbolos	–.18	.02	.07	–.12	.05	–.10	–.05	–.08	–.15	.24	.20	–.09	.11	.06	–.09	.03
Cancelación	–.11	.18	–.26	.09	.08	–.21	.16	–.10	–.09	.14	.01	–.25	.13	–.04	–.08	.08
ICV	.34	.27	.34	.28	.22	.41	.65	.75	.43	.32	.18	.28	.44	.15	.47	.61
IVE	.25	.22	.13	.26	.30	.35	.23	.59	.33	.69	.39	.20	.54	.22	.32	.51
IRF	.35	.24	.15	.18	.22	.26	.34	.45	.30	.49	.28	.14	.57	.18	.26	.47
IMT	.31	.06	.58	–.05	.14	.16	.15	.43	.43	.27	.24	.51	.36	.40	.38	.30
IVP	–.11	.04	.01	–.13	.10	–.10	.01	–.01	–.18	.29	.19	–.10	.09	.12	–.15	.06
CIT	.37	.26	.41	.13	.34	.39	.46	.70	.40	.56	.37	.35	.60	.35	.41	.66
IRC	.29	.24	.28	.19	.15	.36	.34	.52	.36	.43	.37	.27	.49	.15	.31	.49
IMTA	.34	–.07	.64	–.07	.18	.20	.03	.40	.39	.28	.33	.47	.30	.33	.35	.28
INV	.34	.26	.27	.18	.27	.30	.36	.58	.37	.64	.40	.26	.63	.37	.37	.55
ICG	.33	.28	.30	.22	.28	.42	.51	.68	.43	.50	.31	.27	.61	.25	.41	.66
ICC	.19	.06	.47	–.06	.13	.09	.14	.33	.21	.41	.33	.34	.38	.46	.18	.28
KABC-II																
Media[a]	10.3	11.3	10.7	10.0	11.0	9.9	10.1	10.6	10.5	10.5	10.9	10.7	11.4	10.3	9.8	10.3
Dt[a]	2.7	2.9	2.5	2.7	3.3	2.3	3.1	2.1	2.7	2.7	2.7	2.7	2.8	3.3	2.5	2.7
n	86	89	89	88	88	88	88	89	89	88	87	89	88	88	87	89

Nota: Las abreviaciones de la KABC-II son: AT = Atlantis (Atlántida), SC = Story Completion (Finalización de historias), NR = Number Recall (Repetición de números), GC = Gestalt Closure (Cierre gestáltico), RO = Rover (Rover), AD = Atlantis Delayed (Atlántida demorada), EV = Expressive Vocabulary (Vocabulario expresivo), VK = Verbal Knowledge (Conocimiento verbal), RB = Rebus (Jeroglífico), TR = Triangles (Triángulos), BC = Block Counting (Recuento de cubos), WO = Word Order (Orden de palabra), PR = Pattern Reasoning (Razonamiento de patrones), HM = Hand Movements (Movimientos de manos), RD = Rebus Delayed (Jeroglífico demorado), RI = Riddles (Adivinanzas). Las correlaciones se calcularon por separado para cada orden de aplicación en un diseño contrabalanceado y se corrigieron de acuerdo con la variabilidad de la muestra de tipificación estadounidense de la WISC-V (Guilford y Fruchter, 1978). La media del coeficiente de correlación de las dos aplicaciones se obtuvo mediante la transformación *z* de Fisher.

[a] Los valores de las columnas «Media» y «*Dt*» son el promedio de las medias y desviaciones típicas de los dos órdenes de aplicación.

* A pesar de no ser un test adaptado a la población española, se han traducido el nombre de las pruebas y de los índices para facilitar la lectura.

Tabla 6.8. Correlaciones entre la WISC-V y la KABC-II *(continuación)*

Prueba/ punt. compuesta de la WISC-V	Puntuaciones compuestas de la KABC-II								WISC-V		
	Secuencial/ *Gsm*	Simultáneo/ *Gv*	Aprendizaje/ *Glr*	Planificación/ *Gf*	Conocimientos/ *Gc*	IFC	IPM	INV	Media[a]	*Dt*[a]	*n*
Semejanzas	.26	.33	.32	.39	.60	.61	.54	.35	10.7	2.8	88
Vocabulario	.40	.27	.48	.38	.74	.71	.63	.38	10.1	2.7	88
Información	.28	.39	.34	.23	.64	.60	.51	.25	10.3	2.5	89
Comprensión	.35	.26	.31	.38	.56	.60	.54	.41	10.4	2.9	89
Cubos	.11	.57	.25	.50	.51	.62	.60	.59	10.2	2.2	88
Puzles visuales	.15	.35	.31	.37	.40	.49	.48	.36	10.0	2.5	88
Matrices	.12	.42	.32	.45	.38	.54	.55	.52	10.3	2.7	89
Balanzas	.17	.25	.28	.37	.45	.47	.43	.29	9.9	2.8	89
Aritmética	.37	.36	.37	.39	.53	.63	.62	.53	10.3	2.7	89
Dígitos	.65	.35	.32	.16	.34	.58	.60	.33	10.0	2.8	88
Span de dibujos	.44	.11	.41	.24	.33	.50	.52	.38	9.9	2.8	89
Letras y números	.43	.19	.43	.08	.26	.41	.44	.28	9.8	2.6	89
Claves	–.07	.23	–.12	.09	.12	.08	.07	.21	10.4	2.4	88
Búsqueda de símbolos	–.02	.20	–.20	.03	.02	.01	.02	.17	10.3	2.7	88
Cancelación	–.29	.06	–.13	.19	.02	–.04	–.04	.10	10.5	2.7	88
ICV	.36	.34	.45	.44	.74	.72	.64	.41	102.3	13.6	87
IVE	.20	.53	.33	.51	.55	.65	.64	.60	100.7	11.0	87
IRF	.17	.41	.38	.50	.52	.63	.60	.49	100.4	13.4	89
IMT	.63	.27	.43	.23	.38	.63	.65	.41	99.6	14.2	88
IVP	–.06	.25	–.19	.05	.06	.04	.04	.21	102.0	12.6	87
CIT	.44	.55	.44	.54	.72	.81	.77	.64	101.7	12.2	84
IRC	.31	.35	.37	.44	.55	.63	.61	.48	100.4	13.6	89
IMTA	.64	.32	.43	.15	.35	.58	.60	.36	99.2	13.1	88
INV	.31	.51	.41	.57	.58	.73	.73	.68	100.9	12.0	86
ICG	.34	.47	.44	.56	.71	.78	.72	.60	101.8	13.0	86
ICC	.46	.34	.22	.25	.33	.51	.54	.49	100.6	11.7	86
KABC-II											
Media[a]	104.2	105.3	102.3	107.5	102.4	105.4	106.1	106.3			
Dt[a]	13.1	14.5	13.2	14.2	11.6	11.7	11.8	13.9			
n	89	86	86	84	89	86	86	89			

Nota: Las abreviaciones de la KABC-II son: IFC = Índice fluido-cristalizado, IPM = Índice de procesamiento mental, INV = Índice no verbal. Las correlaciones se calcularon por separado para cada orden de aplicación en un diseño contrabalanceado y se corrigieron de acuerdo con la variabilidad de la muestra de tipificación estadounidense de la WISC-V (Guilford y Fruchter, 1978). La media del coeficiente de correlación de las dos aplicaciones se obtuvo mediante la transformación *z* de Fisher.

[a] Los valores de las columnas «Media» y «*Dt*» son el promedio de las medias y desviaciones típicas de los dos órdenes de aplicación.

* A pesar de no ser un test adaptado a la población española, se han traducido el nombre de las pruebas y de los índices para facilitar la lectura.

Las medias de los índices primarios de la WISC-V se sitúan en el rango promedio. Las puntuaciones compuestas de la KABC-II son consistentemente superiores a las de la WISC-V, lo cual es coherente con el efecto Flynn.

Las correlaciones corregidas del CIT de la WISC-V con el IFC y el IPM son altas, y las de las puntuaciones de los índices primarios de la WISC-V con las puntuaciones compuestas correspondientes de la KABC-II son moderadas. El ICV y el IMT de la WISC-V, respectivamente, correlacionan en mayor medida con las escalas Conocimientos/*Gc* y Secuencial/*Gsm* de la KABC-II. El IVE de la WISC-V muestra correlaciones similares con las escalas Simultáneo/*Gv*, Planificación/*Gf* y Conocimientos/*Gc* de la KABC-II, lo que refleja que estas escalas miden varias habilidades mediante estímulos visuales. El IRF de la WISC-V muestra correlaciones altas con las escalas Conocimientos/*Gc* y Planificación/*Gf* de la KABC-II; probablemente esto se deba a que la prueba *Adivinanzas* de la KABC-II, pese a que en ella también interviene el razonamiento fluido, contribuye en la escala Conocimientos/*Gc*. El IVP de la WISC-V, como se esperaba, no muestra correlaciones

altas con ninguna de las escalas de la KABC-II, dado que explícitamente no mide la velocidad de procesamiento. Estas correlaciones demuestran la validez convergente de las nuevas escalas.

En cuanto a las puntuaciones de los índices secundarios, las correlaciones son las esperadas. El INV de la WISC-V correlaciona en mayor medida con el INV de la KABC-II, lo que ofrece una sólida evidencia de la validez convergente de este nuevo índice secundario. El IMTA de la WISC-V muestra correlaciones más altas con la escala Secuencial/*Gsm* de la KABC-II, lo que también demuestra la validez convergente de este índice como medida de la memoria de trabajo. Y, finalmente, el ICG y el ICC de la WISC-V, respectivamente, correlacionan en mayor medida con el IFC y el IPM de la KABC-II.

Estos resultados son coherentes con los de estudios anteriores sobre distintas versiones de la WISC y de la KABC (Kaufman y Kaufman, 2004). La magnitud de las correlaciones y la media consistente de las puntuaciones compuestas sugieren que ambos instrumentos miden constructos similares.

Correlaciones con el Vineland-II

El Vineland-II es una escala de aplicación individual, con un rango de edad que abarca desde el nacimiento hasta los 90 años, y que mide la conducta adaptativa en cuatro amplios dominios: Comunicación, Habilidades de la vida diaria, Socialización y Habilidades motoras. Además incluye el dominio Comportamientos desadaptativos, que mide problemas conductuales. Las puntuaciones de los dominios del Vineland-II tienen una media de 100 y una desviación típica de 15, pero los subdominios y las subescalas se miden mediante puntuaciones *v*, con una media de 15 y una desviación típica de 3. Las puntuaciones de los dominios comprendidas entre 71 y 85 reflejan niveles de habilidades adaptativas moderadamente bajos, y las puntuaciones iguales o inferiores a 70 reflejan niveles bajos. Asimismo, las puntuaciones *v* de los subdominios y subescalas con valores entre 10 y 12 se consideran moderadamente bajos, y los valores iguales o inferiores a 9 se consideran bajos.

Con las subescalas Internalizante y Externalizante y el correspondiente Índice de comportamientos desadaptativos, de puntuación inversa, las puntuaciones *v* de 18 a 20 se consideran altas y las de 21 a 24 se consideran clínicamente significativas.

Se aplicó el Formulario para padres y cuidadores a 61 padres de niños de 6 a 16 años, a los que se les había aplicado la WISC-V, con un intervalo de aplicación de 0 a 53 días (media = 8 días). La tabla 6.9 presenta las medias, las desviaciones típicas y las correlaciones corregidas de las pruebas y las puntuaciones compuestas de la WISC-V y de las pruebas/subdominios y puntuaciones compuestas de la Vineland-II. Todos los niños de este estudio pertenecían a una población no clínica.

Dado que el dominio Habilidades motoras, y sus subdominios, miden el desarrollo, se da una pequeña variabilidad en el rendimiento de los sujetos de entre 7 y 50 años y no se dispone de baremos para este rango de edad. Asimismo, no se ofrecen datos sobre las puntuaciones de Habilidades motoras en los niños de 6 años porque la muestra es demasiado pequeña. No obstante, cuando se disponía de una puntuación *v* del dominio Habilidades motoras, se obtuvo la Puntuación compuesta Conducta adaptativa mediante la suma de las cuatro puntuaciones típicas.

Las medias de los índices primarios y del CIT de la WISC-V se sitúan en el rango promedio, al igual que las medias de los dominios del Vineland-II.

La correlación global entre los dos instrumentos es baja, situándose la correlación entre el CIT y la Puntuación compuesta Conducta adaptativa próxima al cero. La Puntuación compuesta Conducta adaptativa muestra correlaciones de pequeñas a insignificantes con todas las puntuaciones compuestas de la WISC-V.

El dominio Comunicación, que requiere una mayor aptitud cognitiva que el resto de dominios de funcionamiento adaptativo, muestra las correlaciones más altas con las puntuaciones de los índices primarios de la WISC-V, observándose la mayor correlación con el IMT. El ICV y el IRF muestran correlaciones bajas con el dominio Comunicación.

De los subdominios para los que se presentan datos, el que muestra las correlaciones más altas con las puntuaciones de los índices primarios y secundarios de la WISC-V es el de Escrita. Respecto a las puntuaciones de los índices primarios, se observan correlaciones de bajas a moderadas con el

Tabla 6.9. Correlaciones entre la WISC-V y el Vineland-II (datos estadounidenses)

	Subdominio del Vineland-II										
	Comunicación			Habilidades de la vida diaria			Socialización			Comportamiento	
Prueba/ punt. compuesta de la WISC-V	**Receptiva**	**Expresiva**	**Escrita**	**Personales**	**Domésticas**	**En comunidad**	**Relaciones interpersonales**	**Juego y ratos libres**	**Habilidades de afrontamiento**	**Internalizante**	**Externalizante**
Semejanzas	–.10	.22	.34	.05	–.14	.09	.03	.05	–.09	.03	.05
Vocabulario	–.04	.14	.32	.06	–.06	.17	.00	.15	–.15	–.04	.08
Información	–.14	.17	.38	.11	–.09	.21	.08	.21	–.14	–.13	–.11
Comprensión	.04	.20	.14	.15	.01	.20	.11	.16	.07	.04	–.05
Cubos	–.12	–.02	.17	.16	–.21	–.05	–.20	–.26	–.33	.03	–.02
Puzles visuales	–.11	.03	.33	.13	–.02	.11	–.15	–.14	–.24	.01	.08
Matrices	.01	.35	.58	.07	.02	.23	.02	.02	–.17	–.15	.01
Balanzas	–.14	.04	.34	–.07	–.18	–.10	–.08	–.15	–.25	.00	.17
Aritmética	–.01	–.02	.20	.22	–.16	.00	.10	.07	–.20	–.07	.01
Dígitos	.00	.35	.62	–.02	.01	.28	.23	.20	.08	–.33	–.21
Span de dibujos	–.05	.13	.38	–.04	–.13	.07	.02	.10	–.10	–.12	.00
Letras y números	–.02	.09	.37	.04	–.02	.16	.05	.21	–.06	–.23	–.04
Claves	–.02	.10	–.03	.15	–.07	.12	.01	.03	–.03	–.16	–.18
Búsqueda de símbolos	–.03	–.04	.08	.13	–.12	.00	–.20	–.13	–.20	.01	–.04
Cancelación	.01	.12	.10	.05	.01	.10	.27	.02	.05	–.21	–.18
ICV	–.08	.19	.34	.06	–.10	.13	.01	.10	–.13	–.01	.07
IVE	–.13	.01	.29	.16	–.12	.04	–.18	–.21	–.31	.02	.05
IRF	–.08	.18	.47	–.01	–.11	.04	–.04	–.08	–.24	–.07	.12
IMT	–.01	.27	.55	–.03	–.06	.20	.13	.18	–.01	–.25	–.12
IVP	–.03	.02	.04	.17	–.11	.07	–.12	–.07	–.15	–.07	–.12
CIT	–.08	.21	.44	.07	–.12	.13	.01	.02	–.18	–.11	–.01
IRC	–.11	.03	.33	.05	–.21	–.07	–.01	–.07	–.27	–.04	.13
IMTA	–.01	.22	.52	.01	.00	.23	.13	.22	.00	–.29	–.11
INV	–.10	.13	.40	.08	–.13	.08	–.07	–.08	–.24	–.08	.03
ICG	–.08	.16	.41	.06	–.12	.08	–.04	–.03	–.22	–.03	.07
ICC	–.05	.20	.41	.08	–.12	.18	.01	.08	–.11	–.22	–.16
Vineland-II											
Media	15.6	15.5	15.2	15.8	15.2	15.6	15.1	14.7	17.1	15.9	15.1
Dt	2.5	2.5	2.8	2.9	3.0	2.7	2.5	2.7	3.7	2.7	2.8
n	60	61	61	60	61	61	61	61	58	60	60

Nota: Se corrigieron las correlaciones de acuerdo con la variabilidad de la muestra de tipificación estadounidense de la WISC-V (Guilford y Fruchter, 1978).

* A pesar de no ser un test adaptado a la población española, se han traducido el nombre de las pruebas y de los índices para facilitar la lectura.

ICV, el IMT y el IRF. Y respecto a las puntuaciones de los índices secundarios, la correlación más alta se observa entre el IMTA y Escrita. Las correlaciones entre el resto de los subdominios y las puntuaciones compuestas de la WISC-V son todas de muy bajas a insignificantes.

Tabla 6.9. Correlaciones entre la WISC-V y el Vineland-II *(continuación)*

	Dominio del Vineland-II					WISC-V		
Prueba/ punt.compuesta de la WISC-V	**Comunicación**	**Habilidades de la vida diaria**	**Socialización**	**Puntuación compuesta Conducta adaptativa**	**Índice de comportamientos desadaptativos**	**Media**	***Dt***	***n***
Semejanzas	.19	–.01	–.07	.01	.09	10.3	3.0	61
Vocabulario	.20	.08	–.04	.07	.04	9.8	3.2	60
Información	.18	.09	–.02	.06	–.07	10.2	2.9	61
Comprensión	.19	.17	.11	.21	.03	9.8	3.4	61
Cubos	.00	–.05	–.36	–.23	.03	10.3	2.6	61
Puzles visuales	.13	.09	–.24	–.02	.06	9.9	3.0	61
Matrices	.46	.12	–.09	.13	–.08	9.8	2.6	61
Balanzas	.14	–.15	–.24	–.15	.14	9.5	3.1	61
Aritmética	.07	–.02	–.06	–.05	–.01	9.4	2.4	61
Dígitos	.49	.06	.17	.25	–.25	9.7	2.7	61
Span de dibujos	.24	–.08	.02	.07	–.03	9.7	3.0	61
Letras y números	.23	.06	.08	.14	–.15	9.6	3.1	61
Claves	–.01	.10	.01	.01	–.23	10.2	2.6	61
Búsqueda de símbolos	–.02	.02	–.25	–.13	–.05	9.9	3.0	61
Cancelación	.09	.08	.11	.08	–.21	9.4	3.0	61
ICV	.21	.04	–.06	.04	.07	100.3	16.0	60
IVE	.09	.03	–.33	–.12	.06	100.5	14.1	61
IRF	.28	–.05	–.20	–.04	.05	97.7	15.3	61
IMT	.41	–.01	.11	.19	–.16	98.2	14.5	61
IVP	–.02	.07	–.16	–.08	–.16	100.2	13.4	61
CIT	.27	.02	–.12	.01	–.04	99.4	15.3	60
IRC	.14	–.12	–.21	–.12	.09	96.8	13.8	61
IMTA	.37	.06	.12	.19	–.20	98.2	14.8	61
INV	.21	–.01	–.19	–.04	–.01	99.1	14.9	61
ICG	.23	.00	–.17	–.03	.06	99.4	16.0	60
ICC	.27	.03	–.03	.07	–.20	98.9	13.3	61
Vineland-II								
Media	104.1	103.6	104.4	103.9	15.5			
Dt	13.8	14.5	15.4	13.9	2.7			
n	61	60	60	60	60			

Nota: Se corrigieron las correlaciones de acuerdo con la variabilidad de la muestra de tipificación estadounidense de la WISC-V (Guilford y Fruchter, 1978).

* A pesar de no ser un test adaptado a la población española, se han traducido el nombre de las pruebas y de los índices para facilitar la lectura.

Correlaciones con el BASC-2 CPP

El Cuestionario para padres del BASC-2 (BASC-2 CPP) es una escala en la que los padres indican la frecuencia en que observan determinadas conductas en su hijo. Forma parte del BASC-2, un sistema multimétodo y multidimensional de evaluación para las edades de 2 a 25 años, que se utiliza en la realización de diagnósticos diferenciales y en la clasificación educativa de trastornos emocionales y de la conducta.

En este estudio comparativo con la WISC-V se seleccionaron dos escalas primarias y dos escalas de contenido del BASC-2 CPP con el objetivo de obtener más información sobre las relaciones entre la aptitud cognitiva y las dimensiones conductuales medidas por las escalas seleccionadas. Las escalas primarias del BASC-2 CPP utilizadas fueron Problemas de atención (tendencia a la distracción y falta de capacidad para mantener la concentración) y Problemas de conducta (tendencia a las conductas socialmente infrecuentes y perturbadoras). Y las escalas de contenido utilizadas fueron Funcionamiento ejecutivo (aptitud para controlar la conducta mediante la pla-

nificación, la anticipación, la inhibición o la realización de una actividad dirigida, de forma que la reacción sea adecuada al entorno) y Resiliencia (aptitud para acceder a sistemas de ayuda interna y externa para mitigar el estrés y sobreponerse a las adversidades).

Las puntuaciones principales y de contenido tienen una media de 50 y una desviación típica de 10. Las puntuaciones más altas indican más problemas conductuales en la mayoría de las escalas: las puntuaciones entre 60 y 69 indican que se da una situación de riesgo, y las puntuaciones iguales o superiores a 70 se consideran clínicamente significativas. Aunque el funcionamiento ejecutivo es un rasgo positivo, las puntuaciones más altas en la escala homónima indican que el funcionamiento ejecutivo es más bajo. Respecto a la escala Resiliencia, las puntuaciones altas reflejan características positivas, mientras que las puntuaciones bajas sugieren la posibilidad de áreas problemáticas.

Se aplicó la WISC-V y las escalas seleccionadas del BASC-2 CPP a 2302 niños no clínicos de 6 a 16 años, así como a sus cuidadores, con un intervalo de aplicación de 0 a 126 días (media = 1 día). El 89 % de la muestra completaron el BASC-2 el mismo día en que los niños completaron la WISC-V. La tabla 6.10 presenta las medias, las desviaciones típicas y las correlaciones corregidas de la WISC-V y del BASC-2 CPP.

Las medias de las puntuaciones compuestas de la WISC-V y la media de las puntuaciones del BASC-2 CPP se sitúan en el rango promedio. Las correlaciones entre la WISC-V y las puntuaciones seleccionadas del BASC-2 CPP son negativas y bajas. Las mayores correlaciones, aun siendo bajas, se observan con la escala Problemas de atención del BASC-2 CPP. En cuanto a las pruebas de la WISC-V, *Letras y números* muestra la correlación más alta con Problemas de atención.

Tabla 6.10. Correlaciones entre la WISC-V y el BASC-2 CPP (datos estadounidenses)

Prueba/puntuación compuesta de la WISC-V	Escala del BASC-2 CPP				WISC-V		
	Resiliencia	Problemas de conducta	Funcionamiento ejecutivo	Problemas de atención	Media	*Dt*	*n*
Semejanzas	.09	−.14	−.10	−.13	10.4	2.9	2297
Vocabulario	.09	−.13	−.11	−.14	10.3	3.0	2298
Información	.08	−.09	−.09	−.11	10.4	2.9	2301
Comprensión	.07	−.11	−.07	−.11	10.4	2.9	2302
Cubos	.06	−.07	−.05	−.07	10.3	2.9	2302
Puzles visuales	.07	−.10	−.07	−.07	10.3	2.9	2302
Matrices	.07	−.10	−.09	−.14	10.3	2.9	2301
Balanzas	.06	−.12	−.07	−.11	10.2	2.9	2301
Aritmética	.09	−.13	−.13	−.17	10.4	2.9	2302
Dígitos	.09	−.10	−.09	−.16	10.4	2.9	2296
Span de dibujos	.07	−.10	−.06	−.11	10.3	2.9	2296
Letras y números	.11	−.13	−.13	−.20	10.4	2.9	2298
Claves	.05	−.08	−.08	−.16	10.3	2.9	2299
Búsqueda de símbolos	.07	−.06	−.06	−.10	10.2	2.9	2292
Cancelación	.05	−.01	−.02	−.04	10.2	3.0	2301
ICV	.10	−.16	−.11	−.16	102.2	14.5	2293
IVE	.07	−.10	−.07	−.08	101.7	14.6	2302
IRF	.08	−.13	−.10	−.15	101.6	14.6	2300
IMT	.10	−.12	−.08	−.16	101.9	14.3	2290
IVP	.07	−.08	−.08	−.15	101.4	14.4	2289
CIT	.12	−.17	−.14	−.20	102.3	14.1	2282
IRC	.09	−.15	−.11	−.16	101.7	14.4	2301
IMTA	.11	−.14	−.12	−.20	102.2	14.4	2293
INV	.09	−.15	−.12	−.17	102.0	14.3	2291
ICG	.10	−.16	−.12	−.17	102.2	14.4	2291
ICC	.11	−.13	−.11	−.19	101.9	14.1	2278
BASC-2 CPP							
Media	49.6	48.6	49.1	50.8			
Dt	8.4	8.5	8.4	8.5			
n	2223	2213	2222	2240			

Nota: Se corrigieron las correlaciones de acuerdo con la variabilidad de la muestra de tipificación estadounidense de la WISC-V (Guilford y Fruchter, 1978).

Grupos especiales

Los resultados de los estudios realizados con grupos especiales proporcionan pruebas adicionales de validez, basadas en las relaciones entre el test y diversos criterios. La demostración de la validez de una escala, aplicada a grupos clínicos y especiales, es fundamental cuando sus resultados forman parte de una evaluación diagnóstica exhaustiva. Durante las fases de tipificación estadounidense y española de la WISC-V se llevaron a cabo varios estudios con grupos especiales a fin de determinar si los constructos medidos por la escala resultaban como se esperaba en determinados grupos con características definidas.

Un equipo de examinadores e investigadores recopilaron los datos para los estudios con grupos especiales. Los sujetos que participaron en estos estudios se seleccionaron en distintos ámbitos clínicos y fueron aceptados para participar en la muestra de grupos especiales de acuerdo con ciertos criterios de inclusión específicos (en el anexo B de este manual figuran los criterios seguidos para cada uno de los grupos especiales).

Es importante señalar las limitaciones de estos estudios. En primer lugar, cabe destacar que las muestras no se seleccionaron aleatoriamente, sino en función de la disponibilidad. Por ello, es posible que estos estudios no sean representativos del rendimiento en la WISC-V de todos los

niños pertenecientes a una categoría diagnóstica. En segundo lugar, dado que los datos de cada uno de los grupos especiales se recogieron en diferentes ámbitos clínicos, los diagnósticos de los niños del mismo grupo pueden haberse basado en criterios y procedimientos diferentes, en especial por lo que respecta al grupo con trastornos del aprendizaje y al grupo con comportamiento disruptivo. Además, el tamaño de las muestras de algunos de estos grupos es pequeño y cubre solo una parte del rango de edad de la WISC-V. Finalmente, solo se describe el rendimiento del grupo. Por todo ello, los datos de estas muestras se presentan a modo de ejemplo, y no pretenden ser totalmente representativos de los grupos diagnósticos. El objetivo de estos estudios es demostrar que la WISC-V puede ofrecer estimaciones válidas de la aptitud intelectual de los niños de estos grupos especiales, y que los niños que presentan trastornos específicos obtienen las puntuaciones esperadas. Sin embargo, las puntuaciones de la WISC-V no deberían usarse nunca como el único criterio para establecer un diagnóstico o una clasificación.

Los datos presentados en esta sección demuestran la sensibilidad diferencial de la WISC-V a los déficits cognitivos generales y específicos mostrados por los niños que más frecuentemente son evaluados en diferentes ámbitos clínicos. Asimismo, demuestra la utilidad clínica y la validez discriminante de las pruebas y las puntuaciones compuestas de la WISC-V. Queda fuera del alcance de este capítulo ofrecer un análisis completo de la literatura relevante sobre cada uno de los grupos clínicos. Cada una de las tablas muestra la media y la desviación típica de cada grupo, así como las diferencias, incluido el tamaño del efecto. Aunque las medias del grupo se presentan en las tablas expresadas hasta el primer decimal, las diferencias de puntuaciones, los valores *t* y las probabilidades (*p*) se calcularon teniendo en cuenta el segundo decimal. Por este motivo, los cálculos realizados a partir de estas medias pueden variar ligeramente de las mostradas en las tablas debido al redondeo. Las comparaciones entre los grupos se calcularon a partir de la muestra de tipificación de la WISC-V, teniendo en cuenta la edad, el sexo y el nivel educativo de los padres. Luego se seleccionaron aleatoriamente los casos, y se emparejó la muestra obtenida con el grupo clínico según las variables definidas.

Las características demográficas de las muestras de los grupos especiales de la WISC-V se presentan en la tabla 6.11. Todos los grupos control, seleccionados aleatoriamente de la muestra de tipificación, se clasificaron en función de estas variables demográficas. Las tablas 6.12-6.23 presentan las medias de las puntuaciones de la WISC-V para los grupos especiales y para sus correspondientes grupos control.

Tabla 6.11. Datos demográficos de los grupos especiales

	ACI*	ACI	DI*	DIL	DIM	CIL	TA-L	TA-LEE	TA-M	TDAH	CD	TCE	TEA-L	TEA-NL
N	36	95	20	74	37	20	30	22	28	48	21	20	30	32
Edad														
Media	10.7	11.1	10.8	12.1	11.8	12.6	11.5	11.2	14.0	11.4	10.6	12.9	11.5	11.5
Dt	2.7	3.1	3.1	3.0	3.4	3.6	2.5	2.5	2.2	2.9	3.6	3.0	3.0	2.6
Sexo														
Mujer	44.4	34.7	60.0	44.6	51.4	70.0	56.7	31.8	50.0	37.5	47.6	40.0	23.3	25.0
Varón	55.6	65.3	40.0	55.4	48.6	30.0	43.3	68.2	50.0	62.5	52.4	60.0	76.7	75.0

Nota: Las abreviaciones de los grupos especiales son: ACI = alta capacidad intelectual, DI = discapacidad intelectual, DIL = discapacidad intelectual leve, DIM = discapacidad intelectual moderada, CIL = capacidad intelectual límite, TA-L = trastorno específico del aprendizaje-lectura, TA-LEE = trastorno específico del aprendizaje-lectura y expresión escrita, TA-M = trastorno específico del aprendizaje-matemáticas, TDAH = trastorno por déficit de atención/hiperactividad, CD = comportamiento disruptivo, TCE = traumatismo craneoencefálico, TEA-L = trastorno del espectro autista con deterioro de lenguaje, TEA-NL = trastorno del espectro autista sin deterioro de lenguaje. Excepto para el tamaño de la muestra (*n*) y la edad, los datos se muestran en porcentajes. Si el porcentaje total no suma 100, es debido al redondeo.

*Datos españoles.

Niños con alta capacidad intelectual

Los niños con alta capacidad intelectual muestran una gran aptitud en medidas del funcionamiento intelectual, la flexibilidad cognitiva, la creatividad y/u otras áreas específicas (Geake, 2008; Koziol, Budding y Chidekel, 2010; Munro, 2013; Valdés, Vera y Carlos, 2013). Las medias de las puntuaciones obtenidas por estos niños son significativamente más altas que las obtenidas por niños de la población general en las mismas medidas de aptitud cognitiva (Rimm et al., 2008; Rowe, Kingsley y Thompson, 2010; Sweetland, Reina y Tatti, 2006; Wechsler, 2002, 2003). Es más, estos niños tienden a obtener resultados más altos en todas las puntuaciones compuestas que los niños de su misma edad.

Aunque las personas con alta capacidad intelectual muestran un buen rendimiento en las pruebas de las escalas de inteligencia tradicionales de Wechsler, en algunas ocasiones poco frecuentes pueden mostrar grandes divergencias entre las puntuaciones obtenidas en pruebas verbales y pruebas no verbales (Sweetland et al., 2006). Normalmente, los puntos fuertes de los niños con alta capacidad intelectual se reflejan en las áreas de comprensión verbal, aptitud visoespacial y razonamiento fluido. Su rendimiento en memoria de trabajo y velocidad de procesamiento también es, por lo general, superior al de la población general (Elliot, 2007; Kaufman y Kaufman, 2004; Wechsler, 2003, 2008, 2012), aunque no suele ser tan alto como en las tres primeras áreas citadas (Raiford et al., 2005; Rimm et al., 2008; Rowe et al., 2010). El rendimiento relativamente alto de estos niños en memoria de trabajo parece favorecer la creatividad, que es un aspecto importante de la alta capacidad intelectual (Geake, 2008).

Puesto que está demostrado que los niños con alta capacidad intelectual muestran puntos débiles intraindividuales en las áreas de la memoria de trabajo y la velocidad de procesamiento, algunos autores recomiendan utilizar el CI total y el ICG cuando se usan las puntuaciones de la aptitud cognitiva como criterio para identificar la alta capacidad intelectual en estos niños (Rimm et al., 2008). No obstante, el CI total parece ser un mejor indicador del rendimiento académico, incluidas las matemáticas y la comprensión lectora, de los niños con alta capacidad intelectual (Rowe, Miller, Ebenstein y Thompson, 2012). De entre los índices de la WISC-IV, el ICV y el IMT son indicadores sólidos del rendimiento académico de estos niños (Rowe et al., 2010).

Los resultados obtenidos en estudios sobre la alta capacidad intelectual pueden verse afectados por diversos factores. En primer lugar, las distintas medidas de aptitud cognitiva aportan estimaciones altamente correlacionadas, pero diferentes. En segundo lugar, los criterios para identificar a los niños con alta capacidad intelectual varían de un lugar a otro (Pfeiffer, 2013). En tercer lugar, los niños que participan en programas especializados en alta capacidad intelectual son admitidos, a veces, en función de las puntuaciones que obtienen en medidas del rendimiento, y no siempre las puntuaciones altas en estas medidas se corresponden con una aptitud cognitiva alta.

Durante el estudio de validez español, se aplicó la WISC-V a 36 niños de entre 6 y 16 años, que habían sido previamente identificados como niños con alta capacidad intelectual. Para poder formar parte de este estudio, los niños tenían que haber obtenido una puntuación Escala total por encima de 2 desviaciones típicas en alguna prueba estandarizada de aptitud intelectual general, aplicada individualmente, y estar recibiendo atención específica en la escuela. La tabla 6.12 presenta las medias de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas de la WISC-V, para el grupo de niños con alta capacidad intelectual y para el grupo control correspondiente.

Tabla 6.12. Resultados del grupo con alta capacidad intelectual y del grupo control (datos españoles)

Prueba/ punt. de procesamiento/ punt. compuesta	Alta capacidad intelectual		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	12.5	2.5	10.8	2.0	36	−1.72	−3.80	<.01	−0.76
Vocabulario	13.5	3.2	11.1	2.5	36	−2.39	−4.13	<.01	−0.83
Información	13.7	2.9	10.7	2.8	36	−3.06	−5.21	<.01	−1.07
Comprensión	12.6	2.8	10.9	2.9	36	−1.72	−2.53	.02	−0.60
Cubos	13.4	2.4	11.2	1.9	36	−2.19	−4.26	<.01	−1.01
Puzles visuales	13.9	2.8	11.4	2.5	36	−2.47	−4.35	<.01	−0.93
Matrices	13.8	2.5	11.4	2.9	36	−2.39	−3.75	<.01	−0.88
Balanzas	13.6	2.6	10.8	2.9	36	−2.81	−4.88	<.01	−1.02
Aritmética	13.8	2.5	10.6	2.5	36	−3.19	−5.96	<.01	−1.28
Dígitos	13.7	2.8	10.6	2.5	36	−3.11	−5.38	<.01	−1.17
Span de dibujos	13.5	2.6	10.6	3.0	36	−2.86	−4.72	<.01	−1.02
Letras y números	12.7	2.5	10.8	2.5	36	−1.97	−3.62	<.01	−0.79
Claves	12.7	2.6	10.9	2.4	36	−1.81	−3.14	<.01	−0.72
Búsqueda de símbolos	12.4	2.5	11.0	2.6	36	−1.36	−2.10	.04	−0.53
Cancelación	11.1	2.3	10.8	2.8	36	−0.36	−0.52	.60	−0.14
Cs	13.4	2.5	11.3	1.9	36	−2.06	−3.84	<.01	−0.93
Cp	13.2	2.7	10.1	1.9	15	−3.13	−3.33	<.01	−1.34
Dd	12.0	2.6	10.2	2.0	36	−1.83	−3.43	<.01	−0.79
Di	12.4	2.8	10.3	2.8	36	−2.11	−3.08	<.01	−0.75
Dc	13.9	3.0	10.8	3.0	36	−3.06	−4.78	<.01	−1.02
CAa	10.8	2.1	10.9	2.7	36	0.14	0.22	.82	0.06
CAe	11.5	2.7	10.6	2.7	36	−0.89	−1.26	.22	−0.33
ICV	116.4	13.6	105.4	10.4	36	−11.06	−5.01	<.01	−0.91
IVE	120.7	13.3	107.1	11.1	36	−13.67	−4.90	<.01	−1.12
IRF	121.2	12.6	106.4	13.1	36	−14.78	−5.37	<.01	−1.15
IMT	120.0	12.4	103.5	12.1	36	−16.50	−7.11	<.01	−1.35
IVP	114.4	10.5	105.5	12.8	36	−8.94	−3.10	<.01	−0.76
CIT	123.6	11.1	107.0	10.5	36	−16.61	−7.27	<.01	−1.54
IRC	121.5	12.2	103.8	13.5	36	−17.72	−6.36	<.01	−1.38
IMTA	117.5	12.7	103.6	12.1	36	−13.94	−5.31	<.01	−1.12
INV	125.3	12.0	107.7	10.7	36	−17.58	−6.73	<.01	−1.55
ICG	121.8	12.2	107.1	10.0	36	−14.72	−6.60	<.01	−1.32
ICC	120.9	9.9	105.5	12.3	36	−15.42	−6.28	<.01	−1.38

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Las medias de las puntuaciones de los índices primarios y del CI total obtenidas por el grupo con alta capacidad intelectual son significativamente superiores a las del grupo control. En el grupo con alta capacidad intelectual las medias de los índices primarios oscilan entre 114.4 (IVP) y 121.2 (IRF), y la media del CI total es de 123.6. Los tamaños del efecto para las diferencias de las puntuaciones de los índices primarios de dichos grupos son grandes, siendo moderado el del IVP. La media del IVP es más baja que las medias del resto de las puntuaciones compuestas; de hecho, las medias más bajas en el grupo de alta capacidad intelectual se obtuvieron en las pruebas *Cancelación* y *Búsqueda de símbolos*. Las puntuaciones de los índices secundarios oscilan entre 117.5 (IMTA) y 125.3 (INV), y los tamaños del efecto de estos índices son grandes.

En general, las medias de las puntuaciones escalares de las pruebas son significativamente superiores a las del grupo control, con tamaños del efecto de moderados a grandes. Los mayores tamaños del efecto se observaron en *Aritmética, Dígitos* e *Información*, y los resultados son consistentes con los obtenidos en estudios previos (Wechsler, 2003, 2008).

Las puntuaciones más altas del grupo de niños con alta capacidad intelectual confirman su identificación como tales, y siguen la misma línea tanto del estudio de validez estadounidense realizado con un grupo de 95 niños con alta capacidad intelectual (véase tabla 6.13; Wechsler, 2014a) como de los estudios comparativos realizados previamente (Rimm et al., 2008; Rowe et al., 2010; Wechsler, 2002, 2003, 2008, 2012). Estos estudios también indican que los niños con alta capacidad intelectual muestran un rendimiento más bajo en las pruebas diseñadas para medir la velocidad de procesamiento que en las pruebas de otros dominios cognitivos. Los resultados de la WISC-V y los obtenidos por grupos con alta capacidad intelectual evaluados con otras escalas de inteligencia de Wechsler previas también indican que estos instrumentos miden constructos similares. Estos resultados demuestran que la WISC-V es una herramienta útil para la detección de la alta capacidad intelectual. Este estudio se llevó a cabo con niños a los que ya se les había diagnosticado alta capacidad intelectual; por lo que los resultados no deberían generalizarse a niños que presenten un alto rendimiento en otros dominios.

Tabla 6.13. Resultados del grupo con alta capacidad intelectual y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Alta capacidad intelectual		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	15.1	2.4	10.9	2.7	95	−4.16	−10.53	<.01	−1.63
Vocabulario	14.9	2.3	11.2	2.8	95	−3.77	−9.62	<.01	−1.47
Información	14.6	2.5	11.0	2.8	95	−3.53	−9.93	<.01	−1.33
Comprensión	14.1	2.5	10.9	2.9	95	−3.19	−7.64	<.01	−1.18
Cubos	13.9	2.4	10.9	2.6	95	−3.03	−9.50	<.01	−1.21
Puzles visuales	13.5	2.1	11.0	2.4	95	−2.47	−6.97	<.01	−1.10
Matrices	13.3	2.8	10.9	2.6	95	−2.38	−6.19	<.01	−0.88
Balanzas	13.8	2.5	10.8	2.4	95	−2.95	−8.55	<.01	−1.20
Aritmética	13.9	2.5	10.6	2.4	95	−3.24	−8.67	<.01	−1.32
Dígitos	14.0	2.5	11.0	2.6	95	−3.00	−7.69	<.01	−1.18
Span de dibujos	12.3	2.9	10.4	2.5	95	−1.93	−4.52	<.01	−0.71
Letras y números	14.3	2.5	11.1	2.8	93	−3.20	−7.91	<.01	−1.21
Claves	12.1	2.7	9.9	2.6	94	−2.19	−5.78	<.01	−0.83
Búsqueda de símbolos	12.5	2.7	10.2	2.9	91	−2.27	−5.23	<.01	−0.81
Cancelación	11.8	3.2	9.8	2.9	95	−2.01	−4.68	<.01	−0.66
Cs	13.9	2.7	10.8	2.6	95	−3.11	−8.87	<.01	−1.17
Cp	13.8	2.4	10.8	2.7	95	−3.01	−9.20	<.01	−1.18
Dd	12.9	2.9	10.6	2.6	95	−2.29	−5.13	<.01	−0.83
Di	12.9	2.9	10.9	2.8	95	−1.97	−4.75	<.01	−0.69
Dc	13.6	2.4	10.9	3.0	95	−2.74	−6.72	<.01	−1.01
CAa	11.6	3.0	9.9	3.1	95	−1.69	−3.93	<.01	−0.55
CAe	11.8	3.2	9.7	2.9	95	−2.14	−4.82	<.01	−0.70
ICV	127.7	12.3	105.8	12.9	95	−21.97	−11.09	<.01	−1.74
IVE	121.2	11.5	105.2	12.2	95	−15.98	−9.55	<.01	−1.35
IRF	120.3	12.0	105.1	12.3	95	−15.26	−8.80	<.01	−1.26
IMT	117.9	11.7	104.0	12.1	95	−13.86	−7.33	<.01	−1.16
IVP	112.9	13.5	100.4	13.5	90	−12.44	−6.05	<.01	−0.92
CIT	127.5	8.8	105.7	12.2	94	−21.85	−13.19	<.01	−2.05
IRC	122.1	11.8	104.1	11.4	95	−18.04	−10.41	<.01	−1.55
IMTA	123.0	12.9	105.9	13.0	93	−17.13	−8.49	<.01	−1.32
INV	122.9	10.5	104.6	11.8	94	−18.28	−10.80	<.01	−1.64
ICG	127.1	9.6	106.3	12.4	95	−20.83	−12.54	<.01	−1.88
ICC	118.8	11.0	102.1	12.4	90	−16.73	−8.71	<.01	−1.43

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Niños con discapacidad intelectual leve o moderada

De acuerdo con los criterios especificados en el DSM-5 (American Psychiatric Association, 2013), un niño diagnosticado con discapacidad intelectual debe mostrar déficits tanto en las funciones intelectuales como en el funcionamiento adaptativo en, al menos, una actividad cotidiana realizada en distintos ámbitos. Los déficits han de estar confirmados tanto por la evaluación clínica como por los resultados de la evaluación de la aptitud intelectual mediante pruebas estandarizadas aplicadas individualmente; asimismo dichos déficits deben haberse mostrado durante el periodo de desarrollo. De acuerdo con la American Association on Intellectual and Developmental Disabilities (AAIDD), la discapacidad intelectual se origina antes de los 18 años y se caracteriza por una serie de limitaciones significativas tanto en el funcionamiento intelectual como en el comportamiento adaptativo expresado en las habilidades prácticas, conceptuales y sociales (AAIDD, 2010). Tanto el DSM-5 como la AAIDD coinciden al señalar que un rendimiento significativamente bajo en un test de aptitud intelectual general, como la WISC-V, es un criterio necesario pero no suficiente para diagnosticar discapacidad intelectual; también hay que evaluar el funcionamiento adaptativo.

Asimismo destacan la necesidad de considerar el nivel de error de medida de la escala al interpretar las puntuaciones, haciendo hincapié en la importancia de tener en cuenta distintas fuentes de información y de aplicar cuidadosamente el juicio clínico en la identificación de niños con discapacidad intelectual. En una distribución normal de puntuaciones CI, alrededor del 2.2% de sujetos obtienen puntuaciones como mínimo de 2 desviaciones típicas por debajo de la media (p. ej., CI ≤ 70). La prevalencia de la discapacidad intelectual en la población varía de un estudio a otro. Un metaanálisis exhaustivo indicó que la prevalencia de la discapacidad intelectual en la población general era de 10.37 por 1000 (Maulik, Mascarenhas, Mathers, Dua y Saxena, 2011), o aproximadamente del 1 %. Este porcentaje más bajo se debe a que los criterios diagnósticos de la discapacidad intelectual requieren que el rendimiento sea bajo tanto en la aptitud intelectual como en el funcionamiento adaptativo.

Se han llevado a cabo numerosos estudios para evaluar el rendimiento de los sujetos con discapacidad intelectual en versiones anteriores de las escalas de inteligencia de Wechsler. Al respecto se ha demostrado que la prevalencia de diferencias marcadas e inusuales entre los índices verbales y los no verbales disminuye al mismo tiempo que disminuye el nivel de aptitud (Gordon, Duff, Davidson y Whitaker, 2010; Spruill, 1998; Wechsler, 2002, 2003). En algunos estudios se ha observado que el rendimiento en comprensión verbal y en memoria de trabajo es relativamente más bajo que el rendimiento en razonamiento perceptivo y en velocidad de procesamiento (Gordon et al., 2010). Además, las desviaciones típicas de las pruebas y de las puntuaciones compuestas son menores para los sujetos con discapacidad intelectual que para los sujetos de la población general (Nunes et al., 2012; Wechsler, 2002, 2003, 2008, 2012). Así pues, parece que la variabilidad en el rendimiento, tanto en las pruebas como en las puntuaciones compuestas, es menor en los sujetos con discapacidad intelectual que en los sujetos de la población general. Un estudio, llevado a cabo por Van der Molen, Henry y Van Luit (2014), que incluía tareas de memoria de trabajo verbal y visual indica que pueden observarse puntuaciones más bajas en el IMTA que en el IMT.

En el estudio de validez español, se aplicó la WISC-V a 20 sujetos no institucionalizados, de entre 6 y 16 años, que habían sido diagnosticados con discapacidad intelectual. En pruebas estandarizadas de aptitud cognitiva, los niños habían obtenido puntuaciones de 2-4 desviaciones típicas por debajo de la media (p. ej., CI total = 40-70) o cumplían los criterios del DSM-5 para el diagnóstico de discapacidad intelectual leve o moderada. La tabla 6.14 presenta las medias de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas obtenidas por niños con discapacidad intelectual y por el grupo control correspondiente.

Tabla 6.14. Resultados del grupo con discapacidad intelectual y del grupo control (datos españoles)

Prueba/ punt. de procesamiento/ punt. compuesta	Discapacidad intelectual		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	3.7	1.9	9.3	2.4	20	5.55	9.05	<.01	2.56
Vocabulario	3.1	1.9	9.2	2.5	20	6.10	9.99	<.01	2.75
Información	3.2	1.7	9.3	2.5	20	6.10	9.18	<.01	2.85
Comprensión	3.0	2.3	9.5	2.5	20	6.45	7.04	<.01	2.69
Cubos	4.3	2.4	9.8	2.5	20	5.45	10.79	<.01	2.22
Puzles visuales	3.6	1.6	10.0	2.1	20	6.35	10.59	<.01	3.40
Matrices	3.5	1.7	9.6	2.4	20	6.05	9.13	<.01	2.91
Balanzas	3.7	1.7	10.4	2.6	20	6.65	9.20	<.01	3.03
Aritmética	2.7	1.8	10.4	2.3	20	7.70	11.57	<.01	3.73
Dígitos	2.2	1.6	10.7	2.7	20	8.45	12.67	<.01	3.81
Span de dibujos	3.4	1.7	10.6	3.0	20	7.20	8.34	<.01	2.95
Letras y números	2.3	1.6	10.3	3.5	19	7.95	9.17	<.01	2.92
Claves	3.9	2.7	10.4	2.7	20	6.50	7.21	<.01	2.41
Búsqueda de símbolos	4.4	2.5	10.6	3.3	20	6.15	6.94	<.01	2.10
Cancelación	4.7	3.2	11.2	2.8	20	6.50	6.18	<.01	2.16
Cs	4.2	2.5	9.6	2.4	20	5.35	10.10	<.01	2.18
Cp	5.5	1.5	10.3	2.5	6	4.83	7.39	<.01	2.34
Dd	2.8	1.9	10.8	2.6	20	8.00	11.43	<.01	3.51
Di	3.2	2.2	10.5	2.6	20	7.30	10.54	<.01	3.03
Dc	2.9	2.1	10.5	2.4	20	7.60	9.55	<.01	3.37
CAa	5.2	3.3	11.1	2.8	20	5.85	5.35	<.01	1.91
CAe	4.5	3.2	11.1	3.4	20	6.55	6.25	<.01	1.98
ICV	63.1	10.4	95.6	11.8	20	32.55	10.71	<.01	2.93
IVE	65.7	10.9	99.1	11.5	20	33.40	11.99	<.01	2.98
IRF	63.8	8.7	99.7	13.6	20	35.90	9.25	<.01	3.14
IMT	58.9	9.1	103.3	13.7	20	44.35	11.51	<.01	3.81
IVP	65.8	15.1	102.7	15.2	20	36.95	7.45	<.01	2.44
CIT	56.2	10.0	99.0	12.8	20	42.85	12.76	<.01	3.73
IRC	60.4	9.8	101.9	12.1	20	41.50	11.61	<.01	3.77
IMTA	54.2	10.3	102.8	16.3	19	48.63	11.20	<.01	3.57
INV	58.3	9.2	100.6	13.6	20	42.25	11.58	<.01	3.64
ICG	61.7	7.6	97.2	12.7	20	35.45	11.87	<.01	3.39
ICC	55.8	13.0	103.7	14.9	20	47.90	10.35	<.01	3.43

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Las medias de los índices primarios del grupo con discapacidad intelectual oscilan entre 58.9 (IMT) y 65.8 (IVP), y la media del CI total es de 56.2. Tanto la media del CI total como las medias de los índices primarios son significativamente más bajas que las medias correspondientes del grupo control. Las puntuaciones relativamente altas en el IVP son consistentes con los resultados aportados por Gordon et al. (2010), Elliot (2007) y Wechsler (2003). Los tamaños del efecto para las puntuaciones de los índices primarios fueron grandes.

Las medias de los índices secundarios del grupo con discapacidad intelectual oscilan entre 58.3 (INV) y 61.7 (ICG). La media del INV obtenida por este grupo es de 58.3, un valor similar al del CI total, aunque ligeramente superior a este. Aunque el tamaño del efecto del INV es ligeramente más pequeño que el del CI total, sigue siendo grande.

La variabilidad en las puntuaciones de las pruebas principales y en los índices primarios obtenidas por los grupos con discapacidad intelectual es, por lo general, más pequeña que en el grupo control en todos los dominios. Las desviaciones típicas de los índices primarios y del CI total

oscilaron entre 8.7 (IRF) y 10.9 (IVE), excepto para el IVP que mostró un valor de 15.1. En cuanto a las desviaciones típicas de las pruebas se observa un patrón similar.

El grupo de discapacidad intelectual obtuvo, en todas las pruebas, puntuaciones significativamente inferiores a las correspondientes del grupo control. Los tamaños del efecto más grandes se obtuvieron en *Dígitos, Aritmética, Puzles visuales, Balanzas, Span de dibujos, Letras y números, Matrices* e *Información.*

Estos resultados demuestran que la WISC-V es una herramienta útil para la detección de la discapacidad intelectual y siguen la misma línea del estudio de validez estadounidense realizado con un grupo de 111 sujetos no institucionalizados, de entre 6 y 16 años, que habían sido diagnosticados con discapacidad intelectual, 74 niños con discapacidad intelectual leve y 37 con discapacidad intelectual moderada (véanse las tablas 6.15 y 6.16; Wechsler, 2014a).

Tabla 6.15. Resultados del grupo con discapacidad intelectual leve y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Discapacidad intelectual leve		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	3.8	2.1	9.5	3.2	74	5.74	13.92	<.01	2.12
Vocabulario	3.9	1.9	9.0	3.2	74	5.07	12.40	<.01	1.93
Información	3.8	1.9	9.6	3.1	74	5.76	15.00	<.01	2.24
Comprensión	4.0	1.9	8.7	2.9	74	4.69	12.37	<.01	1.91
Cubos	3.9	2.2	10.1	2.9	74	6.22	14.18	<.01	2.42
Puzles visuales	4.1	1.6	10.3	2.9	74	6.22	16.03	<.01	2.66
Matrices	4.1	2.6	10.0	3.4	74	5.89	11.62	<.01	1.95
Balanzas	4.3	2.1	9.8	3.3	74	5.46	11.96	<.01	1.97
Aritmética	3.2	1.8	9.6	3.2	74	6.34	13.39	<.01	2.44
Dígitos	3.4	1.9	9.9	2.8	73	6.51	15.55	<.01	2.72
Span de dibujos	4.3	2.3	9.7	2.9	74	5.42	11.49	<.01	2.07
Letras y números	3.5	2.1	9.7	2.7	73	6.18	14.23	<.01	2.56
Claves	4.6	2.9	9.6	2.8	70	5.00	11.06	<.01	1.75
Búsqueda de símbolos	5.3	3.3	9.5	2.2	67	4.15	8.51	<.01	1.48
Cancelación	6.1	3.4	10.8	2.5	73	4.67	8.61	<.01	1.56
Cs	3.8	2.2	10.1	3.1	74	6.28	13.91	<.01	2.34
Cp	4.1	1.9	10.1	3.0	74	6.04	14.09	<.01	2.41
Dd	4.7	2.6	10.4	3.2	74	5.66	11.05	<.01	1.94
Di	4.5	2.2	9.8	2.7	74	5.28	13.08	<.01	2.14
Dc	3.9	2.0	9.7	2.8	73	5.75	14.94	<.01	2.36
CAa	6.9	3.6	10.8	2.6	73	3.85	7.33	<.01	1.23
CAe	5.8	3.3	10.8	2.6	73	5.00	9.24	<.01	1.68
ICV	66.0	10.9	96.1	16.4	74	30.14	14.21	<.01	2.16
IVE	66.0	9.9	101.1	14.6	74	35.14	16.72	<.01	2.82
IRF	67.0	11.0	99.3	16.1	74	32.34	13.49	<.01	2.35
IMT	65.1	10.5	98.7	14.6	73	33.60	14.93	<.01	2.64
IVP	71.6	16.2	97.3	10.9	67	25.78	11.45	<.01	1.87
CIT	60.9	8.9	98.0	15.6	69	37.07	16.22	<.01	2.92
IRC	64.2	9.7	98.1	15.1	74	33.86	15.16	<.01	2.67
IMTA	62.2	11.5	99.2	13.8	72	36.96	15.96	<.01	2.91
INV	62.1	8.7	99.5	15.2	70	37.40	16.46	<.01	3.02
ICG	63.5	8.8	97.9	15.7	74	34.46	15.95	<.01	2.71
ICC	63.4	12.3	97.6	13.4	67	34.19	14.75	<.01	2.66

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Tabla 6.16. Resultados del grupo con discapacidad intelectual moderada y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Discapacidad intelectual moderada		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	2.2	1.9	10.1	3.5	37	7.89	11.65	<.01	2.80
Vocabulario	2.4	1.7	9.6	3.1	37	7.22	11.75	<.01	2.89
Información	2.3	1.8	9.6	2.7	37	7.30	13.42	<.01	3.18
Comprensión	2.6	2.4	9.5	2.9	37	6.89	11.37	<.01	2.59
Cubos	2.4	1.5	9.5	2.9	37	7.14	13.18	<.01	3.09
Puzles visuales	2.8	1.9	9.3	3.0	37	6.51	12.72	<.01	2.59
Matrices	2.4	2.1	9.7	3.3	37	7.30	11.16	<.01	2.64
Balanzas	3.2	2.1	9.4	3.6	36	6.17	9.24	<.01	2.09
Aritmética	2.2	1.5	10.0	2.9	37	7.81	14.87	<.01	3.38
Dígitos	2.0	1.4	9.7	3.0	34	7.74	13.32	<.01	3.31
Span de dibujos	3.5	2.3	10.2	3.4	36	6.69	9.94	<.01	2.30
Letras y números	2.6	1.7	9.9	3.0	33	7.27	11.72	<.01	2.98
Claves	3.1	2.6	10.4	2.7	37	7.32	10.86	<.01	2.76
Búsqueda de símbolos	3.2	2.7	9.9	2.8	37	6.73	10.05	<.01	2.45
Cancelación	4.6	3.5	10.2	2.8	37	5.68	8.56	<.01	1.79
Cs	2.4	1.5	9.4	2.8	37	7.03	13.57	<.01	3.13
Cp	2.8	1.8	9.4	3.0	36	6.67	12.61	<.01	2.70
Dd	3.8	2.2	9.7	3.0	37	5.95	9.82	<.01	2.26
Di	3.0	2.0	9.6	3.1	37	6.57	10.69	<.01	2.52
Dc	2.6	1.8	10.0	2.8	34	7.35	13.13	<.01	3.12
CAa	5.3	3.6	10.4	2.8	37	5.11	7.37	<.01	1.58
CAe	4.3	3.2	10.1	2.8	37	5.73	9.00	<.01	1.91
ICV	55.2	11.3	99.4	15.2	37	44.19	13.37	<.01	3.30
IVE	56.8	9.6	96.6	14.6	37	39.86	14.71	<.01	3.23
IRF	58.6	12.0	97.1	17.3	36	38.44	11.36	<.01	2.58
IMT	58.3	10.6	99.4	17.0	33	41.09	11.48	<.01	2.90
IVP	59.3	15.8	101.1	12.6	37	41.76	11.31	<.01	2.92
CIT	49.7	8.9	98.5	16.8	33	48.79	13.72	<.01	3.63
IRC	57.1	10.6	98.2	15.2	36	41.08	13.15	<.01	3.14
IMTA	54.1	9.4	99.0	15.2	32	44.94	14.12	<.01	3.56
INV	53.4	10.1	98.0	17.2	35	44.54	13.27	<.01	3.16
ICG	54.3	9.4	97.8	16.4	36	43.50	13.40	<.01	3.25
ICC	52.5	12.7	99.9	15.8	33	47.45	12.26	<.01	3.31

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Niños con capacidad intelectual límite

La capacidad intelectual límite describe un funcionamiento intelectual inferior a la media y superior al rango de discapacidad intelectual (característicamente, un CI total de 71 a 84), que es relevante clinicamente para el funcionamiento de un sujeto (American Psychiatric Association, 2013; Bonifacci y Snowling, 2008; Van der Molen, Van Luit, Van der Molen y Jongmans, 2010). A veces se puede diagnosticar discapacidad intelectual leve a sujetos con un CI total de 71 a 75; no obstante, tal diagnóstico requiere la presencia de otros déficits funcionales adaptativos. Para diferenciar la discapacidad intelectual y la capacidad intelectual límite, es imprescindible realizar una evaluación detallada del funcionamiento intelectual y adaptativo.

Se aplicó la WISC-V a 20 niños no institucionalizados, de entre 6 y 16 años, con capacidad intelectual límite. En pruebas estandarizadas de aptitud cognitiva, los participantes del grupo con capacidad intelectual límite habían obtenido puntuaciones de entre 71 y 84. La tabla 6.17 presenta las medias de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas de la WISC-V, para el grupo de niños con capacidad intelectual límite y para el grupo control correspondiente.

Tabla 6.17. Resultados del grupo con capacidad intelectual límite y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Capacidad intelectual límite		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	7.0	1.4	9.0	2.9	20	2.00	2.59	.02	0.88
Vocabulario	6.2	2.0	8.7	2.9	19	2.53	3.33	<.01	1.02
Información	6.8	2.2	9.1	2.7	20	2.30	2.86	.01	0.93
Comprensión	6.6	2.3	9.2	3.1	20	2.60	3.25	<.01	0.95
Cubos	6.8	2.1	9.6	2.5	20	2.85	4.53	<.01	1.23
Puzles visuales	7.1	1.7	9.4	2.4	20	2.30	3.15	<.01	1.11
Matrices	8.5	3.3	9.8	2.2	20	1.30	1.44	.17	0.46
Balanzas	7.1	1.9	9.3	1.8	20	2.25	3.39	<.01	1.22
Aritmética	6.5	2.2	9.7	2.7	20	3.20	4.38	<.01	1.30
Dígitos	5.9	2.6	9.6	2.8	20	3.75	5.10	<.01	1.39
Span de dibujos	6.5	2.3	9.4	2.9	19	2.84	3.57	<.01	1.09
Letras y números	5.7	2.0	9.7	3.3	20	4.05	5.00	<.01	1.48
Claves	9.4	2.2	9.3	2.7	20	−0.10	−0.13	.90	−0.04
Búsqueda de símbolos	8.9	3.0	9.5	3.0	19	0.53	0.63	.54	0.18
Cancelación	10.4	2.5	11.1	3.2	20	0.70	0.70	.49	0.24
Cs	6.6	2.2	9.8	2.4	20	3.15	4.52	<.01	1.37
Cp	7.2	2.6	9.7	2.5	20	2.50	3.17	<.01	0.98
Dd	6.8	3.5	10.3	3.4	20	3.45	2.94	<.01	1.00
Di	7.0	2.0	8.8	2.4	20	1.85	3.37	<.01	0.84
Dc	6.3	2.5	9.5	3.6	20	3.20	3.71	<.01	1.03
CAa	10.4	2.3	10.7	3.2	20	0.25	0.27	.79	0.09
CAe	10.0	2.5	11.3	3.7	20	1.25	1.15	.26	0.40
ICV	81.7	7.6	93.1	13.8	19	11.42	3.08	<.01	1.03
IVE	83.1	8.3	97.2	11.6	20	14.10	5.06	<.01	1.40
IRF	87.1	11.7	97.3	9.0	20	10.25	2.88	<.01	0.98
IMT	78.2	11.9	97.1	14.2	19	18.95	5.10	<.01	1.45
IVP	95.1	11.9	96.4	13.1	19	1.37	0.33	.75	0.11
CIT	80.4	5.7	94.5	9.2	19	14.16	5.53	<.01	1.85
IRC	81.3	7.8	97.1	10.4	20	15.75	4.97	<.01	1.71
IMTA	76.6	11.1	98.0	14.9	20	21.35	6.07	<.01	1.63
INV	82.1	7.9	95.0	8.6	19	12.89	4.18	<.01	1.56
ICG	80.8	7.5	94.6	11.8	19	13.79	4.26	<.01	1.39
ICC	84.0	10.4	95.7	11.5	19	11.74	3.33	<.01	1.07

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Las medias de los índices primarios del grupo con capacidad intelectual límite oscilan entre 78.2 (IMT) y 95.1 (IVP), y la media del CI total es de 80.4. Todas las medias, a excepción del IVP, son significativamente inferiores en comparación con las medias del grupo control, y los tamaños del efecto son grandes. Los niños con capacidad intelectual límite obtuvieron, en la mayoría de las pruebas, puntuaciones significativamente inferiores a las correspondientes del grupo control. Los

tamaños del efecto más grandes para las diferencias entre las medias de las puntuaciones escalares se observan en las pruebas *Letras y números, Dígitos* y *Aritmética.*

Las medias de los índices secundarios oscilan entre 76.6 (IMTA) y 84.0 (ICC). Todos los niños del grupo con capacidad intelectual límite obtuvieron un CI total igual o inferior a 90; mientras que solo el 30 % de los niños del grupo control obtuvieron un CI total igual o inferior a 90. Además, aproximadamente el 90 % de los niños con capacidad intelectual límite obtuvieron un INV igual o inferior a 90; mientras que solo el 35 % de los niños del grupo control obtuvieron un INV igual o inferior a 90.

Niños con trastorno específico del aprendizaje

Los criterios para la identificación de niños con trastornos del aprendizaje pueden variar en función del sistema educativo local, regional o nacional. No existe un modelo diagnóstico estándar para la identificación de los trastornos del aprendizaje. La ley IDEIA (*Individuals with Disabilities Education Improvement Act*; 2004) contempla distintos enfoques para identificar a los niños que deben clasificarse como de educación especial por algún trastorno de aprendizaje, incluidos los puntos fuertes y débiles en el procesamiento, la respuesta a la intervención, y la discrepancia aptitud-rendimiento grave. Los criterios diagnósticos del DSM-5 para los trastornos del aprendizaje son unas directrices generales que indican que si existen dificultades en la lectura, la ortografía, la escritura o las matemáticas durante, al menos, seis meses y persisten a pesar de que se lleve a cabo algún tipo de intervención, puede determinarse que el niño tiene algún trastorno específico del aprendizaje. En una evaluación clínica ha de confirmarse que las destrezas afectadas se encuentran, sustancial y cuantificablemente, en un nivel inferior al de las puntuaciones esperadas en niños de la misma edad. Además se ofrecen indicadores para identificar el tipo de problema del aprendizaje (p. ej., lectura, matemáticas y/o expresión escrita).

El DSM-5 no proporciona una metodología específica para determinar las dificultades en el funcionamiento académico; por ello, en la aplicación de este procedimiento diagnóstico puede darse cierta variabilidad. Los niños que participaron en estos estudios se seleccionaron en distintos ámbitos clínicos y aplicando criterios diversos; por lo que se predijo que el rendimiento del grupo en las pruebas de la WISC-V sería bastante heterogéneo. En algunos casos se pudieron usar patrones de aptitud para identificar los problemas del aprendizaje, pero en otros casos no. A pesar de los diferentes métodos seguidos para identificar a los niños con trastornos del aprendizaje, se predijo que se observaría un patrón de puntos fuertes y débiles en la WISC-V.

Los modelos contemporáneos de trastornos del aprendizaje se centran en tres componentes del aprendizaje y del desarrollo académico: los procesos cognitivos, las destrezas y conocimientos académicos, y los puntos fuertes y débiles en el procesamiento (Fiorello et al., 2006; Flanagan, Fiorello y Ortiz, 2010). La WISC-V se ha diseñado para ayudar a los profesionales en la evaluación de los procesos cognitivos generales y específicos que permiten la adquisición y el uso de las aptitudes académicas. Por ejemplo, las aptitudes generales, como la memoria de trabajo, la velocidad de procesamiento y la inteligencia cristalizada, y las aptitudes específicas, como la denominación automática rápida y la memoria a largo plazo, son importantes para la adquisición y el desarrollo normales de las destrezas de lectura (Fiorello et al., 2006). Para validar las destrezas cognitivas generales y específicas medidas por la WISC-V y que podían verse afectadas por dificultades del aprendizaje, se recogieron varias muestras de niños con trastornos del aprendizaje.

Trastorno específico del aprendizaje-lectura

Son numerosos los estudios centrados en las dificultades cognitivas generales y específicas asociadas al trastorno de la lectura. Aunque el análisis de estos estudios queda fuera del alcance de este capítulo, sí que se exponen aquí algunas observaciones pertinentes. En el dominio verbal, los estudios muestran que el conocimiento léxico se asocia con el desarrollo de las destrezas lectoras (Ouellette, 2006). Los niños con trastorno específico del aprendizaje-lectura tienen dificultades con la búsqueda y recuperación semántica (Booth, Bebko, Burman y Bitan, 2007), y este trastorno se asocia con un rendimiento más bajo en medidas del lenguaje expresivo, pero no receptivo (Cutting, Materek, Cole, Levine y Mahone, 2009). Los niños con déficits de comprensión lectora muestran un deterioro en el funcionamiento del lenguaje en comparación con los niños del

grupo control y con niños con déficits solo de decodificación (Catts, Adlof y Weismer, 2006). En una amplia muestra de niños diagnosticados con TDAH y trastornos del aprendizaje, las puntuaciones obtenidas en las pruebas de comprensión verbal y memoria de trabajo fueron los mejores indicadores de la aptitud lectora de la WISC-III/WISC-IV; en cambio, las puntuaciones obtenidas en las pruebas de memoria de trabajo y velocidad de procesamiento fueron los mejores indicadores de trastorno del aprendizaje (Mayes y Calhoun, 2007). Los niños diagnosticados con trastorno específico del aprendizaje-lectura muestran una menor memoria de trabajo verbal (Kibby y Cohen, 2008) y déficits de velocidad de procesamiento (Shanahan et al., 2006).

Asimismo se ha observado que los niños con trastornos de la lectura obtienen resultados más bajos en medidas de aprendizaje verbal (Kibby y Cohen, 2008) y en tareas de aprendizaje de pares asociados que implican emparejar un input visual (un símbolo) con una respuesta verbal, es decir, una palabra con o sin sentido (Litt y Nation, 2014; Messbauer y de Jong, 2003). Los puntos débiles en la comprensión verbal, la memoria de trabajo y la velocidad de procesamiento se asocian tanto con el trastorno de la lectura como con el de matemáticas. Además, los problemas de la lectura se relacionan con la conciencia fonológica y la velocidad de denominación (Willcutt et al., 2013).

Se aplicó la WISC-V a 30 niños de 7 a 16 años, que habían sido identificados como sujetos con trastorno específico del aprendizaje-lectura de acuerdo con los criterios del DSM-5. La tabla 6.18 presenta las medias de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas de la WISC-V, para el grupo de niños con trastorno específico del aprendizaje-lectura y para el grupo control correspondiente.

Tabla 6.18. Resultados del grupo con trastorno específico del aprendizaje-lectura y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Trastorno específico del aprendizaje-lectura		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	8.2	2.2	10.3	2.4	30	2.07	3.81	<.01	0.90
Vocabulario	7.7	2.6	10.0	3.0	30	2.23	3.22	<.01	0.79
Información	8.5	1.9	10.1	2.4	30	1.60	2.91	<.01	0.74
Comprensión	8.6	3.2	10.2	2.9	30	1.60	2.30	.03	0.52
Cubos	9.1	2.6	10.3	3.0	30	1.13	1.64	.11	0.40
Puzles visuales	8.5	3.1	10.3	2.2	30	1.87	3.23	<.01	0.70
Matrices	8.6	2.1	10.4	2.6	30	1.83	3.22	<.01	0.77
Balanzas	8.9	2.3	10.3	3.0	30	1.37	1.85	.08	0.51
Aritmética	8.4	2.1	10.7	3.2	30	2.30	3.80	<.01	0.85
Dígitos	8.2	1.6	10.5	2.3	30	2.33	5.04	<.01	1.18
Span de dibujos	7.7	2.6	10.8	2.1	30	3.10	5.34	<.01	1.31
Letras y números	8.2	2.3	9.9	2.1	30	1.73	3.45	<.01	0.79
Claves	8.7	3.1	10.1	3.0	30	1.40	2.44	.02	0.46
Búsqueda de símbolos	8.8	2.7	10.0	2.9	30	1.23	1.90	.07	0.44
Cancelación	9.5	3.1	9.5	3.3	29	0.03	0.04	.97	0.01
Cs	9.1	2.8	10.3	3.1	30	1.17	1.68	.10	0.40
Cp	8.7	2.7	10.5	3.2	30	1.83	2.82	<.01	0.62
Dd	8.4	1.7	9.6	2.6	30	1.20	2.52	.02	0.55
Di	8.5	1.8	10.5	2.2	30	2.03	3.45	<.01	1.01
Dc	8.6	2.4	10.8	2.9	30	2.13	3.75	<.01	0.80
CAa	9.7	2.8	9.7	3.1	29	0.00	0.00	1.00	0.00
CAe	9.4	3.1	9.5	3.3	29	0.10	0.12	.91	0.03
ICV	89.1	11.2	100.7	12.6	30	11.63	4.34	<.01	0.98
IVE	93.3	14.1	101.6	12.4	30	8.27	2.90	<.01	0.62
IRF	92.5	10.8	101.9	13.5	30	9.40	2.97	<.01	0.77
IMT	87.8	10.1	104.1	11.2	30	16.23	6.65	<.01	1.52
IVP	93.0	15.3	100.3	14.4	30	7.37	2.53	.02	0.50
CIT	88.9	10.5	102.0	13.9	30	13.07	5.19	<.01	1.06
IRC	92.2	9.5	102.7	16.1	30	10.57	3.18	<.01	0.80
IMTA	90.1	9.1	101.2	10.3	30	11.07	5.51	<.01	1.14
INV	89.6	11.8	102.6	13.2	30	13.03	5.55	<.01	1.04
ICG	90.0	11.0	101.6	13.1	30	11.63	4.22	<.01	0.96
ICC	88.6	12.6	102.7	13.3	30	14.17	6.02	<.01	1.09

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

En comparación con el grupo control, los niños con trastorno específico del aprendizaje-lectura obtuvieron medias más bajas en todos los índices primarios, con tamaños del efecto de moderados a grandes en todas las medidas. Los tamaños del efecto más grandes se observan en el IMT y el ICV, resultados que son consistentes con las investigaciones contemporáneas que indican que existe una relación entre el rendimiento lector y las dificultades con distintos componentes de la memoria de trabajo (Wang y Gathercole, 2013). Todas las puntuaciones compuestas globales (CIT, INV, ICG e ICC) tienen tamaños del efecto grandes.

Se observa un rendimiento significativamente más bajo en la mayoría de las pruebas, con las medias más bajas en las pruebas *Vocabulario* y *Span de dibujos*. Los tamaños del efecto más grandes se corresponden a las pruebas *Span de dibujos* y *Dígitos*. Los resultados indican dificultades significativas con la comprensión verbal y la memoria de trabajo.

Trastorno específico del aprendizaje-lectura y expresión escrita

Las investigaciones sugieren que las dificultades tanto en la lectura como en la escritura pueden guardar relación con procesos neurológicos y cognitivos que tienen en común y que intervienen en la adquisición y el desarrollo de estas destrezas (Berninger y Richards, 2010). El procesamiento fonológico se asocia con deterioros en la lectura y el deletreo de palabras; mientras que otros déficits en la codificación ortográfica expresiva y receptiva se relacionan con la caligrafía, el deletreo y la composición (Berninger y Richards, 2010).

Se aplicó la WISC-V a 22 niños, de 6 a 14 años, que cumplían los criterios diagnósticos del DSM-5 para el trastorno específico del aprendizaje-lectura y expresión escrita. La tabla 6.19 presenta las medias de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas de la WISC-V, para el grupo de niños con trastorno específico del aprendizaje-lectura y expresión escrita para el grupo control correspondiente.

Tabla 6.19. Resultados del grupo con trastorno específico del aprendizaje-lectura y expresión escrita y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Trastorno específico del aprendizaje-lectura y expresión escrita		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	7.2	2.0	9.1	2.1	22	1.91	3.78	<.01	0.93
Vocabulario	7.8	2.4	8.9	2.9	22	1.09	1.24	.23	0.41
Información	8.0	1.8	9.3	1.9	22	1.23	2.13	.05	0.66
Comprensión	7.5	2.1	8.5	2.6	22	0.91	1.43	.17	0.39
Cubos	9.1	2.8	10.1	2.5	22	1.00	1.57	.13	0.38
Puzles visuales	9.6	2.6	9.2	2.8	22	−0.45	−0.70	.49	−0.17
Matrices	8.2	2.2	9.4	2.8	22	1.14	1.29	.21	0.45
Balanzas	7.8	3.2	9.9	2.8	22	2.14	2.38	.03	0.71
Aritmética	7.4	2.0	9.6	2.8	22	2.27	3.70	<.01	0.93
Dígitos	7.2	2.6	10.2	3.1	22	3.05	3.05	<.01	1.07
Span de dibujos	8.0	2.0	9.4	3.0	22	1.36	2.73	.01	0.53
Letras y números	7.4	2.2	9.8	2.0	22	2.45	4.10	<.01	1.17
Claves	7.8	3.3	9.0	2.8	22	1.14	1.31	.20	0.37
Búsqueda de símbolos	9.6	2.9	9.7	2.5	22	0.09	0.12	.91	0.03
Cancelación	9.8	3.3	9.7	3.1	22	−0.09	−0.11	.91	−0.03
Cs	9.1	2.8	10.1	2.9	22	1.05	1.56	.13	0.37
Cp	8.9	2.8	9.9	2.6	22	0.95	1.41	.17	0.35
Dd	8.4	3.5	9.9	3.1	22	1.45	1.32	.20	0.44
Di	6.9	2.4	10.2	3.0	22	3.27	3.46	<.01	1.20
Dc	7.8	2.8	10.2	2.6	22	2.41	2.67	.01	0.89
CAa	9.9	3.1	9.8	3.0	22	−0.09	−0.11	.92	−0.03
CAe	9.6	3.3	9.6	3.1	22	−0.05	−0.06	.95	−0.02
ICV	86.5	10.1	94.6	11.7	22	8.09	2.50	.02	0.74
IVE	96.2	13.3	98.0	12.5	22	1.73	0.64	.53	0.13
IRF	88.4	12.2	97.8	13.3	22	9.45	2.50	.02	0.74
IMT	85.8	9.7	98.7	13.9	22	12.95	3.68	<.01	1.08
IVP	93.0	15.8	96.3	11.3	22	3.36	0.82	.42	0.24
CIT	84.8	11.1	96.2	10.5	22	11.41	3.81	<.01	1.06
IRC	85.9	12.8	98.5	13.0	22	12.64	3.66	<.01	0.98
IMTA	85.0	11.1	100.2	12.4	22	15.14	3.99	<.01	1.29
INV	88.6	12.7	95.9	11.5	22	7.27	2.41	.02	0.60
ICG	87.0	10.9	96.4	12.2	22	9.45	2.75	.01	0.82
ICC	87.2	12.0	96.8	10.0	22	9.64	4.06	<.01	0.87

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Se observan diferencias significativas en el ICV, el IRF y el IMT, siendo este último el que muestra el mayor tamaño del efecto. Asimismo se observan tamaños del efecto significativos en las puntuaciones de los índices secundarios nuevos de la WISC-V: IRC, IMTA, INV, ICG e ICC, siendo el IMTA el que muestra el mayor tamaño del efecto. Estos resultados son similares a los obtenidos por el grupo de niños con trastorno específico del aprendizaje-lectura, en los que las tareas de memoria de trabajo también presentan tamaños del efecto grandes.

Las puntuaciones obtenidas por el grupo con trastorno específico del aprendizaje-lectura y expresión escrita son más bajas en las pruebas *Semejanzas, Información, Balanzas, Aritmética, Dígitos, Span de dibujos* y *Letras y números* que en el grupo control. De todas las pruebas, *Dígitos* y *Letras y números* muestran los mayores tamaños del efecto. Estos resultados indican dificultades significativas con la memoria de trabajo.

Trastorno específico del aprendizaje-matemáticas

Las dificultades matemáticas y el bajo rendimiento en matemáticas son problemas comunes en la infancia (Geary, 2011b). Hasta un 7 % de los niños tienen un trastorno de las matemáticas, mientras que otro 10 % muestran un bajo rendimiento persistente en matemáticas (Geary, 2011b). Aunque las investigaciones acerca del trastorno de las matemáticas son menos extensivas que las centradas en el trastorno de la lectura, se ha demostrado que ambos trastornos del aprendizaje tienen en común determinadas dificultades cognitivas, incluidas la comprensión verbal, la memoria de trabajo y la velocidad de procesamiento (Willcutt et al., 2013). Geary (2011a) señala que el funcionamiento cognitivo general, la velocidad de procesamiento y los componentes de la memoria de trabajo son predictores longitudinales del rendimiento en matemáticas. Asimismo, las destrezas numéricas tempranas y las habilidades de razonamiento conceptual predicen el rendimiento en matemáticas (Fuchs, Geary, Compton, Fuchs, Hamlett y Bryant, 2010); y el lenguaje, el razonamiento no verbal y la atención están significativamente relacionados con el rendimiento en problemas matemáticos verbales (Fuchs, Geary, Compton, Fuchs, Hamlett, Seethaler et al., 2010; Tolar et al., 2012). Aunque el funcionamiento cognitivo general es un predictor del rendimiento en matemáticas en los niños con un desarrollo normal, no es una causa primaria del trastorno de las matemáticas (Geary, 2011b). Las dificultades en matemáticas se relacionan con las dificultades en la memoria de trabajo (Geary, 2010), la atención (Raghubar et al., 2009) y la recuperación semántica y las destrezas visoespaciales (Cirino, Morris y Morris, 2007). La WISC-V mide habilidades que pueden ser útiles en la identificación de dificultades cognitivas asociadas con el rendimiento en matemáticas y el trastorno de las matemáticas.

Se aplicó la WISC-V a 28 niños de 9 a 16 años, que habían sido identificados como sujetos con trastorno específico del aprendizaje-matemáticas de acuerdo con los criterios del DSM-5. La tabla 6.20 presenta las medias de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas de la WISC-V, para el grupo de niños con trastorno específico del aprendizaje-matemáticas y para el grupo control correspondiente.

Tabla 6.20. Resultados del grupo con trastorno específico del aprendizaje-matemáticas y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Trastorno específico del aprendizaje-matemáticas		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	8.1	2.3	10.3	3.3	27	2.11	3.07	<.01	0.74
Vocabulario	8.1	3.1	9.4	3.1	28	1.32	2.10	.04	0.43
Información	7.5	2.3	10.3	3.7	27	2.78	3.91	<.01	0.90
Comprensión	8.0	2.7	10.4	3.0	28	2.43	3.79	<.01	0.85
Cubos	7.1	2.3	9.8	3.3	28	2.64	3.48	<.01	0.93
Puzles visuales	7.6	2.6	10.3	2.6	28	2.64	4.07	<.01	1.02
Matrices	7.2	3.4	9.6	3.3	28	2.36	2.61	.01	0.70
Balanzas	6.5	3.2	9.4	2.9	28	2.82	3.73	<.01	0.92
Aritmética	6.4	2.4	9.3	3.2	28	2.86	3.82	<.01	1.01
Dígitos	7.9	2.7	9.9	3.5	28	1.93	2.17	.04	0.62
Span de dibujos	8.2	2.7	9.4	4.1	28	1.18	1.29	.21	0.34
Letras y números	7.9	2.1	9.7	3.7	27	1.81	2.21	.04	0.60
Claves	7.6	2.6	9.8	3.7	28	2.14	2.68	.01	0.67
Búsqueda de símbolos	8.8	2.9	9.4	2.9	28	0.57	0.87	.39	0.20
Cancelación	10.3	2.9	10.6	3.2	28	0.32	0.44	.66	0.10
Cs	7.2	2.3	10.1	3.5	28	2.89	3.59	<.01	0.98
Cp	6.5	2.0	9.9	3.1	28	3.39	4.67	<.01	1.30
Dd	8.8	2.8	9.8	3.7	28	1.04	1.12	.27	0.32
Di	8.3	2.7	9.8	3.3	28	1.54	1.82	.08	0.51
Dc	8.1	2.7	9.9	3.0	28	1.75	2.18	.04	0.61
CAa	10.2	2.8	10.4	3.0	28	0.21	0.31	.76	0.07
CAe	10.3	2.8	10.5	3.2	28	0.21	0.28	.78	0.07
ICV	90.3	13.7	99.5	16.4	27	9.19	2.80	<.01	0.61
IVE	85.4	12.6	100.0	15.3	28	14.61	4.01	<.01	1.04
IRF	82.2	15.4	96.7	16.2	28	14.46	3.63	<.01	0.91
IMT	88.7	13.5	97.7	20.4	28	9.00	1.91	.07	0.52
IVP	90.2	14.2	97.7	15.6	28	7.46	2.24	.03	0.50
CIT	83.6	11.9	98.4	16.4	27	14.85	4.44	<.01	1.04
IRC	79.9	13.7	96.2	16.2	28	16.29	4.16	<.01	1.09
IMTA	88.3	11.2	99.1	19.2	27	10.78	2.43	.02	0.69
INV	81.5	13.7	97.6	15.9	28	16.11	4.69	<.01	1.09
ICG	84.2	12.0	98.6	16.5	27	14.44	4.26	<.01	1.00
ICC	87.3	13.9	97.0	16.6	28	9.71	2.55	.02	0.63

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Las medias de los índices primarios y secundarios para el grupo con trastorno específico del aprendizaje-matemáticas fueron significativamente inferiores a las del grupo control, excepto para el IMT. Los mayores tamaños del efecto se observan en el IRC, el INV y el IVE, puntuaciones compuestas nuevas en la WISC-V; ello indica dificultades con las aptitudes de razonamiento cuantitativo, conceptual y espacial.

Todas las puntuaciones de las pruebas son significativamente más bajas en el grupo con trastorno específico del aprendizaje-matemáticas, a excepción de *Span de dibujos, Búsqueda de símbolos* y *Cancelación*. Los tamaños del efecto más grandes se observaron en *Aritmética* y *Puzles visuales*. En conjunto, los resultados sugieren que las dificultades más significativas se encuentran en el razonamiento cuantitativo, conceptual y espacial; y la memoria de trabajo verbal.

Niños con trastorno por déficit de atención con hiperactividad

El trastorno por déficit de atención con hiperactividad (TDAH) es un trastorno común del neurodesarrollo que afecta aproximadamente al 5 % de los niños (American Psychiatric Association, 2013; Polanczyk, de Lima, Horta, Biederman y Rohde, 2007). Se caracteriza por la existencia de dificultades conductuales persistentes en relación con la atención y/o la hiperactividad-impulsividad. Las dificultades generales con la atención y la inhibición coinciden considerablemente con los conceptos neuropsicológicos de *memoria de trabajo* y *función ejecutiva* (Barkley, 2007; Hale, Fiorello y Brown, 2005; Hale et al., 2012; Martinussen, Hayden, Hogg-Johnson y Tannock, 2005).

Las puntuaciones tradicionales del CI total generalmente no han servido para discriminar a los niños o adultos con TDAH de la población no clínica; no obstante, algunos estudios indican que el funcionamiento intelectual global de los niños o adultos con TDAH refleja un leve deterioro (Hale et al., 2012). Las puntuaciones obtenidas por los niños con TDAH muestran que las aptitudes de la comprensión verbal y del razonamiento perceptivo se mantienen relativamente intactas, y que el rendimiento es más bajo en la memoria de trabajo y la velocidad de procesamiento (Hale et al., 2012; Mayes, Calhoun, Chase, Mink y Stagg, 2009; Mayes, Calhoun, Mayes y Molitoris, 2012; Wakkinen, 2008; Wechsler, 2012; Zieman, 2010). Estos puntos débiles se reflejan en puntuaciones más bajas en el ICC que en el ICG, aunque esta discrepancia también se observa en otros grupos (Devena y Watkins, 2012).

En cuanto a las pruebas, el rendimiento en *Claves* es inferior en los niños con TDAH, resultado que se atribuye a un tiempo de elección de la respuesta más lento, no a la velocidad grafomotora; ello sugiere que la eficiencia de la velocidad de respuesta es menor (Jacobson et al., 2011; Metin et al., 2013). El rendimiento en medidas de denominación verbal rápida indica tiempos de respuesta más lentos y puntuaciones más bajas en los niños con TDAH (Elliott, 2007; Jacobson et al., 2011; Korkman et al., 2007). En diversos estudios se ha observado que los tiempos de respuesta son más lentos y la variabilidad en el tiempo de respuesta es mayor en los grupos con TDAH respecto a los niños con un desarrollo normal (Chiang, Huang, Gau y Shang, 2013; Crosbie, et al., 2013; Lipszyc y Schachar, 2010; Rosch, Dirlikov y Mostofsky, 2013). Es más, Gau y Huang (2014) señalan que los sujetos con TDAH cometen más errores de omisión y de comisión, además de requerir un mayor tiempo de latencia, en tareas de atención sostenida y de tiempo de reacción que los grupos control.

Se aplicó la WISC-V a 48 niños de 6 a 16 años, que habían sido previamente identificados como niños con TDAH, de acuerdo con los criterios diagnósticos del DSM-5 y los resultados significativos de las *Brown Attention-Deficit Disorder Scales for Children and Adolescents* (Brown, 2001) en su forma para los padres. El grupo de niños con TDAH incluía todos los subtipos de este trastorno (tipo con predominio del déficit de atención, tipo con predominio hiperactivo-impulsivo y tipo combinado). Los niños del grupo con TDAH debieron estar un mínimo de 24 horas antes de la aplicación de la WISC-V sin tomar medicación psicoestimulante. Asimismo debían haber obtenido previamente una puntuación de la aptitud general (p. ej., el CI total) igual o superior a 80 en alguna medida estandarizada de la aptitud intelectual y de aplicación individual, o bien disponer de un informe en el que se estimara una puntuación de la aptitud general igual o superior a 80. La tabla 6.21 presenta las medias de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas de la WISC-V, para el grupo de niños con TDAH y para el grupo control correspondiente.

Aunque todas las medias de los índices primarios obtenidas por el grupo con TDAH se sitúan en el rango promedio, se observan diferencias significativas entre las puntuaciones obtenidas por el grupo con TDAH en el ICV, el IMT y el IVP y las obtenidas por el grupo control. La media más baja de todos los índices primarios para el grupo con TDAH se observa en el IVP, aunque es el IMT el que muestra el mayor tamaño del efecto. Las puntuaciones de los índices secundarios son más bajas en el grupo con TDAH que en el grupo control, y muestran tamaños del efecto moderados.

Los resultados obtenidos en varias pruebas muestran diferencias estadísticamente significativas entre el grupo con TDAH y el grupo control. Los mayores tamaños del efecto se observan en *Aritmética* y *Claves*; y *Vocabulario*, *Matrices* y *Dígitos* también muestran tamaños del efecto significativos. Estos resultados indican dificultades en la memoria de trabajo y la velocidad de procesamiento grafomotor.

Tabla 6.21. Resultados del grupo con TDAH y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Trastorno por déficit de atención/hiperactividad		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	9.6	2.2	10.3	3.0	48	0.67	1.27	.21	0.25
Vocabulario	9.6	2.4	10.7	2.4	48	1.17	2.41	.02	0.49
Información	9.6	2.7	10.3	2.7	48	0.71	1.34	.19	0.26
Comprensión	9.5	2.8	10.3	2.7	48	0.79	1.61	.11	0.29
Cubos	9.4	3.1	9.8	2.4	48	0.35	0.70	.49	0.13
Puzles visuales	9.6	3.3	10.8	2.7	48	1.13	1.94	.06	0.37
Matrices	9.4	2.5	10.4	2.4	48	1.02	2.14	.04	0.42
Balanzas	9.8	3.0	10.5	3.0	48	0.73	1.18	.24	0.24
Aritmética	8.5	2.6	10.6	2.4	48	2.19	4.31	<.01	0.88
Dígitos	9.2	2.6	10.5	2.3	47	1.23	2.99	<.01	0.50
Span de dibujos	8.9	2.8	10.1	2.8	48	1.23	2.08	.04	0.44
Letras y números	9.0	3.0	10.0	2.7	47	1.02	1.64	.11	0.36
Claves	8.4	2.7	10.0	2.5	48	1.69	3.48	<.01	0.65
Búsqueda de símbolos	9.5	3.0	9.9	2.7	47	0.36	0.57	.57	0.13
Cancelación	9.5	2.6	10.0	2.8	48	0.52	0.99	.33	0.19
Cs	9.5	3.0	9.8	2.3	48	0.33	0.67	.51	0.12
Cp	9.6	2.9	9.7	2.4	47	0.06	0.13	.90	0.02
Dd	9.7	3.1	10.4	2.5	48	0.71	1.26	.21	0.25
Di	9.2	3.0	10.6	2.7	48	1.40	2.84	<.01	0.49
Dc	9.4	2.2	10.3	3.0	47	0.91	1.99	.05	0.35
CAa	10.0	2.5	10.5	2.9	48	0.50	0.94	.35	0.18
CAe	9.1	2.8	9.6	2.8	48	0.54	1.02	.31	0.19
ICV	97.8	11.4	102.7	13.2	48	4.90	1.98	.05	0.40
IVE	97.3	16.7	101.5	12.4	48	4.17	1.51	.14	0.28
IRF	97.6	13.4	102.6	13.1	48	5.04	1.94	.06	0.38
IMT	94.8	13.3	101.7	12.3	47	6.91	2.98	<.01	0.54
IVP	94.2	13.9	99.9	12.5	47	5.70	2.19	.03	0.43
CIT	95.6	11.7	102.2	10.2	47	6.66	3.20	<.01	0.61
IRC	94.8	14.2	103.1	12.7	48	8.33	3.13	<.01	0.62
IMTA	95.2	13.3	101.4	11.5	47	6.23	2.73	<.01	0.50
INV	94.4	13.1	101.7	12.3	48	7.29	3.01	<.01	0.57
ICG	97.1	13.3	102.3	11.1	48	5.21	2.23	.03	0.43
ICC	92.8	12.6	100.8	12.2	46	8.00	3.18	<.01	0.65

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Niños con comportamiento disruptivo

El comportamiento disruptivo puede afectar al rendimiento académico de un niño, así como a su desarrollo social (Rudy y Levinson, 2008). Parte de estos niños reciben, además, un diagnóstico de trastorno negativista desafiante y/o trastorno de la conducta. El trastorno negativista desafiante y, en menor medida, el trastorno de la conducta son comunes en la infancia y la adolescencia, con unas tasas de prevalencia de vida estimadas del 10 % para el trastorno negativista desafiante Nock, Kazdin, Hiripi y Kessler, 2007) y del 3.4 % para el trastorno de la conducta (Morcillo et al., 2012). Además, estos trastornos se presentan en coocurrencia con el TDAH (van Lier, van der Ende, Koot y Verhulst, 2007).

Los niños con comportamiento disruptivo muestran distintas conductas inapropiadas que afectan negativamente a su rendimiento escolar y a sus relaciones interpersonales. Un estudio longitudinal indica que los déficits del lenguaje que no son debidos a un deterioro intelectual son un factor de riesgo independiente para el desarrollo de problemas de comportamiento disruptivo (Petersen et al., 2013). Otros estudios han señalado que las aptitudes verbales bajas (Närhi, Lehto-Salo, Ahonen y Marttunen, 2010), los déficits del vocabulario receptivo y la aptitud cognitiva general ligeramente baja (Barker et al., 2011; Dougherty et al., 2007; Isen, 2010; Séguin, Parent, Tremblay y Zelazo, 2009) se asocian con el comportamiento disruptivo. Los niños con trastornos por comportamiento disruptivo tienen dificultades cognitivas similares a las observadas en niños diagnosticados con TDAH, incluidos los problemas con el control inhibitorio (Rubia, 2011; Schoemaker et al., 2012; Skogan et al., 2013), la atención sostenida, la flexibilidad cognitiva (Rubia, 2011) y la memoria de trabajo (Rhodes, Park, Seth y Coghill, 2012; Saarinen, Fontell, Vuontela, Carlson y Aronen, 2014; Skogan et al., 2013). Así como se da una considerable comorbilidad entre el TDAH y los trastornos por comportamiento disruptivo, los déficits cognitivos identificados en niños diagnosticados con trastorno negativista desafiante o trastorno de la conducta se observan independientemente de un diagnóstico de TDAH (Rhodes et al., 2012; Rubia, 2011; Saarinen et al., 2014; Schoemaker et al., 2012; Skogan et al., 2013). El funcionamiento cognitivo de niños con problemas por comportamiento disruptivo es considerablemente heterogéneo (Närhi et al., 2010; Rubia, 2011), y los déficits en el funcionamiento cognitivo pueden relacionarse con un subtipo específico de estos trastornos (Boylan, 2014). La investigación ha señalado que los niños con problemas de conducta significativos pueden tener intactas las funciones ejecutivas (Drabick, Bubier, Chen, Price y Lanza, 2011). Asimismo sugiere que algunos niños con problemas por comportamiento disruptivo mostrarán dificultades cognitivas comórbidas. Cuando existen problemas cognitivos, se dan con más frecuencia en los dominios del lenguaje, la memoria de trabajo y el funcionamiento ejecutivo.

Se aplicó la WISC-V a 21 niños de 6 a 15 años y con comportamiento disruptivo. Los niños del grupo con comportamiento disruptivo habían sido diagnosticados con trastorno negativista desafiante o con trastorno de la conducta, o bien se les había determinado un alto grado de problemas de conducta a partir de los resultados obtenidos en la BASC-2, en su forma para los padres (p. ej., puntuación $T > 70$). Los niños con un diagnóstico comórbido de TDAH o con síntomas de TDAH no se incluyeron en el estudio de la WISC-V. La tabla 6.22 presenta las medias de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas de la WISC-V, para el grupo de niños con comportamiento disruptivo y para el grupo control correspondiente.

Todas las medias de las puntuaciones de los índices primarios obtenidas por el grupo con comportamiento disruptivo se sitúan en el rango promedio. No se observan diferencias significativas entre el grupo con comportamiento disruptivo y el grupo control en ningún índice ni en ninguna prueba. La media de puntuación compuesta más baja para el grupo con comportamiento disruptivo es de 92.8 tanto en el IVP como en el ICC. Como grupo, los niños con comportamiento disruptivo no son significativamente distintos de los niños del grupo control; sin embargo, el nivel de rendimiento cognitivo en el grupo con comportamiento disruptivo es similar al del grupo con TDAH visto anteriormente.

Tabla 6.22. Resultados del grupo con comportamiento disruptivo y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Comportamiento disruptivo		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	9.4	2.6	10.4	3.5	21	1.00	1.10	.28	0.32
Vocabulario	8.4	2.3	8.9	2.8	21	0.52	0.89	.38	0.20
Información	9.0	3.0	9.9	2.5	21	0.90	1.17	.25	0.33
Comprensión	9.3	3.3	9.5	3.1	21	0.19	0.18	.86	0.06
Cubos	9.5	2.7	8.2	2.9	21	−1.24	−1.34	.20	−0.44
Puzles visuales	9.5	2.7	9.1	2.6	21	−0.43	−0.57	.58	−0.16
Matrices	9.2	3.1	9.4	2.7	21	0.19	0.26	.80	0.07
Balanzas	8.8	2.8	9.2	2.8	21	0.38	0.50	.62	0.14
Aritmética	9.0	3.1	9.3	3.3	21	0.29	0.31	.76	0.09
Dígitos	9.3	3.4	9.4	3.1	21	0.10	0.09	.93	0.03
Span de dibujos	9.0	3.0	8.8	2.6	21	−0.24	−0.24	.81	−0.09
Letras y números	8.7	3.8	9.3	2.2	21	0.62	0.55	.59	0.20
Claves	8.7	3.2	8.2	3.4	21	−0.52	−0.55	.59	−0.16
Búsqueda de símbolos	8.8	3.2	9.0	3.1	21	0.24	0.24	.81	0.08
Cancelación	9.6	2.4	9.4	2.4	21	−0.19	−0.24	.82	−0.08
Cs	9.4	3.0	8.1	3.1	21	−1.24	−1.22	.24	−0.41
Cp	9.8	2.9	8.1	3.2	21	−1.67	−1.76	.09	−0.55
Dd	9.7	2.3	9.6	4.2	21	−0.14	−0.15	.89	−0.04
Di	9.6	4.0	10.0	2.7	21	0.33	0.28	.78	0.10
Dc	8.9	3.3	9.1	2.4	21	0.24	0.24	.81	0.08
CAa	9.7	2.6	9.5	2.3	21	−0.19	−0.25	.81	−0.08
CAe	9.6	2.0	9.7	2.9	21	0.14	0.17	.87	0.06
ICV	94.1	11.8	98.2	13.3	21	4.14	1.27	.22	0.33
IVE	97.1	13.9	92.6	13.3	21	−4.52	−1.10	.29	−0.33
IRF	94.4	15.2	95.9	13.3	21	1.52	0.44	.66	0.11
IMT	95.3	13.7	95.0	13.2	21	−0.29	−0.06	.95	−0.02
IVP	92.8	17.1	92.1	16.5	21	−0.62	−0.12	.91	−0.04
CIT	93.3	12.4	93.5	12.8	21	0.19	0.06	.96	0.02
IRC	93.8	13.7	95.5	13.7	21	1.71	0.47	.64	0.12
IMTA	94.5	18.0	96.6	13.1	21	2.10	0.39	.70	0.13
INV	93.6	12.5	91.3	12.7	21	−2.29	−0.73	.48	−0.18
ICG	94.1	12.0	95.1	13.2	21	0.95	0.31	.76	0.08
ICC	92.8	14.3	92.6	14.7	21	−0.24	−0.05	.96	−0.02

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Niños con traumatismo craneoencefálico

El traumatismo craneoencefálico es una de las principales causas de muerte y discapacidad en niños y adolescentes (Keenan y Bratton, 2006; McCrea, Janecek, Powell y Hammeke, 2014). El traumatismo craneoencefálico es muy variable en los efectos que causa, y puede asociarse con numerosos deterioros cognitivos y conductuales (Lezak et al., 2004; Li y Liu, 2013). Las secuelas cognitivas específicas de un traumatismo craneoencefálico varían y se relacionan con la localización, gravedad y características propias de la lesión (p. ej., abierta o cerrada, focal o difusa), así como con la edad y el funcionamiento premórbido del niño (Allen et al., 2010; Anderson, Catroppa, Godfrey y Rosenfeld, 2012; Donders y Nesbit-Greene, 2004; Gorman, Barnes, Swank, Prasad y Ewing-Cobbs, 2012; Lezak et al., 2004; Prigatano, Gray y Gale, 2008; Recla et al., 2013). Los lóbulos frontal y temporal anterior son particularmente vulnerables a las contusiones, hemorragias y hematomas relacionados con las fuerzas de aceleración y desacelera-

ción del traumatismo craneal cerrado (Mattson y Levin, 1990). El daño en el lóbulo frontal se asocia principalmente a una disfunción ejecutiva. Cuando la lesión es grave, la disfunción puede ser incapacitante porque deteriora de manera importante la flexibilidad cognitiva del sujeto en la resolución de problemas o la adaptabilidad (Lezak et al., 2004).

Las conclusiones de estudios previos sobre el traumatismo craneoencefálico se ven en parte limitadas por la variabilidad inherente a esta lesión, así como por inconsistencias metodológicas (p. ej., definición de gravedad, variabilidad en el tiempo transcurrido desde que se produjo el traumatismo), un conocimiento o un control inadecuados del funcionamiento intelectual premórbido y un control poco preciso de las funciones de la anatomía cerebral (Kaufman, 1994; Tremont, Mittenberg y Miller, 1999). Los efectos a largo plazo del traumatismo craneoencefálico son muy variables; al respecto, las investigaciones señalan que el traumatismo craneoencefálico se relaciona con déficits cognitivos en áreas como las siguientes:

- lenguaje (Dikmen, Machamer, Winn y Temkin, 1995; Green, Rosenfield y Catroppa, 2008),
- atención (Gualtieri y Johnson, 2006; Mathias y Wheaton, 2007),
- memoria de trabajo (Anderson, Catroppa et al., 2012; Anderson, Godfrey, Rosenfield y Catroppa, 2012; Donders y Nesbit-Greene, 2004; Newman, Reesman, Vaughan y Gioia, 2013),
- aprendizaje y memoria (Babikian et al., 2011; Carlozzi, Grech y Tulsky, 2013; De Monte, Geffen, May y McFarland, 2010; Newman et al., 2013),
- velocidad de procesamiento (Anderson, Catroppa et al., 2012; Anderson, Godfrey et al., 2012; Ben-David, Nguyen y van Lieshout, 2011; Rackley, Allen, Fuhrman y Mayfield, 2012; Recla et al., 2013),
- visopercepción y procesamiento (Donders y Nesbit-Greene, 2004; Donders, Tulsky y Zhu, 2001; Green et al., 2008; Hiscock, Inch y Gleason, 2002; Rackley et al., 2012),
- destrezas visoespaciales y visomotoras (Donders et al., 1997; Green et al., 2008; Hiscock et al., 2000), y
- razonamiento abstracto (Dikmen et al., 1995; Green et al., 2008).

Se aplicó la WISC-V a 20 niños de 7 a 16 años, identificados con un historial de traumatismo craneoencefálico moderado o grave. Solo participaron niños que habían sufrido el traumatismo craneoencefálico, abierto o cerrado, entre 6 y 18 meses antes de la fecha de aplicación, y se excluyó a los niños cuyas puntuaciones de inteligencia premórbida se situaban en el rango definido para los niños con discapacidad intelectual (es decir, CIT < 70). Los niños del grupo con traumatismo craneoencefálico podían presentar diagnósticos comórbidos de TDAH, trastorno de la conducta, ansiedad o trastornos del ánimo, así como estar recibiendo servicios logopédicos por dificultades comunicativas relacionadas con el traumatismo craneoencefálico. La tabla 6.23 presenta las medias y las desviaciones típicas de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas de la WISC-V, para el grupo de sujetos con traumatismo craneoencefálico y para el grupo control correspondiente.

Las diferencias en las medias de las puntuaciones de los índices primarios obtenidas por ambos grupos son significativas y muestran tamaños del efecto grandes para el IVE, el IRF y el IMT, observándose la diferencia más grande en el IRF. El índice primario con la puntuación más baja es el IVP. En todas las puntuaciones de los índices secundarios se observan diferencias significativas con tamaños del efecto grandes.

Respecto a las pruebas, se observan tamaños del efecto grandes en *Información, Cubos, Puzles visuales, Matrices, Balanzas, Aritmética, Dígitos, Span de dibujos* y *Claves*; y tamaños del efecto moderados en *Semejanzas, Vocabulario, Comprensión, Letras y números* y *Búsqueda de símbolos*. La mayoría de estos resultados son consistentes con las indicaciones de los estudios anteriormente descritos, observándose tamaños del efecto más grandes en las medidas de percepción y procesamiento visual y de memoria de trabajo en comparación con las medidas de comprensión verbal (Lezak et al., 2004; Tremont et al., 1999). Las puntuaciones escalares más bajas son las de *Aritmética, Claves* y *Búsqueda de símbolos*.

Tabla 6.23. Resultados del grupo con traumatismo craneoencefálico y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Traumatismo craneoencefálico		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	8.5	2.6	9.6	2.0	17	1.18	1.47	.16	0.51
Vocabulario	8.1	3.1	10.1	2.7	19	2.00	2.76	.01	0.69
Información	7.9	2.8	10.2	2.6	20	2.25	3.10	<.01	0.83
Comprensión	8.1	2.4	9.5	3.1	20	1.40	1.71	.10	0.51
Cubos	7.7	3.1	11.1	3.6	19	3.37	3.00	<.01	1.00
Puzles visuales	7.6	3.4	10.5	2.6	20	2.85	2.72	.01	0.94
Matrices	8.2	3.4	11.5	2.7	20	3.35	3.19	<.01	1.09
Balanzas	7.8	3.6	10.9	3.1	20	3.10	2.98	<.01	0.92
Aritmética	7.3	3.0	9.8	1.8	20	2.50	2.86	.01	1.01
Dígitos	7.7	3.1	10.4	3.0	20	2.65	2.74	.01	0.87
Span de dibujos	7.6	2.6	9.8	2.6	20	2.25	2.64	.02	0.87
Letras y números	7.9	3.1	9.8	2.0	20	1.90	2.93	<.01	0.73
Claves	7.2	4.0	9.8	2.0	19	2.58	2.71	.01	0.82
Búsqueda de símbolos	7.4	3.9	9.1	2.4	19	1.74	1.67	.11	0.54
Cancelación	8.5	3.8	10.1	3.0	19	1.53	1.08	.29	0.45
Cs	7.9	3.8	11.2	3.8	17	3.24	2.48	.02	0.85
Cp	7.6	3.3	10.6	3.8	17	3.06	2.49	.02	0.86
Dd	8.2	2.6	10.6	3.6	20	2.40	2.82	.01	0.76
Di	8.2	3.4	10.0	3.5	20	1.80	1.73	.10	0.52
Dc	8.6	3.5	10.1	2.0	20	1.50	1.73	.10	0.53
CAa	8.8	4.0	10.0	3.3	19	1.16	0.76	.45	0.32
CAe	8.6	3.4	10.2	3.2	19	1.53	1.24	.23	0.46
ICV	88.9	12.9	98.3	11.7	17	9.41	2.51	.02	0.76
IVE	87.5	15.9	104.2	14.6	19	16.68	2.99	<.01	1.09
IRF	88.4	18.0	106.9	13.6	20	18.45	3.56	<.01	1.16
IMT	86.2	15.5	100.2	12.8	20	14.00	2.94	<.01	0.98
IVP	84.1	22.2	97.2	8.2	19	13.05	2.55	.02	0.78
CIT	83.3	14.1	103.3	13.0	17	20.00	4.10	<.01	1.47
IRC	85.6	17.5	101.7	11.7	20	16.05	3.31	<.01	1.08
IMTA	87.7	16.3	100.3	11.9	20	12.65	3.07	<.01	0.89
INV	85.0	17.3	104.3	11.8	19	19.32	3.57	<.01	1.30
ICG	85.4	13.5	103.9	12.1	17	18.53	4.15	<.01	1.45
ICC	83.5	18.1	98.2	10.0	19	14.68	3.13	<.01	1.00

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Niños con trastorno del espectro autista

Con la publicación del DSM-5, el trastorno autista y el trastorno de Asperger se conceptualizan como un único diagnóstico: el trastorno del espectro autista. Los niños diagnosticados con trastorno del espectro autista se caracterizan por presentar déficits en la comunicación verbal y no verbal y en la interacción y comunicación social, asimismo muestran patrones de conducta, intereses y actividades restringidos. Para describir más detalladamente la sintomatología, pueden usarse indicadores específicos, incluidas la gravedad de los síntomas, la existencia de deterioro intelectual o lingüístico y la presencia de otros factores médicos, genéticos o ambientales, así como de trastornos del neurodesarrollo, mentales o conductuales. Se aplicó la WISC-V a dos grupos de niños con trastorno del espectro autista: un grupo con deterioro del lenguaje acompañante (anteriormente considerado *trastorno autista*) y otro grupo sin deterioro del lenguaje acompañante (anteriormente considerado *trastorno de Asperger*).

Trastorno del espectro autista con deterioro del lenguaje acompañante

En investigaciones anteriores se sugiere que el funcionamiento intelectual general de los niños con trastorno del espectro autista es más bajo que el de los grupos control; no obstante, estos estudios también sugieren que existe un patrón de puntos fuertes y débiles. Distintos estudios demuestran que las puntuaciones en medidas del funcionamiento intelectual general son más bajas, pero que el rendimiento en medidas del razonamiento fluido es relativamente mejor (Dawson, Soulières, Gernsbacher y Mottron, 2007; Mayes y Calhoun, 2008; Stevenson, 2011). Por lo general, el rendimiento en las tareas verbales es más bajo en la mayoría de los niños con trastorno del espectro autista con deterioro del lenguaje acompañante en comparación con los niños con un desarrollo normal (Joseph, Tager-Flusberg y Lord, 2002; Klinger, O'Kelley, Mussey, Goldstein y DeVries, 2012; Mayes y Calhoun, 2008; Wechsler, 2003, 2012). Las pruebas de comprensión verbal reflejan un típico patrón de rendimiento: la puntuación más alta se obtiene en *Semejanzas,* prueba en la que interviene el razonamiento fluido; y la puntuación más baja se obtiene en *Comprensión,* prueba que requiere cierto nivel de juicio social, un punto débil característico de los sujetos con trastorno del espectro autista (Mayes y Calhoun, 2008; Zayat, Kalb y Wodka, 2011). Además, algunos estudios señalan relativos puntos fuertes de los niños con trastorno autista en las tareas visoespaciales (Mayes y Calhoun, 2008; Wechsler, 2003, 2012).

Se aplicó la WISC-V a 30 niños de 6 a 16 años, que habían sido diagnosticados como sujetos con trastorno del espectro autista con deterioro del lenguaje de acuerdo con los criterios del DSM-5. Se excluyó del estudio a los niños con trastorno autista que habían obtenido puntuaciones en pruebas de aptitud cognitiva general superiores a 2.67 desviaciones típicas por debajo de la media (p. ej., CI total < 60), así como a los niños que carecían de las destrezas comunicativas necesarias para poder completar el test. La sobrerrepresentación de niños respecto a las niñas refleja el efecto combinado de la mayor tasa de prevalencia de varones respecto a mujeres en el trastorno autista con deterioro del lenguaje, así como la mayor prevalencia de discapacidad intelectual en las mujeres con trastorno autista (American Psychiatric Association, 2013). La tabla 6.24 presenta las medias de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas de la WISC-V, para el grupo de niños con trastorno autista con deterioro del lenguaje y para el grupo control correspondiente.

Para el grupo con trastorno autista con deterioro del lenguaje, todas las medias de los índices primarios son significativamente más bajas que las medias correspondientes del grupo control, y los tamaños del efecto son grandes. El IRF y el IVE muestran puntuaciones relativamente más altas y producen tamaños del efecto más pequeños que el ICV; estos resultados son consistentes con los de estudios previos. De entre las puntuaciones de los índices primarios, el ICV produce el segundo tamaño del efecto más grande, siendo el mayor el del IMT. Estos valores concuerdan con las investigaciones que indican la existencia de un punto débil en el funcionamiento ejecutivo, la memoria de trabajo y la comprensión verbal en los niños con trastorno del espectro autista (Boucher y Mayes, 2012; Corbett, Constantine, Hendren, Rocke y Ozonoff, 2009; Englund, Decker, Allen y Roberts, 2014; Mayes y Calhoun, 2007).

En cuanto a las pruebas, todas las medias de las puntuaciones escalares son significativamente más bajas en el grupo con trastorno autista con deterioro del lenguaje que en el grupo control. El tamaño del efecto más grande se observa en *Comprensión,* a la que le siguen *Letras y números, Aritmética, Información* y *Dígitos*; y los tamaños del efecto más pequeños se observan en *Balanzas, Matrices* y *Cubos.*

Todas las puntuaciones de los índices secundarios son significativamente inferiores a las puntuaciones correspondientes obtenidas por el grupo control, y los tamaños del efecto son grandes. Las medias de puntuaciones más bajas y los tamaños del efecto más pequeños se observan en el IMTA y el ICC, índices a los que contribuye, como mínimo, una prueba de memoria de trabajo auditiva.

Estos resultados coinciden con los de investigaciones anteriores que indican déficits cognitivos globales, un rendimiento relativamente más bajo en las tareas verbales y un rendimiento relativamente más alto en las tareas visoespaciales (Barbeau, Soulières, Dawson, Zeffiro y Mottron, 2013; Klinger et al., 2012; Mayes y Calhoun, 2008; Soulières, Dawson, Gernsbacher y Mottron, 2011; Wechsler, 2003, 2012).

Tabla 6.24. Resultados del grupo con trastorno del espectro autista con deterioro del lenguaje, y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Trastorno del espectro autista con deterioro del lenguaje		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	6.5	3.3	10.9	2.8	29	4.38	5.22	<.01	1.43
Vocabulario	6.1	3.6	10.6	2.6	29	4.48	5.64	<.01	1.43
Información	5.8	3.5	11.2	2.9	29	5.34	6.39	<.01	1.66
Comprensión	4.8	3.0	11.1	3.0	30	6.23	8.56	<.01	2.08
Cubos	7.2	4.1	10.5	2.8	30	3.30	3.67	<.01	0.94
Puzles visuales	6.7	4.3	11.1	2.4	29	4.45	4.86	<.01	1.28
Matrices	7.1	4.2	10.3	2.2	30	3.13	3.77	<.01	0.93
Balanzas	7.4	3.6	10.3	3.5	30	2.93	3.07	<.01	0.83
Aritmética	5.4	3.5	10.6	2.5	30	5.23	6.68	<.01	1.72
Dígitos	5.4	3.9	10.8	2.8	30	5.43	6.37	<.01	1.60
Span de dibujos	6.9	3.8	10.7	3.0	30	3.77	4.07	<.01	1.10
Letras y números	5.0	3.6	10.3	2.3	29	5.24	7.40	<.01	1.73
Claves	5.2	3.5	9.3	3.2	29	4.07	4.94	<.01	1.21
Búsqueda de símbolos	6.2	3.7	9.9	2.7	26	3.69	5.30	<.01	1.14
Cancelación	5.8	3.4	9.7	3.0	29	3.93	4.92	<.01	1.23
Cs	7.5	4.4	10.5	2.8	30	3.00	3.32	<.01	0.81
Cp	7.4	4.2	10.8	2.5	30	3.40	4.26	<.01	0.98
Dd	6.0	3.2	11.0	2.4	30	4.97	6.79	<.01	1.76
Di	6.6	4.0	10.6	2.8	30	3.97	4.56	<.01	1.15
Dc	5.5	3.8	10.8	3.8	30	5.27	6.03	<.01	1.39
CAa	6.4	3.3	9.7	2.6	29	3.24	4.45	<.01	1.09
CAe	5.7	3.7	10.0	3.0	30	4.30	4.89	<.01	1.28
ICV	80.4	18.2	104.1	13.6	28	23.68	5.51	<.01	1.47
IVE	82.8	22.3	104.4	13.4	29	21.62	4.49	<.01	1.18
IRF	84.3	20.6	101.6	13.9	30	17.30	3.75	<.01	0.98
IMT	77.6	19.4	104.1	13.9	30	26.47	6.05	<.01	1.57
IVP	75.8	19.0	96.9	14.8	26	21.12	5.41	<.01	1.24
CIT	76.3	19.1	102.1	14.5	28	25.82	5.75	<.01	1.52
IRC	78.9	19.2	102.5	15.8	30	23.67	5.07	<.01	1.35
IMTA	72.3	21.6	102.4	12.6	29	30.14	7.14	<.01	1.70
INV	79.9	20.1	102.8	13.8	28	22.86	4.78	<.01	1.33
ICG	81.8	18.6	102.9	14.2	28	21.18	4.82	<.01	1.28
ICC	74.4	18.4	100.0	13.5	26	25.62	6.24	<.01	1.59

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

Trastorno del espectro autista sin deterioro del lenguaje acompañante

Los niños diagnosticados con trastorno del espectro autista sin deterioro del lenguaje acompañante tienden a mostrar menos retrasos en el desarrollo cognitivo relacionado con las destrezas de aprendizaje y con la conducta adaptativa, y mantienen relativamente intacta la aptitud verbal, a diferencia de los niños con trastorno del espectro autista con deterioro del lenguaje. Diversos estudios indican que los sujetos con trastorno del espectro autista sin deterioro del lenguaje muestran puntos débiles en la velocidad de procesamiento y en la memoria de trabajo, pero conservan relativamente intactas las aptitudes de razonamiento perceptivo y comprensión verbal; asimismo señalan que las puntuaciones bajas en la prueba *Comprensión* son habituales en relación con otras pruebas de comprensión verbal (Boucher y Mayes, 2012; Englund et al., 2014; Mayes y Calhoun, 2007, 2008; Mayes et al., 2012; Oliveras-Rentas et al., 2012; Wechsler, 2008, 2012).

Al igual que señalan los estudios sobre el trastorno del espectro autista con deterioro del lenguaje, los sujetos con trastorno del espectro autista sin deterioro del lenguaje parecen tener un relativo punto fuerte en el razonamiento fluido (Mayes y Calhoun, 2008; Oliveras-Rentas et al., 2012); y, de entre las pruebas de Comprensión verbal, obtienen la puntuación más alta en *Semejanzas* (Oliveras-Rentas et al., 2012). Otros estudios también indican que, en estos sujetos, el ICG suele ser más alto que el CI total (Raiford et al., 2005; Wechsler, 2008, 2012). Por todo ello se predijo que los resultados de este estudio serían similares.

Se aplicó la WISC-V a 32 niños de 6 a 16 años, que habían sido diagnosticados como sujetos con trastorno del espectro autista sin deterioro del lenguaje de acuerdo con los criterios del DSM-5. Se excluyó del estudio a los niños con trastorno del espectro autista sin deterioro del lenguaje que habían obtenido puntuaciones en pruebas de aptitud cognitiva general de más de 2 desviaciones típicas por debajo de la media (p. ej., CI total < 70). La tabla 6.25 presenta las medias de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas de la WISC-V, para el grupo de niños con trastorno autista sin deterioro del lenguaje y para el grupo control correspondiente.

Tabla 6.25. Resultados del grupo con trastorno del espectro autista sin deterioro del lenguaje, y del grupo control (datos estadounidenses)

Prueba/ punt. de procesamiento/ punt. compuesta	Trastorno del espectro autista sin deterioro del lenguaje		Grupo control		Diferencia de medias de los grupos				
	Media	*Dt*	Media	*Dt*	*n*	Diferencia	*t*	*p*	Diferencia típica[a]
Semejanzas	10.8	3.2	10.3	2.4	31	−0.55	−0.98	.34	−0.19
Vocabulario	9.9	3.0	11.5	2.6	32	1.56	2.13	.04	0.56
Información	10.5	3.7	10.8	2.5	32	0.31	0.47	.64	0.10
Comprensión	8.9	2.9	10.1	3.2	32	1.19	1.48	.15	0.39
Cubos	9.4	3.6	11.2	2.9	32	1.78	2.34	.03	0.54
Puzles visuales	10.8	2.9	10.0	3.2	32	−0.84	−1.22	.23	−0.28
Matrices	10.2	3.2	10.6	2.2	32	0.47	0.63	.53	0.17
Balanzas	10.2	3.5	10.5	2.9	32	0.38	0.49	.62	0.12
Aritmética	10.4	3.8	10.8	2.8	32	0.31	0.38	.71	0.09
Dígitos	9.2	3.2	10.4	2.3	32	1.22	1.69	.10	0.44
Span de dibujos	9.3	3.4	11.0	2.7	32	1.75	2.74	.01	0.57
Letras y números	8.7	3.7	10.9	3.0	32	2.22	2.61	.01	0.66
Claves	7.8	3.4	10.0	3.0	31	2.23	2.39	.02	0.70
Búsqueda de símbolos	8.4	3.4	9.3	3.0	32	0.91	1.02	.32	0.28
Cancelación	9.8	3.4	9.7	2.8	32	−0.09	−0.11	.91	−0.03
Cs	9.4	3.7	11.3	2.8	32	1.81	2.40	.02	0.55
Cp	9.4	3.7	10.9	3.1	32	1.47	1.80	.08	0.43
Dd	9.1	3.1	10.1	2.2	32	0.94	1.51	.14	0.35
Di	9.4	3.2	10.8	2.5	32	1.34	1.78	.08	0.47
Dc	9.6	3.5	10.3	3.2	32	0.75	0.95	.35	0.22
CAa	10.2	3.1	10.0	3.0	32	−0.19	−0.22	.83	−0.06
CAe	9.3	3.1	9.8	2.5	32	0.41	0.56	.58	0.15
ICV	102.5	14.4	104.8	11.6	31	2.26	0.77	.45	0.17
IVE	100.7	17.1	103.4	14.5	32	2.72	0.77	.44	0.17
IRF	100.9	17.5	103.5	12.7	32	2.53	0.65	.52	0.17
IMT	95.4	16.8	104.3	12.1	32	8.81	2.52	.02	0.60
IVP	89.4	18.4	98.0	14.7	31	8.58	1.79	.08	0.52
CIT	98.3	17.4	105.0	11.9	30	6.63	1.75	.09	0.44
IRC	101.7	19.1	103.5	12.4	32	1.78	0.47	.64	0.11
IMTA	94.1	17.2	103.8	12.2	32	9.78	2.59	.01	0.66
INV	97.5	17.7	104.4	13.9	31	6.87	1.72	.10	0.43
ICG	101.1	17.0	105.5	11.9	31	4.45	1.31	.20	0.30
ICC	91.0	17.9	101.4	13.0	31	10.32	2.35	.03	0.66

[a] La diferencia típica es la diferencia entre las medias de las dos aplicaciones dividida por la raíz cuadrada de la varianza común, calculada mediante la fórmula de Cohen (1996, fórmula 10.4).

En general, la mayoría de las puntuaciones del grupo con trastorno del espectro autista sin deterioro del lenguaje no son significativamente diferentes de las del grupo control y son más altas que las obtenidas por el grupo con trastorno del espectro autista con deterioro del lenguaje. Estos resultados son consistentes con investigaciones anteriores que demuestran que los déficits en los niños con trastorno de Asperger son menos graves que los de los niños con trastorno autista (Barbeau et al., 2013; Gilchrist et al., 2001). De entre las puntuaciones de los índices primarios, solo el IMT muestra una diferencia significativas entre el grupo con trastorno del espectro autista sin deterioro del lenguaje y el grupo control; no obstante, los tamaños del efecto tanto para el IMT como para el IVP son moderados, y la media del IVP es la más baja de todas las puntuaciones de los índices primarios. También se observan diferencias significativas y tamaños del efecto moderados en el IMTA y el ICC, lo que indica un punto débil en la memoria de trabajo del grupo con trastorno del espectro autista sin deterioro del lenguaje. Otros resultados que concuerdan con

estudios previos son que el ICC es más bajo que la media del ICG en los niños con trastorno del espectro autista sin deterioro del lenguaje, y que el ICV, el IVE y el IRF son similares a los del grupo control.

En cuanto a las pruebas, las medias de las puntuaciones escalares de *Vocabulario, Cubos, Span de dibujos, Letras y números* y *Claves* son significativamente más bajas en el grupo con trastorno del espectro autista sin deterioro del lenguaje que en el grupo control. Los mayores tamaños del efecto se observan en *Claves* y *Letras y números*. Las diferencias entre las medias de *Comprensión, Puzles visuales, Dígitos* y *Búsqueda de símbolos* no son estadísticamente significativas aunque ofrezcan tamaños del efecto pequeños. De entre las pruebas de Comprensión verbal, la puntuación más baja es la de *Comprensión* y la más alta es la de *Semejanzas*, resultados consistentes con investigaciones anteriores. Todos estos resultados indican la existencia de déficits en la velocidad de procesamiento y en la memoria de trabajo, así como sutiles dificultades en el lenguaje.

Resumen del rendimiento de los grupos especiales en la WISC-V

Los resultados obtenidos por los grupos especiales respaldan la validez y utilidad clínica de la WISC-V, en especial de las pruebas y los índices nuevos; y generalmente son consistentes con los estudios anteriores y con los fundamentos teóricos. Las próximas investigaciones basadas en la WISC-V y en ambientes y poblaciones clínicas diferentes aportarán más datos sobre la utilidad de la escala como parte de una exhaustiva evaluación clínica llevada a cabo con fines diagnósticos y para definir el tipo de intervención adecuado.

Consecuencias de la aplicación de tests

Braden y Niebling (2012) señalan que en los datos sobre las consecuencias de la aplicación de tests debe incluirse una evaluación de la utilidad clínica diagnóstica individualizada. El diagnóstico o clasificación de un niño no debe basarse solo en los resultados de un único test, sino que también debe fundamentarse en la información recopilada mediante las observaciones, otras modalidades de evaluación y la aplicación de distintos tests relevantes, así como en los datos proporcionados por otras fuentes (padres, maestros, etc.). Los resultados procedentes de estudios con grupos especiales y que tradicionalmente utilizan puntuaciones de corte indican el porcentaje de niños del grupo especial (p. ej., con alta capacidad intelectual, con discapacidad intelectual) que se sitúan en un rango de puntuación específico de la WISC-V (p. ej., CIT $\geq$ 120 o CIT $\leq$ 75).

Braden and Niebling (2012) también afirman que, al analizar las consecuencias de la aplicación de tests, deben considerarse también las consecuencias inesperadas, como el sesgo de los ítems y las diferencias entre las puntuaciones obtenidas por los distintos grupos. El procedimiento seguido para analizar el sesgo de los ítems (funcionamiento diferencial de los ítems) y la adecuación del contenido se describe en el capítulo 2 de este manual. Las escalas de inteligencia de Wechsler han sido objeto de diversos estudios que han analizado las diferencias entre las puntuaciones obtenidas por los distintos sujetos que forman las muestras, prestando especial atención a las agrupaciones por sexo y nivel socioeconómico (p. ej., Harris y Llorente, 2005; Weiss, Chen, Harris, Holdnack y Saklofske, 2010).

Los tests psicológicos y educativos se aplican, por lo general, con el convencimiento de que, a partir de los resultados obtenidos, los sujetos evaluados obtendrán algún beneficio. Además, los tests ayudan a los profesionales y a las instituciones a cumplir con los criterios legales y profesionales que conciernen al diagnóstico y clasificación, al plan de tratamiento o programa de rehabilitación, y a actividades de investigación.

Los profesionales deben tener siempre en cuenta las consecuencias esperadas e inesperadas del hecho de usar tests, así como las consecuencias esperadas e inesperadas de confiar en valoraciones informales en vez de en valoraciones basadas en los resultados de los tests. Aunque las consecuencias de la aplicación de tests pueden influir en la decisión sobre el uso de un test (si utilizarlo o no), las consecuencias adversas en sí mismas no disminuyen la validez de las interpretaciones del test previstas (AERA , APA, NCME, 1999).

Capítulo 7

Interpretación de los resultados

Los resultados de la WISC-V proporcionan información importante sobre el funcionamiento cognitivo de un niño, pero nunca deberían interpretarse aisladamente. Las respuestas a los ítems y las puntuaciones obtenidas en las pruebas ofrecen información cualitativa y cuantitativa que se interpreta de forma más adecuada junto con el historial detallado, acompañado de una observación clínica y minuciosa del niño. La WISC-V ofrece múltiples métodos para evaluar las aptitudes cognitivas en el marco de muy diversas evaluaciones psicológicas. Un patrón de puntuaciones obtenidas en un test puede servir para evaluar los puntos fuertes y débiles cognitivos, pero no para diagnosticar trastornos específicos, puesto que en los distintos trastornos pueden observarse patrones similares. El profesional, además, suele disponer de otras fuentes de información: historial médico, educativo y psicosocial; observaciones conductuales directas; puntuaciones obtenidas en tests previos; aspectos cualitativos del rendimiento del niño en un test; y resultados obtenidos en otros instrumentos relevantes. Asimismo, el profesional debería evaluar los resultados en el contexto del motivo o del propósito de la evaluación.

Muchos de los métodos, estrategias y procedimientos de investigación desarrollados y perfeccionados por profesionales e investigadores para las anteriores escalas de inteligencia de Wechsler continúan siendo válidos y útiles (Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Prifitera et al., 2008; Sattler, 2008a, 2008b). No obstante, el profesional ha de tener en cuenta los cambios en el contenido de las distintas versiones cuando utilice estas directrices de interpretación. La interpretación de las puntuaciones de los índices de la WISC-V como medidas fiables y válidas de los principales constructos cognitivos medidos cuenta con un fuerte respaldo psicométrico y clínico. También es importante tener en cuenta que cada prueba requiere la intervención de diversos procesos cognitivos para completar la tarea, y que los procesos requeridos pueden cambiar a lo largo de la prueba.

Este capítulo se centra en los aspectos fundamentales de la interpretación de los resultados de la WISC-V, especialmente en la presentación de las puntuaciones y el uso de las tablas a efectos interpretativos. Queda fuera del alcance de este capítulo proponer un método detallado, sistemático y completo de análisis e interpretación de los resultados; no obstante, en las siguientes secciones se exponen los principios básicos de interpretación, así como información sobre la presentación de los resultados y la elaboración del perfil. Al final de este capítulo se incluye información sobre la interpretación de las discrepancias entre aptitud y rendimiento; es decir, para analizar las discrepancias entre aptitud y rendimiento, así como entre los patrones de puntos fuertes y débiles.

Descripción de los resultados

Las puntuaciones típicas sitúan el rendimiento reflejado por cualquier puntuación de las pruebas o cualquier puntuación compuesta en una métrica común, de modo que indican el rendimiento del niño en comparación con el de otros niños de su misma edad. La conversión de las puntuaciones directas en puntuaciones típicas permite al profesional comparar las puntuaciones de las pruebas entre sí y con otras medidas relacionadas. Aunque las puntuaciones típicas son las más útiles para realizar análisis, a menudo se complementan con otros tipos de puntuaciones, como los percentiles, los intervalos de confianza, las clasificaciones descriptivas y las edades equivalentes, cuya comprensión es más fácil.

Puntuaciones típicas

La WISC-V ofrece dos tipos de puntuaciones ajustadas a la edad: *puntuaciones escalares* (las que tienen una media 10 y una desviación típica 3) y *puntuaciones compuestas* (las que tienen una media 100 y una desviación típica 15). Las puntuaciones de las pruebas y muchas de las puntuaciones de procesamiento son puntuaciones escalares. Las puntuaciones escalares representan dónde se sitúa una puntuación directa, obtenida por un niño, en la distribución de puntuaciones directa en relación con los niños de su misma edad. Una puntuación escalar de 10 refleja el rendimiento medio de un determinado grupo de edad. Las puntuaciones escalares 7 y 13 se alejan una desviación típica por debajo y por encima de la media, respectivamente. Las puntuaciones escalares 4 y 16 se alejan dos desviaciones típicas por debajo y por encima de la media, respectivamente. La tabla 7.1 presenta la relación entre las puntuaciones escalares, las desviaciones típicas respecto a la media y los percentiles equivalentes.

Tabla 7.1. Relación entre las puntuaciones escalares, las desviaciones típicas respecto a la media y los percentiles equivalentes

Puntuación escalar	*Dt* a la media	Percentil[a]
19	+3	99.9
18	$+2\frac{2}{3}$	99.6
17	$+2\frac{1}{3}$	99
16	+2	98
15	$+1\frac{2}{3}$	95
14	$+1\frac{1}{3}$	91
13	+1	84
12	$+\frac{2}{3}$	75
11	$+\frac{1}{3}$	63
10	0 (Media)	50
9	$-\frac{1}{3}$	37
8	$-\frac{2}{3}$	25
7	−1	16
6	$-1\frac{1}{3}$	9
5	$-1\frac{2}{3}$	5
4	−2	2
3	$-2\frac{1}{3}$	1
2	$-2\frac{2}{3}$	0.4
1	−3	0.1

[a] Los percentiles son valores teóricos para una distribución normal.

Las puntuaciones compuestas son puntuaciones típicas basadas en la suma de las puntuaciones escalares de varias pruebas. Una puntuación 100 en cualquiera de las puntuaciones compuestas define el rendimiento medio de los niños de la misma edad, y las puntuaciones 85 y 115 se corresponden con un alejamiento de una desviación típica por debajo y por encima de la media, respectivamente. Las puntuaciones 70 y 130 se alejan dos desviaciones típicas por debajo y por encima de la media, respectivamente. Para las puntuaciones tipificadas en esta métrica, aproximadamente el 68 % de todos los niños obtienen puntuaciones compuestas entre 85 y 115, alrededor del 96 % de los niños logran puntuaciones compuestas entre 70 y 130, y prácticamente todos los niños (alrededor del 99.8 %) obtienen puntuaciones entre 55 y 145 (que se alejan tres desviaciones típicas por debajo y por encima de la media, respectivamente). La tabla 7.2 presenta la relación entre las puntuaciones compuestas, las desviaciones típicas respecto a la media y los percentiles equivalentes.

Tabla 7.2. Relación entre las puntuaciones compuestas, las desviaciones típicas respecto a la media y los percentiles equivalentes

Puntuación compuesta	*Dt* a la media	Percentil[a]
160	+4	>99.9
155	$+3\frac{2}{3}$	>99.9
150	$+3\frac{1}{3}$	>99.9
145	+3	99.9
140	$+2\frac{2}{3}$	99.6
135	$+2\frac{1}{3}$	99
130	+2	98
125	$+1\frac{2}{3}$	95
120	$+1\frac{1}{3}$	91
115	+1	84
110	$+\frac{2}{3}$	75
105	$+\frac{1}{3}$	63
100	0 (Media)	50
95	$-\frac{1}{3}$	37
90	$-\frac{2}{3}$	25
85	−1	16
80	$-1\frac{1}{3}$	9
75	$-1\frac{2}{3}$	5
70	−2	2
65	$-2\frac{1}{3}$	1
60	$-2\frac{2}{3}$	0.4
55	−3	0.1
50	$-3\frac{1}{3}$	<0.1
45	$-3\frac{2}{3}$	<0.1
40	−4	<0.1

[a] Los percentiles son valores teóricos para una distribución normal.

Percentiles

En la WISC-V se ofrecen percentiles para las puntuaciones escalares y las puntuaciones compuestas. Los percentiles indican la posición del niño en relación con otros niños de la misma edad y reflejan el porcentaje de niños de la muestra de tipificación que han obtenido una puntuación inferior a la que se analiza. Los percentiles oscilan normalmente entre 1 y 99, con una mediana de 50. Por ejemplo, un niño con un percentil 15 ha obtenido una puntuación igual o superior al 15 % de los niños de la misma edad (o dicho de otro modo, igual o inferior al 85 % del resto de los niños de la misma edad).

Aunque los percentiles son fáciles de entender y útiles para explicar el rendimiento de un niño en relación con el rendimiento de otros, tienen algunas limitaciones. La escala de percentiles no está compuesta por intervalos iguales. En una distribución normal, como la del CI total, los percentiles tienden a agruparse en torno a la mediana (percentil 50). Por consiguiente, para los niños que obtienen puntuaciones alrededor de la media, una pequeña diferencia en su puntuación directa, apenas 1 o 2 puntos, puede dar a lugar a una gran diferencia en los percentiles; mientras que, para los niños con puntuaciones más extremas, el cambio de 1 o 2 puntos en las puntuaciones directas no dará lugar a una diferencia considerable en los percentiles.

Errores típicos de medida e intervalos de confianza

Las puntuaciones obtenidas en un test de aptitud cognitiva como la WISC-V proporcionan una estimación de la puntuación verdadera del niño. Las puntuaciones verdaderas reflejan la aptitud real de un niño junto con cierto grado de error de medida. Para precisar al máximo la puntuación verdadera de un niño, es más adecuado establecer un intervalo de confianza (un rango de puntuaciones) en el que es probable que se sitúe su puntuación verdadera.

Los intervalos de confianza proporcionan otras maneras de expresar la precisión de la puntuación y sirven como recordatorio de que el error de medida es inherente a todas las puntuaciones de los tests. Las puntuaciones más fiables tienen un intervalo más estrecho, mientras que las puntuaciones menos fiables y precisas tienen un intervalo más amplio. Se anima a los profesionales a registrar los intervalos de confianza de las puntuaciones compuestas y a usarlos para garantizar una mayor precisión en la interpretación de las puntuaciones. En el capítulo 5 de este manual pueden consultarse las fórmulas que permiten calcular los distintos intervalos de confianza de las puntuaciones escalares y de las puntuaciones compuestas.

Las tablas A.2-A.7 y C.1-C.5 del *Manual de aplicación y corrección* presentan los intervalos de confianza de las puntuaciones compuestas utilizando el error típico de estimación (*ETE*) y la puntuación verdadera estimada (*PVE*). Los intervalos de confianza calculados a partir del error típico de medida (*ETM*) se agrupan alrededor de las puntuaciones observadas, mientras que los intervalos de confianza calculados con el error típico de estimación (*ETE*) se agrupan alrededor de las puntuaciones verdaderas estimadas corregidas mediante regresión a la media.

Clasificación cualitativa

A menudo las puntuaciones se describen en términos cualitativos según el nivel de rendimiento. Las clasificaciones resultan útiles para comunicar de forma clara los resultados a personas que no están familiarizadas con los resultados cuantitativos de un test. **Las clasificaciones cualitativas son únicamente sugerencias y no se basan en evidencias; por ello pueden formularse utilizando otros términos.**

Existe un considerable interés por revisar los términos usados tradicionalmente para realizar descripciones cualitativas, y las razones son numerosas. Por una parte, a menudo las personas que reciben los resultados de un test no los entienden muy bien porque los términos que se utilizan en el sistema tradicional no mantienen una relación de correspondencia por debajo y por encima de la media (Raiford y Coalson, 2014). Por otra parte, al comunicar los resultados del test, los términos no se usan consistentemente en las distintas medidas (cognitivas, de rendimiento, de memoria, de habla y de lenguaje), que, además, suelen describirse juntas y comparándolas entre

ellas (Lichtenstein, 2013, 2014; Sattler, 2008a). Por ello es necesario llevar a cabo una revisión de los términos usados en las clasificaciones cualitativas.

En la tabla 7.3 se presenta la clasificación cualitativa de las puntuaciones compuestas. Además del nuevo sistema usado en la WISC-V, se incluye el sistema utilizado en la WISC-IV para mostrar la correspondencia y facilitar su uso, así como por si es necesario comparar los resultados de la WISC-V con los de otra versión de la escala. La clasificación cualitativa de la WISC-V reemplaza la anterior clasificación, y deben utilizarse sus términos en lugar de los anteriormente utilizados. Los resultados de los índices del test pueden describirse de una forma similar a como se describe el CI total en este ejemplo:

> En comparación con los niños de su misma edad, la puntuación obtenida por este niño en una prueba estandarizada de la aptitud intelectual caracteriza su funcionamiento intelectual en el rango [*insertar la clasificación cualitativa correspondiente*] según el rango de inteligencia.

Tabla 7.3. Clasificación cualitativa de las puntuaciones compuestas

Puntuación compuesta	Clasificación cualitativa WISC-V	Clasificación cualitativa anterior (WISC-IV)	Porcentajes	
			Curva normal teórica	Muestra[a]
130 y superior	Muy alto	Muy superior	2.5	2.6
120-129	Alto	Superior	7.2	6.1
110-119	Medio-alto	Normal-alto	16.6	16.4
90-109	Medio	Medio	49.5	53.7
80-89	Medio-bajo	Normal-bajo	15.6	14.4
70-79	Bajo	Inferior	6.5	4.3
69 e inferior	Muy bajo	Muy bajo	2.1	2.7

Nota: Debe usarse la clasificación cualitativa de la WISC-V en lugar de la clasificación tradicional anterior.

[a] Los porcentajes mostrados son para el CI total y se basan en la muestra de tipificación española (N = 1008). Los porcentajes obtenidos para los índices son muy similares.

Puntuaciones de edad equivalente

La puntuación de edad equivalente representa la edad promedio, en años y meses, en la que es más frecuente obtener una determinada puntuación directa. Por ejemplo, una puntuación directa de 22 en la prueba *Cubos* se corresponde con una edad equivalente de 8:2 años, puesto que es la puntuación media obtenida por los niños de 8 años 2 meses. Las puntuaciones de edad equivalente, que figuran en la tabla A.9 del *Manual de aplicación y corrección,* son útiles para describir la aptitud cognitiva de un niño en comparación con el rendimiento promedio o típico de niños de distintas edades, pero deben emplearse con cierta cautela. En primer lugar, las puntuaciones de edad equivalente ofrecen poca información sobre la posición de un niño respecto a los niños de la misma edad. Un niño cuya edad equivalente es aparentemente llamativa puede, o no, situarse en el rango promedio respecto a los niños de la misma edad. Por ejemplo, el niño A, de 8:0 años, ha obtenido una puntuación de edad equivalente de 6:10 años en la prueba *Vocabulario*. Aunque puede parecer que el rendimiento del niño A en esta prueba es bajo, su puntuación directa equivale a una puntuación escalar de 8, que se encuentra en el rango promedio del rendimiento en comparación con los niños de la misma edad.

En segundo lugar, dado que las puntuaciones de edad equivalente se calculan a partir de las puntuaciones directas totales, no producen unidades equivalentes en toda la escala. Por ello, los pequeños cambios de la puntuación directa total pueden convertirse en grandes cambios en las puntuaciones de edad equivalente. Se pueden presentar importantes diferencias entre la edad equivalente y la edad cronológica de un niño, pero es inapropiado interpretar esta diferencia como reflejo de unas aptitudes cognitivas inferiores o superiores a la media. Esto sucede porque dentro del rango de puntuaciones medias se solapan diversos grupos de edad adyacentes. Por ejemplo, el niño B y el niño C tienen 8 años 0 meses, y a ambos se les aplicó la prueba *Semejanzas*. El niño B consiguió una puntuación directa de 21 y una edad equivalente de 7 años 2 meses (7:2). El

niño C consiguió una puntuación directa de 24 y una edad equivalente de 8 años 2 meses (8:2). No obstante, la puntuación escalar del niño B, que es de 9, y la puntuación escalar del niño C, que es de 10, se encuentran en el mismo rango promedio respecto a los niños de la misma edad.

En tercer lugar, las puntuaciones de edad equivalente no son comparables entre las pruebas. Los percentiles de un niño correspondientes a dos pruebas con las mismas puntuaciones de edad equivalente pueden ser considerablemente distintos. Por ejemplo, si el niño D, de 8:0, obtiene una puntuación de edad equivalente de 11:6 tanto en *Cubos* (puntuación escalar de 14) como en *Claves* (puntuación escalar de 17), los percentiles de estas pruebas son 91 y 99, respectivamente.

Finalmente, una puntuación de edad equivalente extrema no significa que las aptitudes cognitivas de un niño se asemejen a las del grupo de menor o mayor edad. Además, las puntuaciones equivalentes en los extremos de edad son más difíciles de interpretar porque de las puntuaciones extremas solo puede decirse que están por debajo de 6 años o por encima de 16 años 11 meses.

Dadas estas limitaciones, se recomienda no utilizar las puntuaciones de edad equivalente como datos principales e interpretarlas con cautela. Para comparar el rendimiento de un niño respecto a los niños de la misma edad, deben usarse las puntuaciones ajustadas a la edad y los percentiles. Las decisiones clínicas han de basarse en el análisis de las puntuaciones ajustadas a la edad del niño, así como en informaciones cualitativas y en su historial. El diagnóstico y la toma de decisiones nunca deben basarse únicamente en las puntuaciones de edad equivalente.

Diferencias entre puntuaciones

La diferencia entre dos puntuaciones obtenidas en una única aplicación es una comparación intraindividual, y esta diferencia debe tenerse en cuenta en la interpretación. Dado que la comparación se basa en el propio rendimiento del niño, al evaluar la diferencia entre puntuaciones, debe considerarse el nivel de aptitud general del niño. Por ejemplo, una puntuación de un índice primario que es alta como valor absoluto (muy por encima de 100) puede representar un relativo punto débil para un niño con una aptitud intelectual alta, respecto a su rendimiento general o a su rendimiento en otro dominio cognitivo. Asimismo debe considerarse el nivel de aptitud general del niño al evaluar las diferencias entre puntuaciones de las pruebas. Por ejemplo, un resultado alto (muy por encima de 10) en una puntuación escalar de una prueba puede representar un relativo punto débil para un niño con una aptitud general alta, respecto a su rendimiento general o a su rendimiento en otra prueba. Por el contrario, un resultado bastante bajo en una puntuación de un índice primario o de una prueba puede indicar un relativo punto fuerte para un niño con una aptitud intelectual baja, respecto a su rendimiento general o su rendimiento en otros dominios cognitivos.

A la hora de interpretar los resultados obtenidos por un niño en la WISC-V, es importante analizar las diferencias que se observan entre las distintas puntuaciones. En este sentido hay dos aspectos relevantes: la significación estadística de la diferencia, determinada por su comparación con los valores críticos; y la frecuencia (tasa base) en que esa diferencia se ha observado en la población. Estos aspectos han de responder dos preguntas, respectivamente: esa diferencia ¿es real o se debe al error de medida? y ¿es clínicamente significativa?

Significación estadística de las diferencias entre puntuaciones

Una diferenciación estadísticamente significativa entre dos puntuaciones, basada en su comparación con los valores críticos, indica que la probabilidad de que esta diferencia se obtenga por azar es muy baja (p. ej., $p < .05$ o $.10$), si la diferencia verdadera entre puntuaciones es 0 (Matarazzo y Herman, 1985). El nivel de significación refleja el grado de confianza que el profesional puede tener respecto a que la diferencia observada entre puntuaciones es una diferencia verdadera; y determina qué diferencia debe haber entre dos puntuaciones (valor crítico) para alcanzar un nivel de confianza específico que indique que la diferencia es una diferencia verdadera y no debida al azar. La elección del nivel de significación del valor crítico se basa en diversos factores, como el objetivo de la evaluación (p. ej., describir las aptitudes cognitivas o realizar el diagnóstico diferen-

cial de un determinado trastorno del aprendizaje) o el número de comparaciones de puntuaciones que pueden llevarse a cabo al mismo tiempo. La WISC-V ofrece valores críticos para evaluar la diferencia entre puntuaciones con los niveles de significación .01, .05, .10 y .15, por grupo de edad y para toda la muestra. El hecho de disponer de más niveles de significación permite que el profesional adecue la selección de los valores críticos a la situación clínica y, con ello, pueda lograr un equilibrio entre la potencia estadística y el posible error de tipo I (falsos positivos).

La elección de un nivel de significación menos estricto (p. ej., .15) puede ser adecuada cuando se van a comparar pocas puntuaciones o en situaciones relativamente de bajo riesgo (p. ej., cuando el error de tipo I no es el problema principal), así como cuando el motivo de consulta es conocer mejor los puntos fuertes y débiles de un niño con una aptitud intelectual media. Por el contrario, cuando se van a comparar varias puntuaciones al mismo tiempo (p. ej., comparaciones entre todas las puntuaciones de los índices primarios) o en situaciones de alto riesgo (p. ej., cuando el error de tipo I es el problema principal), debe elegirse un nivel de significación más estricto (p. ej., .01 o .05) para controlar el posible incremento del error de tipo I. Dado que los niveles de significación más estrictos dan lugar a menos diferencias significativas y pueden suponer la pérdida de la potencia estadística de la hipótesis que se comprueba, es importante considerar tanto el motivo de evaluación como el número de comparaciones simultáneas de puntuaciones que se van a realizar, antes de elegir el nivel de significación de los valores críticos. Si el error de tipo I es un aspecto crítico o si se prevé realizar muchas comparaciones de puntuaciones, el profesional puede optar por dar más relevancia a la evaluación de la tasa base de la diferencia entre puntuaciones de la muestra de tipificación que a la significación de la diferencia entre puntuaciones.

Los valores críticos de la WISC-V se calcularon tomando en cuenta la varianza compartida entre las puntuaciones comparadas. Se utilizan los valores críticos corregidos cuando las puntuaciones que se comparan comparten el contenido de las pruebas, como cuando se compara una puntuación de un índice primario con la media de las puntuaciones de los índices primarios (MIP) o cuando se compara la puntuación de una prueba principal con la media de las puntuaciones escalares de las pruebas principales (MPE-P). En cambio, se utilizan los valores críticos no corregidos cuando se comparan puntuaciones que no comparten contenido del test, como las comparaciones entre índices primarios o puntuaciones de pruebas.

Puntos fuertes y débiles

La comparación de un índice primario o una puntuación de una prueba con una estimación de la aptitud general, como la MIP o la MPE-P, ofrece información relevante sobre los puntos fuertes y débiles intraindividuales en relación con una estimación global del rendimiento en varios dominios cognitivos (Grégoire et al., 2011; Longman, 2004; Naglieri y Paolitto, 2005; Wechsler, 2003, 2008, 2012). Una puntuación de un índice primario que sea significativamente superior a la MIP obtenida por el mismo sujeto puede reflejar un punto fuerte relativo, mientras que si es significativamente inferior a la MIP puede indicar un punto débil relativo. Del mismo modo, una puntuación de una prueba principal que se desvía significativamente de la MPE-P puede indicar un punto fuerte o débil en relación con el rendimiento general del niño.

En la tabla B.1 del *Manual de aplicación y corrección* figuran los valores críticos para comparar uno de los índices primarios y la MIP o el CI total. La tabla B.3 del *Manual de aplicación y corrección* presenta los valores críticos para comparar cada una de las pruebas principales con la MPE-P o con la media de las puntuaciones escalares de las pruebas que conforman el CI total (MPE-CI).

Dado que los valores críticos para realizar comparaciones con la MPE-CI, que figuran en la tabla B.3, se basan en las pruebas para obtener el CI total, si se utiliza la MPE-CI como puntuación de comparación debe comprobarse si se ha realizado alguna sustitución. En tal caso, debe considerarse el valor crítico y la tasa base de la prueba que ha sido sustituida en vez del valor crítico y la tasa base de la prueba sustituta. Por ejemplo, si *Puzles visuales* sustituyó a *Cubos* para obtener el CI total (y, por tanto, contribuyó a la MPE-CI), debe considerarse el valor crítico y la tasa base de *Cubos* (la prueba principal que fue reemplazada) en vez del valor crítico y la tasa base de *Puzles visuales*. Se sigue este proceso aparentemente contradictorio porque la puntuación de *Puzles visuales* se consideró la mejor estimación de la puntuación de *Cubos*.

Comparación entre índices, pruebas o puntuaciones compuestas

Comparar las diferencias entre pares de puntuaciones permite que el profesional pueda evaluar la variabilidad del rendimiento intraindividual en términos de la diferencia entre las puntuaciones de dos índices, las de dos pruebas, la de una prueba y una puntuación de procesamiento, o entre dos puntuaciones de procesamiento. Los valores críticos para comparar las puntuaciones de índices o de pruebas o las puntuaciones de procesamiento figuran en las tablas B.5 y B.7 y C.6 y C.10 del *Manual de aplicación y corrección*.

Tasas base de las diferencias entre puntuaciones

La prevalencia o frecuencia de una diferencia observada en la población general se denomina *tasa base*. A menudo, la diferencia entre dos puntuaciones es estadísticamente significativa aun siendo frecuente. Por consiguiente, la significación estadística de las diferencias entre puntuaciones y la frecuencia de las diferencias tienen implicaciones distintas para la interpretación de los resultados de un test. Para más información sobre la significación, estadística y clínica, de las diferencias entre puntuaciones, véanse los trabajos de Flanagan y Kaufman (2009), Groth-Marnat (2009) y Sattler (2008a).

Dado que la frecuencia de las diferencias entre puntuaciones varía según el nivel de aptitud y no de la edad (Grégoire et al., 2011), en las tablas B.2, B.6, C.7 y C.8 del *Manual de aplicación y corrección* las tasas base para las comparaciones entre índices se presentan para la muestra total y por nivel de aptitud (según rangos de puntuación determinados del CI total o del ICG). Puesto que la frecuencia de las diferencias también puede variar según la dirección de la diferencia (Sattler, 2008a), en las tablas B.2, B.4, B.6, B.8, C.7 y C.8 del *Manual de aplicación y corrección*, las tasas bases para las comparaciones entre puntuaciones de índices o o pruebas o puntuaciones de procesamiento se presentan según la dirección de la diferencia (positiva o negativa).

Análisis de dispersión entre pruebas

La dispersión de las puntuaciones escalares refleja la variabilidad del rendimiento del niño en las diferentes pruebas (Matarazzo, Daniel, Prifitera y Herman, 1988; Sattler, 2008a). Aunque se han propuesto varias medidas de la dispersión (véase Schinka, Vanderploeg y Curtiss, 1997), el índice de dispersión utilizado en el *Manual de aplicación y corrección* se calcula a partir de la diferencia entre la puntuación escalar más alta y la puntuación más baja obtenida por el sujeto en las pruebas.

Para analizar la dispersión, debería considerarse la frecuencia de dicha dispersión en la muestra de tipificación. Los porcentajes acumulados de la dispersión entre pruebas principales y entre pruebas que permiten obtener el CI total figuran en la tabla B.9 del *Manual de aplicación y corrección*. Por ejemplo, entre las diez pruebas que conforman los índices primarios, el 6.8 % de los niños de la muestra de tipificación obtuvieron una dispersión de 11 puntos o más; y, entre las siete pruebas que permiten obtener el CI total, el 3.2 % de los mismos niños obtuvieron una dispersión de 11 puntos o más. Para aquellos casos poco habituales en que el niño obtiene la puntuación más alta y la más baja en pruebas que contribuyen al mismo índice (p. ej., puntuación más alta en *Vocabulario* y más baja en *Semejanzas*), puede ser preferible considerar las tasas bases para la correspondiente comparación entre pruebas, puesto que dichas tasas bases se ofrecen según la dirección de la diferencia y proporcionan datos más descriptivos a efectos de interpretación.

Etapas para el análisis de un perfil

El rendimiento de un niño puede interpretarse en función de su perfil de puntuaciones de las pruebas y de puntuaciones compuestas. El análisis de un perfil puede realizarse desde una perspectiva intraindividual y desde una perspectiva interindividual comparando los patrones de puntuaciones obtenidas por el niño, tanto puntuaciones de las pruebas como puntuaciones compuestas. Estas comparaciones aptitudinales pueden ayudar al profesional a identificar patrones potencialmente significativos de puntos fuertes y débiles, lo que es importante para describir trastornos funcionales y para diseñar y preparar planes de tratamiento (Flanagan y Kaufman, 2009; Sattler, 2008a, 2008b).

Un buen análisis del perfil del niño debe empezar con la generación de hipótesis basadas en el patrón de las puntuaciones obtenidas. Posteriormente, estas hipótesis se corroborarán o rechazarán en función de los resultados de otras evaluaciones, de su historial o de observaciones directas del comportamiento (Flanagan y Kaufman, 2009; Prifitera et al., 2008). Para comprender y determinar los puntos fuertes y débiles cognitivos de un niño, los profesionales siempre deberían obtener cuantos más datos confirmatorios sobre el niño sea posible, incluidos los relativos a su nacimiento y desarrollo; antecedentes familiares y culturales; historial médico, educativo y social; y resultados obtenidos en evaluaciones anteriores. Toda esta información debe tenerse en cuenta al interpretar los resultados de un niño, y debe constar en los informes sobre el rendimiento del niño en la escala.

En la siguiente sección se describen los pasos básicos para realizar el análisis primario y secundario del perfil; no obstante, para llevar a cabo una interpretación más racional, es necesario que un profesional experto integre los datos obtenidos en el cuadro clínico del niño. Si se desea información más detallada sobre la interpretación clínica de las anteriores ediciones de la WISC, véanse Flanagan y Kaufman (2009), Prifitera et al. (2008) y Sattler (2008a, 2008b).

Realización del análisis primario de un perfil

Para empezar, deben anotarse las puntuaciones de las pruebas, las puntuaciones compuestas y las puntuaciones de procesamiento en el cuadernillo de anotación, junto con los percentiles e intervalos de confianza correspondientes.

Paso 1. Descripción del CI total

El CI total es la puntuación más fiable y, tradicionalmente, ha sido la primera que se considera al interpretar un perfil. Se obtiene de la suma de siete puntuaciones escalares de las pruebas. Por lo general, se considera que el CI total es la puntuación más representativa del funcionamiento de la aptitud intelectual general (*g*).

Para interpretar mejor el CI total, deben tenerse en cuenta los índices primarios y analizar las diferencias entre los índices primarios y otras puntuaciones. Los constructos medidos por los índices primarios suponen una parte importante de la inteligencia general. Se recomienda analizar los índices primarios en la primera fase de la interpretación clínica, en especial en los casos en que se da una variabilidad considerable entre los índices primarios y/o las puntuaciones de las pruebas principales (Weiss, Saklofske, Prifitera y Holdnack, 2006).

Cada índice primario puede compararse con una estimación del rendimiento general (la MIP o el CI total), y los índices primarios pueden compararse entre ellos (p. ej., ICV e IRF). La variabilidad entre las puntuaciones escalares de las pruebas pueden aportar más información relevante para interpretar el CI total. Así pues, para disponer de más datos relevantes para llevar a cabo la interpretación de los resultados, pueden compararse las puntuaciones escalares de las pruebas con una estimación del rendimiento general (la MPE-P o la MPE-CI), las puntuaciones escalares de las pruebas entre ellas (p. ej., *Semejanzas* y *Vocabulario*), y puede realizarse el análisis de dispersión de las pruebas. Véanse Flanagan y Kaufman (2009) y Sattler (2008a) para información más detallada sobre la dispersión de las pruebas en la interpretación de las puntuaciones compuestas.

Paso 2. Descripción del ICV

El ICV mide la aptitud del niño para acceder y utilizar el conocimiento léxico adquirido. La utilización de este conocimiento implica la formación de conceptos verbales, el razonamiento y la expresión. Todos los ítems de las pruebas que contribuyen a esta puntuación, incluidos los ítems de *Vocabulario*, requieren una respuesta verbal por parte del niño. Las puntuaciones altas en el ICV indican que el sistema de razonamiento verbal está bien desarrollado, con una adquisición de conocimiento léxico sólida, una eficaz recuperación de la información, una buena capacidad para razonar y resolver problemas verbales, y una efectiva comunicación del conocimiento. Las puntuaciones bajas en el ICV pueden deberse a diversas razones, como que el conocimiento léxico no esté suficientemente desarrollado, que existan dificultades para recuperar la información adquirida o para expresarse verbalmente, o que haya dificultades generales con el razonamiento y la solución de problemas.

Las pruebas que conforman el ICV de la WISC-V son distintas de las que lo conforman en la WISC-IV: *Semejanzas* y *Vocabulario* son pruebas comunes a las dos versiones, pero el ICV de la WISC-IV incluye *Comprensión*, que no forma parte del de la WISC-V. Respecto al ICV de la WISC-IV, el de la WISC-V incide más en el razonamiento mediante el conocimiento léxico y menos en el conocimiento adquirido, el conocimiento práctico y el juicio.

Paso 3. Descripción del IVE

El IVE mide la aptitud del niño para evaluar los detalles visuales y entender las relaciones visoespaciales a fin de construir diseños geométricos a partir de un modelo. La capacidad de construir diseños requiere razonamiento visoespacial, integración y síntesis de las relaciones parte-todo, atención a los detalles visuales, formación de conceptos no verbales e integración visomotora. Las puntuaciones altas en el IVE indican que la aptitud para aplicar el razonamiento espacial y analizar detalles visuales está bien desarrollada. Las puntuaciones bajas en el IVE pueden deberse a déficits del razonamiento espacial, dificultades con la discriminación visual, una atención visual insuficiente, déficits en la integración visomotora o una aptitud de razonamiento general baja.

Las pruebas que conforman el IVE de la WISC-V son distintas de las que conforman el IRP de la WISC-IV. *Cubos* es la única prueba común. El IVE de la WISC-V no incluye *Matrices* ni *Conceptos*, que forman parte del IRP de la WISC-IV; y sí incluye *Puzles visuales*, una prueba nueva de esta última versión de la escala. Respecto al IRP de la WISC-IV, el IVE de la WISC-V incide más en el razonamiento visoperceptivo y visoespacial, y menos en el conceptual.

Paso 4. Descripción del IRF

El IRF mide la aptitud del niño para detectar las relaciones conceptuales subyacentes entre objetos visuales y usar el razonamiento a fin de identificar y aplicar reglas. La identificación y aplicación de relaciones conceptuales en el IRF requiere razonamiento inductivo y cuantitativo, inteligencia visual amplia, procesamiento simultáneo y pensamiento abstracto. Las puntuaciones altas en el IRF indican que la aptitud para abstraer información conceptual a partir de los detalles visuales y aplicar este conocimiento eficazmente está bien desarrollada. Las puntuaciones bajas en el IRF pueden deberse a diversas razones, como dificultades para identificar la información visual importante, para vincular la información visual a conceptos abstractos o para entender y aplicar conceptos conceptuales o cuantitativos; o una aptitud de razonamiento general baja.

Las pruebas que conforman el IRF de la WISC-V son distintas de las que conforman el IRP de la WISC-IV. *Matrices* es la única prueba común. El IRF de la WISC-V no incluye *Cubos* ni *Conceptos*, que forma parte del IRP de la WISC-IV; y sí incluye *Balanzas*, una prueba nueva de esta última versión de la escala. Respecto al IRP de la WISC-IV, el IRF de la WISC-V incide más en el razonamiento conceptual abstracto y menos en las aptitudes constructivas que requieren integración visoperceptiva y razonamiento visoespacial.

Paso 5. Descripción del IMT

El IMT mide la aptitud del niños para registrar, mantener y manipular información visual y auditiva de forma consciente. El registro de la información requiere atención, discriminación visual y auditiva y concentración. El mantenimiento es el proceso por el cual la información se conserva activa de forma consciente mediante el bucle fonológico o la agenda visual (véase Baddeley, 2012). La manipulación es la resecuenciación mental de la información a partir de la aplicación de un regla específica. Las puntuaciones altas en el IMT indican que la aptitud para identificar información visual y auditiva, mantenerla en el almacenamiento temporal y resecuenciarla para usarla en la solución de problemas está bien desarrollada. Las puntuaciones bajas en el IMT pueden deberse a varias razones, como distractibilidad, problemas de discriminación visual o auditiva, dificultades para mantener la información activa de forma consciente, capacidad de almacenamiento insuficiente, dificultades para manipular la información en la memoria de trabajo, o funcionamiento cognitivo general bajo.

Las pruebas que conforman el IMT de la WISC-V son distintas de las que lo conforman en la WISC-IV. *Dígitos* es la única prueba común, pero la versión de *Dígitos* de la WISC-V es el resultado de una considerable revisión llevada a cabo con el fin de aumentar la carga de la memoria de trabajo, lo que supuso la incorporación de la tarea en orden creciente. El IMT de la WISC-V no incluye *Letras y números*, que forma parte del IMT de la WISC-IV, porque ha sido sustituida por *Span de dibujos*, una prueba nueva de memoria de trabajo visual. Respecto al IMT de la WISC-IV, el de la WISC-V incide más en la memoria de trabajo visual y menos en la memoria de trabajo auditiva.

Paso 6. Descripción del IVP

El IVP mide la velocidad y la precisión de la identificación visual, la toma de decisiones y la implementación de estas. El rendimiento en el IVP se relaciona con el rastreo visual, la discriminación visual, la memoria visual a corto plazo, la coordinación visomotora y la concentración (Flanagan y Kaufman, 2009; Groth-Marnat, 2009; Sattler, 2008a, 2008b). Las puntuaciones altas en el IVP indican que la aptitud para identificar información visual rápidamente, tomar decisiones rápidas y certeras e implementarlas rápidamente está bien desarrollada. Las puntuaciones bajas en el IVP pueden deberse a varias razones, como problemas de discriminación visual, distractibilidad, toma de decisiones enlentecida, dificultades motoras o, por lo general, velocidad cognitiva lenta. Diversos estudios señalan una correlación significativa entre la velocidad de procesamiento y la aptitud intelectual (Dodonova y Dodonov, 2012), y subrayan la sensibilidad de las pruebas de velocidad de procesamiento a ciertos estados clínicos, como el TDAH, los trastornos específicos del aprendizaje y el trastorno del espectro autista (Calhoun y Mayes, 2005; Hale et al., 2012; Mayes et al., 2012). Las pruebas que conforman el IVP no son medidas solo del tiempo de reacción o de la discriminación visual, sino que en ellas interviene un componente cognitivo de toma de decisiones y de aprendizaje.

El IVP de la WISC-V está formado por las mismas pruebas que el de la WISC-IV: *Claves* y *Búsqueda de símbolos*. No obstante, en la WISC-V, la dificultad de los ítems en estas pruebas está equilibrada más uniformemente entre las pruebas, en vez de aumentar gradualmente a lo largo de los ítems. *Claves* se rediseñó para que no fuera necesario levantar el lápiz al dibujar cada símbolo. Por lo general, el IVP de la WISC-V y el de la WISC-IV son similares en su composición e interpretación.

Paso 7. Evaluación de los puntos fuertes y débiles en los índices

En la mayoría de los niños suele existir cierto grado de variabilidad cognitiva. Para obtener información sobre los puntos fuertes y débiles cognitivos de un niño respecto a su nivel de aptitud cognitiva general, pueden compararse los índices primarios con un indicador del rendimiento general (la MIP o el CI total). Al analizar los puntos fuertes y débiles específicos de los índices, debe seguirse un criterio claro establecido a priori. Una diferencia significativa entre una puntuación de un índice y un indicador del rendimiento general representa un punto fuerte o débil en el dominio cognitivo correspondiente. Sin embargo, la importancia clínica de un punto fuerte o débil depende del niño y del contexto de la evaluación.

Cuando se obtiene una diferencia significativa entre puntuaciones, su interpretación debe corroborarse mediante otras evidencias, y el uso de otros instrumentos de evaluación puede servir para refutar o confirmar una hipótesis. Si los datos obtenidos no corroboran la hipótesis original, pueden surgir otras hipótesis, y estas pueden evaluarse mediante otros instrumentos u otras técnicas que permiten obtener información. Por ejemplo, si un niño muestra un rendimiento significativamente más bajo en el IVE respecto a la MIP, la baja puntuación puede atribuirse a que su aptitud visoespacial es baja o sus destrezas motoras son insuficientes. En este caso, las evaluaciones adicionales podrían centrarse en tareas de aptitud visoespacial sin requisitos motores, o bien en tareas motoras en las que no interviniera la aptitud visoespacial. Para confirmar o rechazar las hipótesis, deben tenerse en cuenta la información recopilada mediante entrevistas, las observaciones directas, otros aspectos, así como los resultados de otras evaluaciones.

La interpretación de los puntos fuertes o débiles reflejados en los índices dependen, en parte, del indicador usado en la comparación (la MIP o el CI total). La proporción de pruebas de cada escala primaria es distinta en la MIP y el CI total: algunas pruebas que contribuyen a la MIP no contribuyen al CI total. Por lo tanto, el contenido de estos indicadores es ligeramente distinto, y las pruebas que los componen deben tenerse en cuenta al realizar la interpretación.

En la página de análisis primario del cuadernillo de anotación se dispone de una sección para identificar los puntos fuertes y débiles del niño en los índices. El profesional ha de elegir la MIP o el CI total como puntuación de comparación. Si es posible, es preferible utilizar la MIP porque se basa en más pruebas y es una muestra más amplia del rendimiento que el CI total. Asimismo, el profesional debe elegir un nivel de significación de los valores críticos, lo que implica un juicio clínico. La decisión tomada debe basarse en el tipo y número de comparaciones que vayan a realizarse, el objetivo de la evaluación y el motivo de consulta. En el apartado «Significación estadística de las diferencias entre puntuaciones» de este capítulo y en el apartado «Opciones de comparación» del capítulo 2 del *Manual de aplicación y corrección* se exponen más consideraciones psicométricas y clínicas sobre la elección del nivel de significación más adecuado.

La tabla B.1 del *Manual de aplicación y corrección* presenta los valores críticos corregidos para realizar comparaciones entre cada uno de los índices primarios y la MIP o el CI total con los niveles de significación .01, .05, .10 y .15. El profesional puede usar esta tabla para decidir si las diferencias entre puntuaciones son significativas estadísticamente. Cabe señalar que los valores se basan en la aplicación de las pruebas principales que conforman cada puntuación. Si se ha realizado alguna sustitución de las permitidas, los valores de esta tabla han de usarse con cautela al valorar los puntos fuertes y débiles de un niño; dado que es posible que se haya introducido un error de medida adicional, este debe registrarse en el cuadernillo de anotación.

En la tabla B.2 del *Manual de aplicación y corrección* figuran las tasas base de las diferencias entre cada uno de los índices primarios y la MIP o el CI total. Estas tasas base se presentan en las dos direcciones de la diferencia, para la muestra total y según el nivel de aptitud reflejado en el CI total. Con esta tabla, el profesional puede determinar la frecuencia de las diferencias respecto a la muestra de tipificación. Los profesionales a menudo se preguntan qué nivel de ocurrencia (tasa base) debe considerarse poco frecuente: 15 %, 10 %, 5 % o 2 %. El uso de un criterio más o menos restrictivo afecta al número de observaciones poco frecuentes que pueden identificarse, y la decisión de usar un valor de corte u otro implica considerar el impacto de identificar, en exceso o por defecto, puntos fuertes o débiles cognitivos. Sattler (2008a) sugiere que si las diferencias entre puntuaciones se dan en menos del 15 % de la muestra de tipificación, deben considerarse inusuales.

Paso 8. Evaluación de las comparaciones entre los índices

Wechsler (1944) destacó el CI total y lo describió como el promedio del rendimiento mostrado por un sujeto en todas las pruebas. Sin embargo, Wechsler observó que a veces era conveniente analizar por separado las puntuaciones de pruebas verbales y manipulativas, en especial en personas «con trastornos especiales que requieren una atención especial» (Wechsler, 1994, p. 138). Aunque es más probable que se hallen diferencias estadísticamente significativas entre las puntuaciones de los índices primarios en poblaciones clínicas, con frecuencia también se encuentran

diferencias similares en la muestra de tipificación (Flanagan y Kaufman, 2009; Sattler, 2008a). Las comparaciones son muy útiles para evaluar hipótesis interpretativas específicas.

Los profesionales deben tener en cuenta que las puntuaciones de los índices primarios son estimaciones del funcionamiento general de un dominio cognitivo o área de contenido particular. Por ello, estas puntuaciones siempre deben analizarse en el contexto de las dos pruebas que las conforman. Una discrepancia extrema entre las dos puntuaciones escalares de las pruebas que contribuyen a cada índice primario indica que existe una inconsistencia en la expresión de las aptitudes medidas por el índice. Al interpretar estas puntuaciones, el profesional debe examinar minuciosamente los resultados obtenidos por el sujeto en las dos pruebas que conforman el índice. Cuando se observan diferencias significativas y poco frecuentes, debe tenerse en cuenta esta variabilidad al interpretar y comparar las puntuaciones de los índices primarios (Flanagan y Kaufman, 2009; Sattler, 2008a). La diferencia entre dos puntuaciones de índices primarios puede deberse a una dificultad específica en una de las pruebas que contribuye a uno de los índices, y, por tanto, no reflejaría una diferencia real entre los dominios cognitivos.

En la página de análisis primario del cuadernillo de anotación se incluye una tabla que facilita el análisis de las comparaciones entre índices. La tabla B.5 del *Manual de aplicación y corrección* ofrece los valores críticos de las diferencias estadísticamente significativas entre índices primarios con los niveles de significación .01, .05, .10 y .15, por grupo de edad. Si las puntuaciones de los índices primarios no son significativamente diferentes, puede presuponerse que los niveles de aptitud del niño en los dos dominios son iguales. Si las puntuaciones de los índices primarios son significativamente diferentes, es probable que los niveles reales de aptitud del niño en los dos dominios cognitivos sean distintos, y debe tenerse en cuenta si su frecuencia es baja. En la tabla B.6 del *Manual de aplicación y corrección* figuran las tasas base de las diferencias entre las puntuaciones de los índices primarios obtenidas por la muestra de tipificación. Estas tasas base se presentan en las dos direcciones de la diferencia, para la muestra total y según el nivel de aptitud reflejado en el CI total.

ICV e IVE

Se pueden comparar el ICV y el IVE para obtener información sobre las aptitudes visoperceptivas/visoespaciales respecto a las aptitudes de razonamiento verbal. A diferencia del IRF, el IVE no evalúa específicamente aspectos del razonamiento cognitivo, que tiene un componente semántico/lingüístico. Por ello, comparar el ICV y el IVE representa comparar el procesamiento mental de la información visual con el de la información verbal. Una puntuación superior en el ICV respecto al IVE indica un punto fuerte relativo en el uso de los estímulos verbales para resolver problemas respecto a la resolución de problemas visoespaciales. Una puntuación superior en el IVE respecto al ICV indica un punto fuerte en la capacidad de comprender y utilizar información visoperceptiva/visoespacial en comparación con las destrezas de razonamiento verbal.

ICV e IRF

Se pueden comparar el ICV y el IRF para obtener información sobre las destrezas de razonamiento utilizando estímulos verbales y visuales, así como sobre la aptitud cristalizada respecto a la aptitud fluida. Si el ICV y el IRF no son significativamente diferentes, ello sugiere que la aptitud cristalizada corresponde con la fluida, al igual que el razonamiento conceptual con información verbal corresponde con la visual. Una puntuación superior en el ICV respecto al IRF indica un punto fuerte en las aptitudes cristalizadas respecto a las aptitudes de razonamiento fluido; mientras que una puntuación superior en el IRF respecto al ICV indica un punto fuerte en las aptitudes fluidas respecto a las cristalizadas. Para esclarecer si la diferencia se relaciona principalmente con el tipo de estímulo (visual o verbal), puede analizarse el IVE.

IVE e IRF

Se pueden comparar el IVE y el IRF para obtener información sobre el razonamiento visoperceptivo y visoespacial respecto al razonamiento conceptual abstracto evaluado mediante estímulos visuales. Las diferencias entre estas medidas sugieren que existe variabilidad en la capacidad de los niños para usar información visual. Si el IVE y el IRF no son significativamente diferentes, ello sugiere que el procesamiento visoespacial y el razonamiento visual son acordes. Una puntuación

superior en el IVE respecto al IRF indica que el niño tiene un razonamiento visoperceptivo muy sólido respecto al razonamiento fluido. Una puntuación superior en el IRF respecto al IVE indica que el niño tiene una mayor capacidad para entender la relación entre la información visual y los conceptos abstractos que para usar la información visual y espacial para construir diseños. Si el ICV es un punto fuerte relativo, el niño puede mediar verbalmente las tareas (p. ej., transformar la información visual en información verbal describiendo el problema en voz baja o en silencio, para sí mismo) a fin de resolver con éxito las tareas del IRF.

ICV, IVE e IRF

Estas tres puntuaciones de índices pueden usarse juntas para mejorar la comprensión de los puntos fuertes y débiles en la capacidad de razonamiento de un niño. En casos en que el ICV es mayor que el IVE y el IRF, y el IVE y el IRF son similares, hay evidencias que sugieren que existe un punto débil constante en la capacidad para usar información visual en la resolución de problemas, o un punto fuerte en el razonamiento verbal. Si el ICV es menor que el IVE y el IRF, y el IVE y el IRF son similares, puede existir un punto débil general en el funcionamiento del lenguaje del niño, en la resolución de problemas verbales y en el razonamiento verbal. Cuando el IVE es mayor que el ICV y el IRF, y el ICV y el IRF son similares, el niño muestra un punto fuerte en el procesamiento visoperceptivo/visoespacial respecto al pensamiento conceptual y abstracto. Si el IVE es menor que el ICV y el IRF, y el ICV y el IRF son similares, existe un punto débil relativo en las destrezas de razonamiento visoperceptivo y espacial en comparación con las aptitudes de razonamiento conceptual. Cuando el IRF es mayor que el ICV y el IVE, y el ICV y el IVE son similares, el niño muestra un punto fuerte en la capacidad de vincular la información visual a los constructos semántico y cuantitativo respecto al razonamiento conceptual visoespacial y verbal. Si el IRF es menor que el ICV y el IVE, y el ICV y el IVE son similares, existe un punto débil relativo en la capacidad de vincular la información visual a los constructos semántico y cuantitativo respecto al razonamiento conceptual visoespacial y verbal. Cuando existe una gran variabilidad entre estas tres puntuaciones, se recomienda usar otros instrumentos de evaluación del funcionamiento visoespacial y del lenguaje para clarificar cualquier discrepancia observada.

IMT e IVP

El IMT y el IVP identifican diferentes componentes de la eficacia del procesamiento de la información. Mientras que el ICV, el IRF y el IVE están fuertemente asociados con el razonamiento y la resolución de problemas, el IMT y el IVP favorecen la capacidad de resolución de problemas. La memoria de trabajo y la velocidad de procesamiento están relacionadas porque la memoria de trabajo implica identificación, registro y manipulación de la información en el almacenamiento de la memoria a corto plazo, y la velocidad de procesamiento favorece la identificación y el registro rápidos de la información en la memoria a corto plazo para la toma de decisiones. Si el IMT y el IVP no son significativamente diferentes, ello sugiere que la memoria de trabajo y la velocidad de procesamiento son acordes. Una puntuación superior en el IMT respecto al IVP indica que la capacidad de identificar y registrar información en la memoria a corto plazo es un punto fuerte respecto a la velocidad para tomar decisiones utilizando esta información. Una puntuación superior en el IVP respecto al IMT sugiere que el niño puede ser más competente en la toma de decisiones rápidas utilizando la información registrada en la memoria a corto plazo que en la manipulación de esta información. Sin embargo, el hecho de que los estímulos de las pruebas de Velocidad de procesamiento estén presentes durante la realización de las tareas puede servirle al niño para superar ciertas limitaciones en el funcionamiento de la memoria de trabajo.

IMT respecto a ICV, IVE e IRF

Las pruebas de Memoria de trabajo miden la capacidad de manipular información en el conocimiento consciente para favorecer la resolución de problemas complejos. El ICV, el IVE y el IRF representan aptitudes más complejas cuya eficacia requiere, en teoría, la participación de las destrezas de memoria de trabajo. Dicho de otro modo, una memoria de trabajo baja puede interferir con la capacidad de realizar operaciones mentales más complejas, o inhibirla. Cuando el IMT es mayor que el ICV, el IVE o el IRF, la capacidad de manipular mentalmente información es superior a la capacidad de resolver problemas complejos. En este caso, es posible que los déficits de la memoria de trabajo no estén limitando el rendimiento en estas tareas. Si el IMT es menor

que el ICV, el IVE o el IRF, la memoria de trabajo relativamente más baja del niño puede no interferir con la resolución de problemas complejos.

IVP respecto a ICV, IVE e IRF

Las pruebas de Velocidad de procesamiento miden la capacidad del niño para identificar, registrar, tomar e implementar rápidamente decisiones sobre estímulos visuales. El ICV, el IVE y el IRF representan aptitudes más complejas que pueden verse favorecidas por la capacidad de pensar rápidamente y tomar decisiones certeras. Una velocidad de procesamiento cognitivo baja podría interferir con la capacidad para realizar operaciones mentales más complejas, o inhibirla, en especial en las pruebas de Visoespacial y en la prueba *Balanzas*, cuyas puntuaciones pueden verse afectadas por la velocidad. Cuando el IVP es mayor que el ICV, el IVE y el IRF, la capacidad para evaluar rápidamente la información visual y tomar decisiones es superior a la capacidad para resolver problemas complejos. En este caso, es posible que los déficits de la velocidad de procesamiento no estén limitando el rendimiento en estas tareas. Si el IVP es menor que el ICV, el IVE y el IRF, el niño puede resolver problemas complejos a pesar de tener una aptitud de velocidad de procesamiento relativamente más baja.

Paso 9. Evaluación de los puntos fuertes y débiles en las pruebas

Al igual que se ha visto respecto a los índices primarios, la mayoría de los niños presentan puntos débiles y fuertes en las pruebas relacionados con su rendimiento en los dominios de aptitud cognitiva. No obstante, al analizar los puntos fuertes y débiles específicos de los índices, debe seguirse un criterio claro establecido a priori. El significado clínico de estos resultados debe contextualizarse con la información recopilada y con las puntuaciones de las pruebas, y nunca debe confirmarse o rechazarse una hipótesis en función de una única observación.

Las puntuaciones de las pruebas principales pueden compararse con la MPE-P o la MPE-CI. Al interpretar los puntos fuertes y débiles de un niño en las pruebas y comparar la puntuación de una prueba con uno de estos indicadores generales del rendimiento, los profesionales deben considerar qué pruebas contribuyen al indicador de comparación.

En la página de análisis primario del cuadernillo de anotación se dispone de una sección para identificar los puntos fuertes y débiles del niño en las pruebas. El profesional ha de elegir la MPE-P o la MPE-CI como puntuación de comparación. Si es posible, es preferible utilizar la MPE-P porque se basa en más pruebas y es una muestra más amplia del rendimiento que la MPE-CI. Asimismo, el profesional debe elegir un nivel de significación de los valores críticos para identificar las diferencias significativas.

La tabla B.3 del *Manual de aplicación y corrección* presenta los valores críticos para realizar comparaciones entre cada una de las puntuaciones escalares de las pruebas principales y la MPE-P o la MPE-CI con los niveles de significación .01, .05, .10 y .15, por grupo de edad. Dado que los valores para realizar comparaciones con la MPE-CI, que figuran en la tabla B.3, se basan en las pruebas que permiten obtener el CI total, si se utiliza la MPE-CI como puntuación de comparación debe comprobarse si se ha realizado alguna sustitución. En tal caso, debe considerarse el valor crítico y la tasa base de la prueba que ha sido sustituida en vez del valor crítico y la tasa base de la prueba sustituta. En la tabla B.4 del *Manual de aplicación y corrección* figuran las tasas base de las diferencias entre las pruebas principales y la MPE-P o la MPE-CI. Con esta tabla, el profesional puede determinar la frecuencia de las diferencias respecto a la muestra de tipificación.

Paso 10. Evaluación de las comparaciones entre las pruebas

Dado que cada índice está formado por dos pruebas, la comparación de las puntuaciones escalares de estas dos pruebas es informativa. Por ejemplo, la comparación de las puntuaciones obtenidas en *Dígitos* y *Span de dibujos* puede servir para interpretar el IMT y aportar información relevante sobre el rendimiento en la memoria de trabajo auditiva y visual. En la página de análisis primario del cuadernillo de anotación se incluye una tabla en la que realizar las comparaciones entre cada par de pruebas que contribuyen a cada índice primario.

La tabla B.7 del *Manual de aplicación y corrección* presenta los valores críticos de las diferencias estadísticamente significativas entre cada par de pruebas con los niveles de significación .01, .05, .10 y .15. Por lo general, debe seleccionarse un nivel de significación más estricto si se van a realizar varias comparaciones, puesto que la probabilidad de errores tipo I (falsos positivos) es mayor cuando se comparan muchas puntuaciones al mismo tiempo. En el apartado «Opciones de comparación» del capítulo 2 del *Manual de aplicación y corrección* se ofrece información detallada sobre la elección del nivel de significación más adecuado. En la tabla B.8 del *Manual de aplicación y corrección* figuran las tasas base bidireccionales de las diferencias entre las puntuaciones escalares de las pruebas obtenidas por la muestra de tipificación.

A continuación se describen cinco comparaciones entre pares de pruebas que son interesantes para realizar la interpretación de las pruebas que contribuyen a los índices primarios. En cada caso, las dos pruebas que se comparan difieren en las aptitudes específicas que intervienen en ellas, y la comparación de sus puntuaciones sirve para interpretar la puntuación del índice. La interpretación de las diferencias clínicamente significativas se expone en cada comparación.

Semejanzas y Vocabulario

La comparación *Semejanzas-Vocabulario* proporciona información sobre el rendimiento en tareas de Comprensión verbal que se diferencian por el tipo de aptitudes que requieren. En ambas tareas intervienen las aptitudes verbales y cristalizadas, aunque *Semejanzas* puede estar más relacionada con el razonamiento abstracto y conceptual y la flexibilidad cognitiva; y *Vocabulario*, con el conocimiento léxico. Esta comparación permite interpretar el ICV, que se obtiene de la suma de las puntuaciones escalares de estas dos pruebas. Una puntuación superior en *Semejanzas* respecto a *Vocabulario* indica un punto fuerte relativo en el razonamiento abstracto y la flexibilidad cognitiva en comparación con el conocimiento léxico. Una puntuación superior en *Vocabulario* respecto a *Semejanzas* sugiere que el conocimiento léxico está más desarrollado que el razonamiento abstracto y la flexibilidad cognitiva.

Cubos y Puzles visuales

La comparación *Cubos-Puzles visuales* proporciona información sobre el rendimiento en tareas de construcción visual que requieren aptitudes de razonamiento visoperceptivo y espacial y que se diferencian en algunas capacidades necesarias para completarlas. En ambas tareas interviene el razonamiento visoperceptivo y espacial, pero *Cubos* se basa en los mismos estímulos y respuestas en todos los ítems (cubos, componentes similares de las construcciones) y requiere integración visomotora. Dado que en la prueba se utilizan estímulos físicos, el niño recibe *feedback* visual concreto sobre la precisión de su solución. En cambio, *Puzles visuales* se basa en un estímulo distinto para cada ítem y requiere resolver el problema mentalmente, sin ningún *feedback* visual sobre la precisión de su solución. Esta comparación permite interpretar el IVE, que se obtiene de la suma de las puntuaciones escalares de estas dos pruebas. Una puntuación superior en *Cubos* respecto a *Puzles visuales* puede indicar que el aprendizaje procedimental, la resolución de problemas por ensayo-error, el *feedback* visual concreto y/o la integración visomotora favorecen el rendimiento en tareas en las que interviene el razonamiento visoperceptivo y espacial. Una puntuación superior en *Puzles visuales* respecto a *Cubos* puede indicar que las destrezas visomotoras pueden ser un punto débil respecto al razonamiento visoperceptivo y espacial general.

Matrices y Balanzas

La comparación *Matrices-Balanzas* proporciona información sobre el rendimiento en tareas de Razonamiento fluido que se diferencian por el tipo de aptitudes que requieren. En *Matrices* interviene principalmente el razonamiento inductivo, mientras que en *Balanzas* interviene principalmente el razonamiento cuantitativo. Esta comparación permite interpretar el IRF, que se obtiene de la suma de las puntuaciones escalares de estas dos pruebas. Una puntuación superior en *Matrices* respecto a *Balanzas* indica un punto fuerte relativo en el razonamiento inductivo respecto al razonamiento cuantitativo. Una puntuación superior en *Balanzas* respecto a *Matrices* sugiere que el razonamiento cuantitativo es un punto fuerte respecto al razonamiento inductivo.

Dígitos y Span de dibujos

La comparación *Dígitos-Span de dibujos* proporciona información sobre el rendimiento en tareas de Memoria de trabajo que se diferencian por el tipo de estímulo y de respuesta. *Dígitos* se basa en estímulos verbales con significado, un paradigma de recuerdo libre, y requiere una respuesta verbal; mientras que *Span de dibujos* se basa en estímulos visuales con significado (p. ej., estímulos con una asociación semántica conocida), un paradigma de reconocimiento, y requiere una respuesta no verbal (el niño puede señalar la respuesta). Esta comparación permite interpretar el IMT, que se obtiene de la suma de las puntuaciones escalares de estas dos pruebas. Una puntuación superior en *Dígitos* respecto a *Span de dibujos* indica que el niño puede utilizar mejor la memoria de trabajo cuando la información se presenta auditivamente y no visualmente, o cuando se usa un paradigma de recuerdo libre y no de reconocimiento. Una puntuación superior en *Span de dibujos* respecto a *Dígitos* indica que el niño puede utilizar mejor la memoria de trabajo en la resolución de problemas cuando el estímulo es visual en vez de verbal, o cuando se usa un paradigma de reconocimiento y no de recuerdo libre. Asimismo, una diferencia poco frecuente en cualquiera de las dos direcciones puede indicar que el niño ha sufrido un lapsus de atención o motivación durante la aplicación de la prueba en la que ha obtenido la puntuación más baja, dado que en estas pruebas no se puede repetir o volver a presentar el estímulo.

Claves y Búsqueda de símbolos

La comparación *Claves-Búsqueda de símbolos* proporciona información sobre el rendimiento en tareas de Velocidad de procesamiento que se diferencian por el tipo de aptitudes que requieren. *Claves* también implica el aprendizaje de pares asociados y la velocidad grafomotora; mientras que *Búsqueda de símbolos* requiere que el rastreo y la discriminación visuales sean precisos. Además, los requisitos de motricidad fina en *Búsqueda de símbolos* son menores que en *Claves*. Esta comparación permite interpretar el IVP, que se obtiene de la suma de las puntuaciones escalares de estas dos pruebas. Una puntuación superior en *Claves* respecto a *Búsqueda de símbolos* sugiere puntos fuertes relativos en el aprendizaje de pares asociados y/o en la velocidad de la motricidad fina respecto al rastreo y la discriminación visuales. Una puntuación superior en *Búsqueda de símbolos* respecto a *Claves* sugiere que el rastreo visual preciso es un punto fuerte respecto a la memoria asociativa y/o la velocidad grafomotora.

Realización del análisis secundario de un perfil

Los pasos que componen el análisis secundario son opcionales y, en función del motivo de evaluación, pueden ser necesarios únicamente algunos de estos pasos. Generalmente, cuando la evaluación se realiza por un motivo clínico específico, es necesario realizar el análisis secundario. Para empezar, deben anotarse las puntuaciones de las pruebas y las puntuaciones compuestas en el cuadernillo de anotación, junto con los percentiles e intervalos de confianza correspondientes.

Paso 1. Descripción del IRC

El IRC es el resultado de la suma de las puntuaciones escalares de *Balanzas* y *Aritmética*, y es un indicador de las destrezas de razonamiento cuantitativo del niño. El razonamiento cuantitativo está estrechamente relacionado con la inteligencia general (Flanagan y Kaufman, 2009; Lichtenberger y Kaufman, 2013; Weiss et al., 2013a, 2013b). La evaluación del razonamiento cuantitativo permite predecir con mayor precisión las puntuaciones del rendimiento tanto en lectura como en matemáticas, la creatividad, los logros académicos futuros, el éxito en programas especializados en alta capacidad intelectual, el rendimiento en pruebas profesionales y el nivel educativo futuro (Lakin y Lohman, 2011; McGrew y Wendling, 2010; Robertson, Smeets, Lubinski y Benbow, 2010; Schmidt, Homeyer y Walker, 2009). Asimismo, evaluar estas destrezas permite adaptar la formación e intervención a los puntos fuertes del estudiante (Lohman, 2009). El IRC resulta especialmente interesante cuando se sospecha que un niño tiene un trastorno específico del aprendizaje con dificultad matemática, porque el razonamiento cuantitativo puede ser un punto débil que requiera la correspondiente intervención (Wechsler, 2003, 2008; Zheng, Flynn y Swanson, 2013). Las puntuaciones altas en el IRC sugieren que la capacidad para realizar operaciones matemáticas mentalmente y para comprender las relaciones cuantitativas está bien desarrollada. Las puntuacio-

nes bajas en el IRC pueden deberse a diversas razones, como dificultades para realizar operaciones matemáticas mentalmente o para comprender y aplicar relaciones cuantitativas, una memoria de trabajo baja o problemas generales con el razonamiento conceptual abstracto.

Paso 2. Descripción del IMTA

El IMTA es el resultado de la suma de las puntuaciones escalares de *Dígitos* y *Letras y números*, y es un indicador de las destrezas de memoria de trabajo auditiva del niño. Así como el IMT es una medida compuesta de la memoria de trabajo en distintas modalidades (auditiva y visual), el IMTA es una medida más pura de la memoria de trabajo auditiva y, por diversas razones, ofrece información relevante.

Muchos modelos conocidos de memoria de trabajo conceptualizan sistemas de memoria de trabajo específicos según el dominio. Por una parte, el modelo multicomponente establece dos sistemas de almacenamiento específicos según el dominio: el bucle fonológico para la información verbal y la agenda visoespacial para la información visual y espacial (Baddeley, 2000, 2002, 2012). Las puntuaciones altas en el IMTA indican que la capacidad para registrar, mantener y manipular información presentada verbalmente está bien desarrollada. Las puntuaciones bajas en el IMTA pueden deberse a diversas razones, como dificultades en el procesamiento auditivo, falta de atención, distractibilidad, almacenamiento y manipulación de la memoria de trabajo auditiva bajos o aptitud general de memoria de trabajo baja.

Las pruebas que conforman el IMTA de la WISC-V son idénticas a las que lo conforman en el WISC-IV. Sin embargo, respecto al IMTA de la WISC-IV, el de la WISC-V incide más en la secuenciación y la manipulación mental, porque en la prueba *Dígitos* se ha añadido la tarea en orden creciente.

Paso 3. Descripción del INV

El INV es el resultado de la suma de seis puntuaciones escalares de pruebas que no requieren respuestas verbales. Las pruebas que lo conforman miden cuatro de los cinco dominios cognitivos primarios: Visoespacial, Razonamiento fluido, Memoria de trabajo y Velocidad de procesamiento.

El INV puede interpretarse como una medida de la aptitud intelectual general que reduce al mínimo la intervención del lenguaje expresivo para permitir evaluar a niños cuyas circunstancias son especiales (p. ej., niños que están aprendiendo español) o a niños con necesidades clínicas (p. ej., niños con trastorno del espectro autista con deterioro del lenguaje). La ausencia de requisitos expresivos también hace que este índice sea útil para medir la aptitud cognitiva general de los niños sordos o con dificultades de audición.

El INV ha demostrado su utilidad como medida de la aptitud intelectual general con niños que están aprendiendo un idioma (Raiford y Coalson, 2014; Wechsler, 2012). Sin embargo, el INV no debe conceptualizarse como una medida sin intervención del lenguaje; es más apropiado considerarlo una medida «con menor intervención del lenguaje», puesto que se obtiene a partir de pruebas que requieren que el niño entienda consignas verbales. Las puntuaciones compuestas, como el INV, pueden utilizarse adecuadamente en procedimientos establecidos de evaluación intelectual de niños que están aprendiendo un idioma (Ortiz et al., 2012).

El DSM-5 especifica que, en determinadas situaciones, es necesario utilizar una medida no verbal de aptitud para determinar que los síntomas de un niño cumplen los criterios diagnósticos. Por ejemplo, el INV es útil si ya se ha diagnosticado discapacidad intelectual y se está considerando el diagnóstico comórbido de trastorno del lenguaje. Asimismo, si un niño con deterioro del lenguaje, al que se está evaluando por un posible diagnóstico de discapacidad intelectual, no obtiene puntuaciones válidas en las pruebas de Comprensión verbal que contribuyen al CI total, el INV es una alternativa útil (Raiford y Coalson, 2014). Las puntuaciones altas en el INV indican que el funcionamiento intelectual general para los estímulos presentados visualmente está bien desarrollado. Las puntuaciones bajas en el INV pueden deberse a diversas razones, como velocidad de procesamiento baja, memoria de trabajo baja, aptitudes de razonamiento abstracto y conceptual bajas, razonamiento espacial bajo o aptitud intelectual general baja.

En otros casos, el INV puede ser útil para evaluar las discrepancias entre aptitud y rendimiento para establecer el diagnóstico de un trastorno específico del aprendizaje. Por ejemplo, al evaluar a un niño con trastorno del espectro autista con deterioro del lenguaje para determinar un trastorno específico del lenguaje, el deterioro del lenguaje puede dar lugar a menores diferencias entre el CI total y las puntuaciones de rendimiento. En este caso, el INV puede ser una alternativa útil (Raiford y Coalson, 2014).

Paso 4. Descripción del ICG

El ICG se basa en las pruebas de Comprensión verbal, Visoespacial y Razonamiento fluido que contribuyen al CI total. El ICG se desarrolló específicamente para ayudar a los profesionales en la identificación de los puntos fuertes y débiles, a partir de comparaciones entre la aptitud general y otras funciones cognitivas. Conceptualmente, el ICG ofrece una estimación de la aptitud intelectual general menos dependiente de la memoria de trabajo y de la velocidad de procesamiento que el CI total. Las puntuaciones altas en el ICG indican que el razonamiento abstracto y conceptual, el razonamiento visoperceptivo y espacial y la resolución de problemas verbales están bien desarrollados. Las puntuaciones bajas en el ICG pueden deberse a diversas razones, como destrezas de razonamiento insuficientes, dificultades en el procesamiento visoespacial, déficits del lenguaje o aptitud intelectual general baja.

Dado que el CI total está formado por pruebas de memoria de trabajo y de velocidad de procesamiento, los niños con trastornos del neurodesarrollo asociados a dificultades en la memoria de trabajo y la velocidad de procesamiento, tales como trastorno del aprendizaje, TDAH, trastornos del lenguaje o trastornos del espectro autista, pueden obtener puntuaciones del CI total más bajas que los niños que no tienen estos trastornos (Akbar, Loomis y Paul, 2013; Horowitz-Kraus, 2014; Kasper, Alderson y Hudec, 2012; Niileksela y Reynolds, 2014; Pimperton y Nation, 2014; Thaler, Bello y Etcoff, 2012; Travers et al., 2014; Vugs, Cuperus, Hendriks y Verhoeven, 2013). En estos casos, el CI total más bajo puede ocultar diferencias significativas entre la aptitud cognitiva general (representada por el CI total) y otras funciones cognitivas (p. ej., rendimiento y memoria).

Sin embargo, el ICG no sustituye al CI total como estimación de la aptitud general; sino que debe registrarse e interpretarse junto con el CI total y todas las puntuaciones de los índices primarios, incluidos el IMT y el IVP. La memoria de trabajo y la velocidad de procesamiento son fundamentales para realizar una evaluación completa de la aptitud cognitiva general (Blalock y McCabe, 2011; Bunting, 2006; Dodonova y Dodonov, 2012). No incluir las medidas de estas aptitudes en la puntuación global reduce la amplitud de constructo y la validez predictiva de dicha puntuación (Rowe et al., 2010).

Algunos autores abogan a favor de utilizar el ICG como criterio para la asignación de niños a programas especializados en alta capacidad intelectual o similares (Rimm et al., 2008). Se ha observado que algunos niños que participan en programas especializados en alta capacidad intelectual tienden a obtener puntuaciones más altas en el ICG que en el ICC o el CI total. Por ejemplo, el 34% de los niños con alta capacidad intelectual que participaron en el estudio con grupos especiales obtuvieron diferencias en las que el ICG era superior al CI total. Este resultado, aunque en menor porcentaje, es similar al obtenido con la WISC-IV (Raiford et al., 2005), la WPPSI-IV (Wechsler, 2012) y la WAIS-IV (Wechsler, 2008). Además el 72% de los niños del grupo con alta capacidad intelectual obtuvieron diferencias en las que el ICG era superior al ICC. Asimismo se observaron patrones similares y porcentajes más altos en los niños de la muestra de tipificación que mostraron un nivel alto de aptitud general (ICG ≥ 120), lo que indica que frecuentemente la memoria de trabajo y la velocidad de procesamiento no son las áreas de rendimiento más fuertes en los niños con una aptitud intelectual alta.

No obstante, los puntos débiles normativos o intrapersonales en la competencia cognitiva no pueden observarse si no se han aplicado las pruebas de memoria de trabajo y de velocidad de procesamiento, puesto que estos constructos se consideran componentes esenciales del constructo de la inteligencia (Weiss et al., 2006). Asimismo, estudios recientes indican que la velocidad de procesamiento de los niños con alta capacidad intelectual puede incrementarse si participan en

programas especializados en alta capacidad intelectual (Duan, Shi y Zhou, 2010). Por ello es importante realizar una evaluación de la velocidad de procesamiento antes de aplicar un programa especializado en alta capacidad intelectual, para garantizar la comprensión de los resultados y el impacto del programa (Raiford y Coalson, 2014).

El profesional puede calcular e interpretar el ICG y el ICC y evaluar la diferencia correspondiente en diversas situaciones clínicas, como las siguientes:

- una diferencia significativa e infrecuente entre el IMT y la MIP o el CI total;
- una diferencia significativa e infrecuente entre el IVP y la MIP o el CI total;
- una diferencia significativa e infrecuente entre el ICV y el IMT;
- una diferencia significativa e infrecuente entre el ICV y el IVP;
- una diferencia significativa e infrecuente entre el IVE y el IMT;
- una diferencia significativa e infrecuente entre el IVE y el IVP;
- una diferencia significativa e infrecuente entre el IRF y el IMT;
- una diferencia significativa e infrecuente entre el IRF y el IVP;
- una diferencia significativa e infrecuente entre el IMT y el IVP;
- una diferencia significativa e infrecuente entre las pruebas que contribuyen al IMT o al IVP;
- una diferencia significativa e infrecuente entre las pruebas de Memoria de trabajo o de Velocidad de procesamiento y la MPE-P o la MPE-CI.

Paso 5. Descripción del ICC

El ICC es una puntuación compuesta que se basa en las pruebas que contribuyen al IMT y al IVP. Conceptualmente, el ICC ofrece una estimación de la eficacia con la que se procesa la información durante el aprendizaje, la resolución de problemas y el razonamiento de nivel superior. Las puntuaciones altas en el ICC indican un nivel alto de eficacia cognitiva para manipular y procesar rápidamente información. Las puntuaciones bajas en el ICC pueden deberse a diversas razones, como déficits en el procesamiento visual o auditivo, falta de atención, distractibilidad, dificultades visomotoras, capacidad limitada de almacenamiento de la memoria de trabajo o manipulación mental, o aptitud cognitiva generalmente baja.

El ICC ofrece más información cuando se interpreta junto con el ICG. El profesional puede calcular, interpretar y evaluar la diferencia entre el ICG y el ICC en diversas situaciones clínicas. La comparación de la diferencia entre el IMT y el IVP también es relevante para interpretar el ICC; en el apartado «Paso 8. Evaluación de las comparaciones entre los índices» de la sección «Realización del análisis primario de un perfil» de este mismo capítulo se describen los aspectos fundamentales de su interpretación.

Paso 6. Evaluación de las comparaciones entre las puntuaciones compuestas de los índices secundarios

En la página de análisis secundario del cuadernillo de anotación se incluye un cuadro para el análisis de las diferencias entre los índices. La tabla C.6 del anexo C del *Manual de aplicación y corrección* presenta los valores críticos para las diferencias estadísticamente significativas con los niveles de significación .01, .05, .10, y .15, por grupo de edad y muestra total; y las tablas C.7 y C.8 de dicho manual presentan las tasas base de estas diferencias. Las tasas base de la tabla C.7 figuran en las dos direcciones de la diferencia, para la muestra total y según el nivel de aptitud reflejado en el ICG, y las tasas base de la tabla C.8 figuran en las dos direcciones de la diferencia, para la muestra total y según el nivel de aptitud reflejado en el CI total. La interpretación de las diferencias estadísticamente significativas con tasas base bajas se expone en cada comparación.

ICG y CI total

Respecto al CI total, el ICG proporciona una estimación de la aptitud intelectual general menos influida por las dificultades en la memoria de trabajo y la velocidad de procesamiento. Se puede comparar el CI total y el ICG para evaluar los efectos de la competencia cognitiva (medida por las pruebas de Memoria de trabajo y Velocidad de procesamiento) sobre el funcionamiento intelectual general del niño. La evaluación de la significación y frecuencia de las diferencias entre el ICG y el CI total puede servir para decidir cuándo se debe comparar el ICG con otras puntuaciones de procesamiento cognitivo en determinadas situaciones clínicas. Diversas afecciones clínicas, como el trastorno específico del lenguaje, el TDAH, el trastorno del espectro autista y los trastornos del lenguaje (Raiford et al., 2005; Theiling y Petermann, 2014; Wang, Jing, Jin, Liu y Wang, 2013; Wechsler, 2008), así como la epilepsia (Baxendale, McGrath y Thompson, 2014), muestran un patrón de puntuaciones significativamente más altas en el ICG que en el CI total.

Si el ICG y el CI total no son significativamente diferentes, implica que reducir la influencia de la memoria de trabajo y de la velocidad de procesamiento en la estimación de la aptitud intelectual general produce una pequeña diferencia, o ninguna, en el rendimiento general del niño. Una puntuación superior en el ICG respecto al CI total sugiere que la estimación de la aptitud intelectual general es menor por la inclusión de las pruebas de memoria de trabajo y de velocidad de procesamiento, y que estas aptitudes pueden constituir puntos débiles específicos. Una puntuación superior en el CI total respecto al ICG sugiere que la memoria de trabajo y la velocidad de procesamiento son puntos fuertes que refuerzan la aptitud intelectual general del niño.

Los profesionales deben tener en cuenta que diferencias relativamente pequeñas entre el ICG y el CI total pueden ser significativas (estadísticamente significativas e infrecuentes) en la WISC-V. Ello es debido a que el ICG y el CI total comparten el contenido de las pruebas, a excepción de dos pruebas. En comparación, el CI total de la WISC-IV incluye cuatro pruebas que no contribuyen al ICG. Por ello, las diferencias más pequeñas entre el ICG y el CI total pueden ser más significativas en la WISC-V que en la WISC-IV.

ICG e ICC

Dado que el CI total y el ICG de la WISC-V tienen muchas pruebas comunes, la comparación del ICG con el ICC puede ofrecer más información para la interpretación de los resultados. Esta comparación es psicométrica y composicionalmente distinta a la comparación del ICG con el CI total. El ICC incluye pruebas principales de Memoria de trabajo y de Velocidad de procesamiento que no contribuyen al CI total, lo que proporciona una mayor amplitud de constructo de la competencia cognitiva. El ICG puede compararse con el ICC para obtener más información sobre el posible efecto que los déficits cognitivos básicos pueden tener sobre la aptitud intelectual general del niño.

Los datos recopilados sugieren que la comparación del ICG y el ICC puede ser útiles cuando se interpretan en el conjunto de una evaluación completa. Las investigaciones indican que el porcentaje de niños en edad escolar con trastornos específicos del lenguaje que obtienen puntuaciones más bajas en el ICC y en medidas del rendimiento académico que en el ICG es considerablemente más alto que el porcentaje de niños del grupo control (Weiss y Gabel, 2008). También se ha observado un patrón de puntuaciones significativamente más bajas en el ICC que en el ICG en niños en edad escolar con epilepsia (Gottlieb, Zelko, Kim y Nordli, 2012) y en niños con TDAH con predominio del déficit de atención, pero no con TDAH de tipo combinado (Bremner, McTaggart, Saklofske y Janzen, 2011). La diferencia entre el ICG y el ICC es un indicador significativo de pertenencia a un grupo clínico de trastorno específico del aprendizaje-lector, discapacidad intelectual moderada, traumatismo craneoencefálico y deterioro motor (Zhu y Chen, 2013). Algunas afecciones neurocognitivas, como el trastorno específico del aprendizaje, el TDAH y el traumatismo craneoencefálico dan lugar a diferencias más grandes entre el ICC y el ICG (siendo el ICC inferior al ICG) que las observadas en los trastornos psicológicos (Harrison, DeLisle y Parker, 2008). Sin embargo, las diferencias entre el ICG y el ICC por sí solas no deben considerarse indicadores diagnósticos del TDAH (Devena y Watkins, 2012) ni de otros trastornos.

Si el ICG y el ICC no son significativamente diferentes, ello implica que la aptitud general es acorde con la competencia cognitiva. Una puntuación superior en el ICG respecto al ICC indica que las aptitudes cognitivas de nivel superior (p. ej., comprensión verbal, procesamiento visoespacial y razonamiento fluido) pueden ser puntos fuertes relativos en comparación con las aptitudes que favorecen la eficacia del procesamiento cognitivo (p. ej., la memoria de trabajo y la velocidad de procesamiento). Dicho de otro modo, es posible que las limitaciones en la eficacia cognitiva no reduzcan la aptitud de razonamiento general del niño. Una puntuación superior en el ICC respecto al ICG sugiere que las aptitudes que favorecen la eficacia del procesamiento cognitivo pueden ser un punto fuerte en comparación con las aptitudes cognitivas de nivel superior. Por ello, el hecho de que las destrezas de razonamiento sean bajas no se debe a las limitaciones en la eficacia cognitiva.

IMT e IMTA

El IMT está compuesto por pruebas que utilizan estímulos verbales y visuales, y en la prueba *Span de dibujos* se admiten tanto respuestas verbales como no verbales. El IMTA está compuesto por dos pruebas que utilizan estímulos presentados verbalmente y requieren respuestas verbales. La comparación del IMT y el IMTA puede ofrecer información sobre el efecto de la especificidad del dominio (visual y auditivo, respectivamente) en el funcionamiento de la memoria de trabajo. Si el IMT y el IMTA no son significativamente diferentes, ello sugiere que la memoria de trabajo no es distinta en función del dominio. Una puntuación superior en el IMT respecto al IMTA sugiere que la presentación visual de la información puede mejorar el funcionamiento de la memoria de trabajo respecto a la presentación verbal. Una puntuación superior en el IMTA respecto al IMT sugiere que la presentación verbal de la información puede mejorar el funcionamiento de la memoria de trabajo respecto a la presentación visual. En el siguiente apartado, se describe la comparación de las diferencias entre *Dígitos* y *Letras y números* y entre *Dígitos* en orden creciente y *Letras y números*, y se ofrece más información sobre la interpretación de las diferencias significativas relacionadas con la memoria de trabajo.

Paso 7. Evaluación de las comparaciones entre las pruebas de los índices secundarios

La comparación de pruebas que contribuyen a las puntuaciones de índices obtenidos a partir de dos o tres pruebas es relevante para la interpretación de estos índices y permite describir con mayor precisión el rendimiento mostrado en ellos. La tabla B.7 del *Manual de aplicación y corrección* presenta los valores críticos de las diferencias estadísticamente significativas entre cada par de pruebas con los niveles de significación .01, .05, .10 y .15. La tabla B.8 del mismo manual presenta las tasas base bidireccionales de las diferencias entre las puntuaciones de determinadas pruebas obtenidas por la muestra de tipificación. A continuación se describen las comparaciones entre pruebas que figuran en la página de análisis secundario del cuadernillo de anotación.

Balanzas y Aritmética

La comparación *Balanzas-Aritmética* proporciona información sobre el rendimiento en tareas de Razonamiento cuantitativo que se diferencian por el tipo de estímulo y de respuesta. En *Balanzas* se utilizan estímulos abstractos visuales y no se requiere que la respuesta sea verbal. En *Aritmética* se utilizan estímulos verbales concretos, excepto en los ítems gráficos, y se requiere que la respuesta sea verbal. Los requisitos de memoria de trabajo son menores en *Balanzas* que en *Aritmética*, puesto que, en *Balanzas*, el estímulo sigue estando presente mientras se resuelve el ítem. Una puntuación superior en *Balanzas* respecto a *Aritmética* puede indicar que el niño puede expresar mejor su aptitud de razonamiento cuantitativo cuando el estímulo y la respuesta son visuales, en vez de verbales. Además, el niño puede beneficiarse del hecho de que los requisitos de memoria de trabajo son menores en las tareas de razonamiento cuantitativo. Una puntuación superior en *Aritmética* respecto a *Balanzas* puede indicar que el niño puede expresar mejor su aptitud de razonamiento cuantitativo cuando el estímulo y la respuesta son verbales, en vez de visuales. Asimismo, es posible que al niño le resulten más sugerentes las tareas con mayores requisitos de memoria de trabajo.

Dígitos y Letras y números

La comparación de la diferencia entre *Dígitos* y *Letras y números* proporciona información sobre el rendimiento en tareas de Memoria de trabajo que utilizan un único tipo de estímulo y dos tipos de estímulos, respectivamente. En *Dígitos*, el niño debe pasar de una tarea a otra durante la aplicación de la prueba: repetir series de números en orden directo, repetir series de números en orden inverso y reordenar series de números en orden creciente; mientras que en *Letras y números*, el niño debe reordenar series formadas por números y letras en orden numérico y alfabético, respectivamente.

Una puntuación superior en *Dígitos* respecto a *Letras y números* puede indicar que el niño tiene dificultades con el registro inmediato de letras o que no ha aprendido aún las destrezas alfabéticas. Asimismo, el niño puede tener dificultades en la memoria de trabajo por tratarse de una tarea dual. Una puntuación superior en *Letras y números* respecto a *Dígitos* puede indicar que, dado que *Dígitos* se aplica antes que *Letras y números*, el niño ha aprendido a usar determinadas estrategias eficaces mientras realizaba las distintas tareas de la prueba, o bien que mejora fácilmente con la experiencia y la práctica. En *Letras y números,* la tarea de ordenar es menos compleja puesto que, en cada ítem, los elementos a ordenar son menos. Por ejemplo, en *Letras y números*, un ítem puede estar formado por seis elementos, tres de ellos son números, y estos son los que deben compararse y ordenarse correctamente, mientras que los tres elementos restantes, que son tres letras, se mantienen almacenadas para ordenarlas seguidamente. En *Dígitos*, un ítem formado por seis elementos, todos números, requiere que se ordenen todos los elementos, los seis números. Además, esta prueba requiere que el niño repita los elementos de la serie en orden inverso, lo cual puede resultar más complejo para algunos niños que ordenar mentalmente los elementos. Una puntuación superior en *Letras y números* respecto a *Dígitos* puede indicar dificultades con la tarea de ordenar cadenas más largas, mayor implicación en las tareas duales o más dificultad con los requisitos de determinadas tareas (p. ej., presentación en orden inverso, números repetidos). En el siguiente apartado, al describir el análisis de procesamiento de *Dígitos* y *Letras y números* y la interpretación de *Dígitos* en orden creciente y *Letras y números*, se ofrece más información sobre la interpretación de las diferencias significativas relacionadas con la memoria de trabajo.

Paso 8. Realización del análisis de procesamiento

El enfoque de procesamiento para la interpretación defendido por Kaplan (1988) y Kaplan, Fein, Morris y Delis (1991) fue diseñado para ayudar al profesional a determinar la naturaleza de los errores cometidos en los tests estandarizados. Aunque cada una de las pruebas de una escala de inteligencia se ha diseñado para medir un proceso cognitivo determinado, al realizar la tarea, intervienen otros procesos cognitivos. Son varias las razones por las que un sujeto puede obtener 0 puntos en un ítem. Por ejemplo, los niños pueden fallar un ítem de *Cubos* porque no perciben correctamente el diseño, no son capaces de analizar la configuración del diseño o agotan el tiempo concedido para su realización.

Las puntuaciones de procesamiento ofrecen una información más detallada de las aptitudes cognitivas que intervienen en el rendimiento del niño en las distintas pruebas. El análisis de procesamiento puede ayudar a los profesionales a evaluar determinados estilos de procesamiento, que pueden ser importantes para describir los puntos fuertes y débiles cognitivos, tomar decisiones diagnósticas o establecer estrategias de rehabilitación como parte de un plan de tratamiento o un programa educativo. El análisis de procesamiento puede realizarse desde un punto de vista interindividual mediante la comparación de las puntuaciones de procesamiento del niño con las obtenidas por otros niños de su misma edad. El análisis del procesamiento intraindividual se lleva a cabo evaluando las puntuaciones de procesamiento del sujeto en relación con otras puntuaciones obtenidas (puntuaciones de pruebas o de procesamiento).

En la página de análisis de procesamiento del cuadernillo de anotación se incluye un cuadro para que el profesional realice el análisis de procesamiento de cinco pruebas: *Cubos, Dígitos, Cancelación, Span de dibujos* y *Letras y números*. El cuadro «Conversión de puntuaciones directas a puntuaciones de procesamiento escalares» se usa junto con la tabla C.9 del *Manual de aplicación y corrección* para convertir las puntuaciones directas en puntuaciones de procesamiento escalares.

Estas puntuaciones expresan el rendimiento del niño en comparación con el de otros niños de la misma edad que formaron parte de la muestra de tipificación.

Para examinar los patrones intraindividuales de procesamiento de la información, en el cuadernillo de anotación se incluye un cuadro para realizar las comparaciones entre puntuaciones de procesamiento. La tabla C.10 del *Manual de aplicación y corrección* presenta los valores críticos de las diferencias estadísticamente significativas para estas comparaciones con los niveles de significación .01, .05, .10 y .15. La tabla C.11 presenta las tasas base de las diferencias obtenidas por la muestra de tipificación. Estas tasas base se presentan en las dos direcciones de la diferencia.

La interpretación del análisis de procesamiento de las pruebas se ofrece al principio del apartado correspondiente a cada prueba.

Cubos

La puntuación de procesamiento de *Cubos* sin bonificación por tiempo (Cs) se basa en el rendimiento del niño en la prueba *Cubos* sin conceder puntos extra por rapidez en la resolución de los ítems. El hecho de incidir menos en la velocidad puede resultar de gran utilidad cuando se sospecha que las limitaciones físicas del niño, sus estrategias de resolución de problemas o su personalidad afectan a su rendimiento en las pruebas con límite de tiempo.

La puntuación de procesamiento de *Cubos* parcial (Cp) se basa en el número de cubos correctamente colocados dentro del tiempo límite. Se conceden puntos de bonificación por tiempo en los ítems 10-13 cuando el diseño se ha realizado por completo, correcta y rápidamente. La puntuación Cp incide menos en la velocidad porque se puntúa cada uno de los cubos correctamente colocados dentro del tiempo límite; por lo tanto, aunque el diseño no esté completado al agotarse el tiempo límite, igualmente se concederán los puntos correspondientes a los cubos correctamente colocados. La puntuación Cp también incide menos en la atención al detalle y, consecuentemente, puntúa cada cubo correctamente colocado. De este modo proporciona una estimación del rendimiento mostrado por niños que son impulsivos o descuidados o que no perciben correctamente el diseño.

La diferencia entre la puntuación de *Cubos* (C) y la puntuación Cs puede ofrecer información sobre la relativa influencia de la velocidad en el rendimiento mostrado en esta prueba. Una puntuación superior en Cs respecto a C sugiere que el rendimiento en la prueba puede haberse visto afectado por dificultades con el procesamiento rápido de la información visoperceptiva o con la ejecución motora. Una puntuación superior en C respecto a Cs indica que el nivel de las aptitudes en este dominio no se debe a la velocidad de procesamiento ni a la ejecución. La puntuación C y la puntuación Cs son generalmente muy similares. En la muestra de tipificación española es muy poco frecuente que la diferencia sea superior a 1 punto (< 2 %). Si la puntuación C es alta, la diferencia entre esta y la puntuación Cs debe interpretarse con cautela dado que el rango superior de Cs está limitado en los grupos de más edad.

La diferencia entre las puntuaciones C y Cp puede ofrecer información sobre la relativa influencia de la atención al detalle y la velocidad en el rendimiento mostrado en esta prueba. Una puntuación superior en Cp respecto a C sugiere que el rendimiento en la prueba se ha visto afectado por dificultades con el procesamiento rápido de la información visoperceptiva, con la ejecución motora o con la atención al detalle. Una puntuación superior en C respecto a Cp indica que el nivel de las aptitudes en este dominio no se debe a problemas con la velocidad de procesamiento ni la atención al detalle. Al igual que sucede con las puntuaciones C y Cs, las puntuaciones C y Cp son generalmente muy similares. En la muestra de tipificación española es muy poco frecuente que la diferencia sea superior a 2 puntos (< 5 %).

Dígitos

Puntuaciones de procesamiento escalares de *Dígitos*

Las puntuaciones de procesamiento escalares de *Dígitos* en orden directo (Dd), *Dígitos* en orden inverso (Di) y *Dígitos* en orden creciente (Dc) se obtienen a partir de las puntuaciones directas totales de las correspondientes tareas de *Dígitos*. En todas las tareas se usa el mismo tipo de estímulos: los números del 1 al 9; sin embargo el proceso de respuesta que debe seguir el niño es dis-

tinto en cada tarea. En la tarea en orden directo, el niño debe repetir literalmente la secuencia que se le ha leído. La dificultad de la tarea aumenta conforme avanzan los ítems; por ello, se requiere la intervención de la memoria de trabajo cuando el número de dígitos excede la capacidad del niño para repetirlos sin posibilidad de reexposición al estímulo. Esta tarea representa la capacidad básica del bucle fonológico. En la tarea en orden inverso interviene la memoria de trabajo porque el niño debe repetir los dígitos en un orden distinto (inverso) al que se le han presentado, lo que requiere la manipulación de la información antes de dar la respuesta. La tarea en orden creciente requiere que el niño ordene los dígitos por su valor, por lo que interviene el conocimiento cuantitativo además de la memoria de trabajo. Los mayores requisitos de manipulación mental de la información en la tarea en orden inverso implican una mayor participación de la memoria de trabajo y la atención, al igual que sucede en tareas similares, como la de orden creciente (de Jonge y de Jong, 1996; MacDonald et al., 2001; Sattler, 2008a; Wechsler, 2008; Werheid et al., 2002).

La comparación de las diferencias entre las puntuaciones Dd y Di y entre Dd y Dc ofrece información sobre la capacidad de almacenamiento respecto a la aptitud de manipulación mental en tareas con distinto nivel de demanda: cognitivamente más sencilla (en orden directo) frente a cognitivamente más complejas (en orden inverso y en orden creciente). Una puntuación superior en Dd respecto a Di o en Dd respecto a Dc sugiere que el niño tiene una capacidad suficiente, pero no ha aprendido la destreza de invertir el orden de los números u ordenarlos de menor a mayor mentalmente, no ha entendido bien los requisitos de las tareas más complejas (orden inverso y orden creciente) o tiene dificultades con la manipulación mental en las tareas en orden inverso o en orden creciente. Asimismo puede indicar que el niño tiene dificultades con la mayor intervención de la memoria de trabajo. Una puntuación superior en Di respecto a Dd o en Dc respecto a Dd puede deberse a distintos motivos. Es posible que el niño prefiera las tareas más complejas (p. ej., en orden inverso o en orden creciente), porque suponen un reto y se siente más motivado, que las tareas más sencillas (p. ej., en orden directo). También puede deberse a que el niño ha sufrido un lapsus de atención o motivación durante la realización de la tarea en orden directo. Asimismo, esta diferencia puede indicar que, dado que la tarea en orden directo se aplica antes que las tareas en orden inverso y en orden creciente, el niño aprende a usar determinadas estrategias eficaces mientras realiza las tareas, o mejora con la experiencia y la práctica, más fácilmente que los niños de su misma edad. Las puntuaciones de procesamiento escalares Dd, Di y Dc muestran correlaciones moderadas en la muestra de tipificación, lo que también demuestra que las tareas guardan relación entre ellas, pero no son equivalentes desde la perspectiva del procesamiento cognitivo.

La diferencia entre las puntuaciones Di y Dc es la comparación del rendimiento entre dos tareas de memoria de trabajo con requisitos relativamente complejos. Una puntuación superior en Di respecto a Dc puede indicar que el niño aún no ha aprendido la destreza de ordenar los números de menor a mayor o que no ha entendido bien la tarea por el hecho de que en algunos ítems de la tarea en orden creciente hay números repetidos. Una puntuación superior en Dc respecto a Di sugiere que el niño no ha aprendido adecuadamente la destreza de invertir el orden de los números, o bien que ha aprendido con más facilidad a ordenar los números de menor a mayor. Asimismo, esta diferencia puede indicar que, dado que la tarea en orden inverso se aplica antes que la tarea en orden creciente, el niño aprende a usar determinadas estrategias eficaces mientras realiza las tareas, o mejora con la experiencia y la práctica, más fácilmente que los niños de su misma edad. Una diferencia entre estas puntuaciones, en cualquiera de las dos direcciones, también puede indicar que el niño ha sufrido un lapsus de atención o motivación durante la realización de la tarea en la que ha obtenido un resultado más bajo.

Es también posible comparar la puntuación Dc con la de *Letras y números;* la interpretación de esta comparación figura en el apartado correspondiente a la prueba *Letras y números.*

Puntuaciones de procesamiento directas de *Dígitos*

A veces, un niño muestra un rendimiento variable entre los ítems (p. ej., puntuaciones de 1 y de 0 en los intentos de un mismo ítem, y/o entre ítems), en lugar de un patrón más típico (p. ej., respuestas correctas para los primeros ítems, seguidas de alguna puntuación de 0 justo antes de cumplir la regla de terminación). Un patrón de respuesta variable puede dar como resultado puntuaciones Dd, Di y Dc similares a las obtenidas por niños con patrones de respuestas más

típicos, a pesar de que el rendimiento de los niños sea distinto. Para evaluar con mayor detalle el rendimiento en las tres tareas de *Dígitos*, se añadieron tres puntuaciones de procesamiento adicionales: span de *Dígitos* en orden directo (SpanDd), span de *Dígitos* en orden inverso (SpanDi) y span de *Dígitos* en orden creciente (SpanDc). Estas puntuaciones directas representan el número de dígitos recordados en el último intento puntuado con 1 punto en cada una de las tareas de *Dígitos*.

Las diferencias entre las puntuaciones directas de span de *Dígitos* (SpanDd, SpanDi y SpanDc) proporcionan información cuando se observa variabilidad en el rendimiento entre los intentos o ítems de longitud similar, así como entre las tareas en orden directo, en orden inverso y en orden creciente. Estas diferencias pueden interpretarse de forma similar a las descritas anteriormente para las puntuaciones de procesamiento escalares de las tareas de *Dígitos*.

Span de dibujos

Al igual que en *Dígitos*, un niño con un rendimiento variable en los ítems puede obtener una puntuación escalar similar a la obtenida por un niño con un patrón de respuestas típico. Un niño puede obtener esporádicamente puntuaciones de 2, 1 o 0 puntos en ítems que tienen el mismo número de estímulos y/o de opciones de respuesta, en vez de mostrar un patrón de rendimiento más típico (p. ej., respuestas de 2 puntos en los ítems iniciales, seguidas de alguna puntuación de 1 o 0 en los ítems con mayor número de estímulos y/o de opciones de respuesta justo antes de cumplir la regla de terminación). Un patrón de respuesta variable puede dar como resultado puntuaciones escalares en *Span de dibujos* similares a las obtenidas por niños con patrones de respuestas más típicos, a pesar de que el rendimiento de los niños sea distinto. Para evaluar con mayor detalle el rendimiento en *Span de dibujos*, se añadieron dos puntuaciones de procesamiento adicionales: span de estímulos de *Span de dibujos* (SpaneSD) y span de respuestas de *Span de dibujos* (SpanrSD). Dado que el número de opciones de respuesta cambia a lo largo de los ítems (cuando aumenta la longitud del estímulo, el número de opciones de respuesta en la página puede ser menor), la puntuación SpanrSD debe interpretarse en relación con la puntuación SpaneSD. Estas puntuaciones directas representan el número de estímulos y de opciones de respuesta, respectivamente, que forman el último ítem en el que el niño ha obtenido una puntuación perfecta.

Letras y números

Al igual que en *Dígitos*, un niño con un rendimiento variable en los intentos y/o ítems de *Letras y números* puede obtener una puntuación escalar similar a la obtenida por un niño con un patrón de respuestas típico. La puntuación de span de *Letras y números* (SpanLN) permite obtener más información sobre el rendimiento en *Letras y números* de un niño cuyo patrón de respuestas es variable. La puntuación SpanLN es una puntuación directa que representa el número de cifras y letras recordadas en el último intento puntuado con 1 punto.

La comparación de la diferencia entre la puntuación de *Dígitos* en orden creciente (Dc) y la puntuación de *Letras y números* (LN) proporciona información sobre el rendimiento en dos tareas de memoria de trabajo que utilizan un único tipo de estímulo y dos tipos de estímulos, respectivamente. *Dígitos* en orden creciente requiere que el niño reordene series de números de menor a mayor, mientras que Letras y números requiere que el niño reordene, al mismo tiempo, series de números y de letras en orden numérico ascendiente y alfabético, respectivamente. Una puntuación superior en Dc respecto a LN puede indicar que el niño tiene dificultades con el registro inmediato de letras, no ha aprendido aún las destrezas alfabéticas o tiene problemas con el procesamiento de la información de una tarea dual en la memoria de trabajo. Una puntuación superior en LN respecto a Dc puede indicar que el niño no ha entendido bien la tarea en orden creciente por el hecho que en algunos ítems hay números repetidos, o bien tiene dificultades para reordenar series de números de mayor longitud. Asimismo, esta diferencia puede indicar que, dado que *Dígitos* en orden creciente se aplica antes que la prueba *Letras y números*, el niño aprende fácilmente a usar determinadas estrategias eficaces mientras realiza las tareas, o mejora con la experiencia y la práctica. Puesto que en estas pruebas no se permite repetir los intentos, una diferencia entre estas puntuaciones, en cualquiera de las dos direcciones, también puede indicar que el niño ha sufrido un lapsus de atención o motivación durante la realización de la tarea en la que ha obtenido un resultado más bajo.

Cancelación

Las puntuaciones de procesamiento *Cancelación* aleatoria (CAa) y *Cancelación* estructurada (CAe) se crearon para ofrecer más información sobre las aptitudes cognitivas que influyen en el rendimiento del niño en la prueba *Cancelación*. Las tareas correspondientes miden la atención selectiva y la velocidad de procesamiento del niño ante dos presentaciones distintas de los estímulos: aleatoria y estructurada. A diferencia de otras tareas de cancelación que requieren que el niño realice simplemente la correspondencia visoperceptiva entre los estímulos, la prueba *Cancelación* de la WISC-V requiere que el niño determine si los estímulos pertenecen a una determinada categoría (animales). Para discriminar los estímulos que ha de buscar y los distractores, el niño también puede realizar asociaciones visosemánticas.

La comparación de la diferencia entre las puntuaciones CAa y CAe proporciona información sobre la variación del rendimiento en *Cancelación* en función de la disposición de los estímulos visuales. Esta diferencia es útil para discriminar entre los niños del grupo control y los niños con trastornos de la lectura, discapacidad intelectual leve, traumatismo craneoencefálico cerrado y déficits motores (Zhu y Chen, 2013).

Una puntuación superior en CAe respecto a CAa sugiere que al niño le favorece la presentación estructurada. Una puntuación superior en CAa respecto a CAe puede indicar que el niño aplica la estrategia de resolución de una forma más efectiva en la tarea CAa que en la CAe, o que al niño no le favorece la presentación estructurada (el niño puede haber utilizado una estrategia de búsqueda desorganizada a pesar de que la disposición de los estímulos visuales sea estructurada). Los datos obtenidos mediante la observación de las estrategias de búsqueda seguidas por el niño pueden aportar más información que permita ratificar o rechazar hipótesis relativas a la influencia que la presentación estructurada tiene sobre el rendimiento del niño en *Cancelación*.

Realización del análisis cualitativo de las respuestas

El último paso del análisis del perfil es el análisis cualitativo de las respuestas del sujeto. Debe prestarse especial atención a las respuestas poco frecuentes y al patrón de respuestas general del niño. Las respuestas poco frecuentes a menudo ofrecen información muy reveladora. También es conveniente examinar el proceso mediante el cual el niño ha llegado a la respuesta correcta, o las razones por las que ha dado una respuesta incorrecta. El enfoque de procesamiento para la interpretación defendido por Kaplan y sus colegas (Kaplan, 1988; Kaplan et al., 1991) fue diseñado para ayudar al profesional a determinar la naturaleza de los errores que se cometían en los tests estandarizados. Aunque cada una de las pruebas de una escala de inteligencia se ha diseñado para medir un proceso cognitivo determinado, al realizar la prueba, suelen intervenir muchos procesos cognitivos. Son varias las razones por las que un sujeto puede obtener 0 puntos en un ítem. Por ejemplo, los niños pueden fallar un ítem de *Cubos* porque no perciben correctamente el diseño, no son capaces de analizar la configuración del diseño o agotan el tiempo concedido para su realización.

Resumen

La comparación y la interpretación de las puntuaciones que se registran en el cuadernillo de anotación de la WISC-V deben llevarlas a cabo solo profesionales expertos. La evaluación psicológica es una actividad clínica que utiliza las puntuaciones de los tests únicamente como una fuente más de información que permite a un psicólogo con un buen juicio clínico hacer un retrato psicológico exhaustivo de un niño. El profesional que utiliza los tests como componente de la evaluación es el responsable de la interpretación de los resultados y de la toma de decisiones relacionadas con el diagnóstico, el tratamiento y la intervención.

Anexo A

Tablas de intercorrelaciones

Tabla A.1 Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas para las edades 6:0-7:11

Prueba/ punt. de procesamiento/ punt. compuesta	S	V	I	CO	C	PV	M	B	A	D	SD	LN	CL	BS	CA	Cs	Cp	Dd	Di	Dc	CAa	CAe	ICV	IVE	IRF	IMT	IVP	CIT	IRC	IMTA	INV	ICG	ICC
Semejanzas																							.61					.62				.61	
Vocabulario	.61																						.61					.59				.56	
Información	.54	.51																															
Comprensión	.41	.50	.40																														
Cubos	.35	.24	.32	.21																				.49				.47			.50	.44	
Puzles visuales	.36	.33	.37	.28	.49																			.49							.60		
Matrices	.39	.41	.36	.36	.43	.53																			.51			.62			.65	.58	
Balanzas	.45	.41	.40	.27	.34	.47	.51																		.51			.56	.30		.50	.58	
Aritmética	.38	.42	.50	.46	.41	.41	.43	.30																					.30				
Dígitos	.50	.48	.51	.40	.41	.42	.52	.43	.55																	.38		.64		.65			.46
Span de dibujos	.18	.31	.31	.28	.25	.30	.41	.22	.39	.38																.38					.43		.45
Letras y números	.40	.42	.38	.40	.42	.50	.43	.39	.56	.65	.35																			.65			
Claves	.29	.31	.30	.18	.25	.25	.33	.21	.38	.33	.34	.33															.58	.40			.39		.55
Búsqueda de símbolos	.33	.32	.33	.32	.27	.29	.43	.26	.37	.37	.34	.35	.58														.58						.57
Cancelación	.21	.25	.17	.26	.11	.17	.16	.16	.17	.17	.20	.23	.32	.43																			
Cs	.35	.24	.32	.20	.99	.50	.42	.33	.41	.40	.25	.42	.26	.26	.10																		
Cp	.37	.36	.33	.25	.95	.60	.52	.43	.50	.37	.33	.40	.30	.33	.10	.95																	
Dd	.38	.44	.44	.29	.27	.27	.33	.27	.40	.73	.25	.47	.23	.30	.18	.27	.31																
Di	.34	.37	.37	.30	.31	.33	.42	.30	.44	.80	.33	.47	.23	.24	.11	.31	.29	.43															
Dc	.48	.39	.44	.40	.41	.43	.47	.44	.51	.82	.35	.61	.35	.37	.18	.40	.39	.47	.47														
CAa	.19	.20	.14	.25	.12	.17	.17	.18	.14	.16	.15	.19	.22	.34	.90	.11	.13	.15	.11	.15													
CAe	.22	.28	.21	.24	.10	.18	.15	.18	.19	.19	.25	.23	.38	.45	.91	.09	.08	.18	.13	.20	.68												
ICV	.90	.89	.59	.50	.33	.39	.44	.48	.45	.54	.28	.46	.34	.36	.26	.33	.40	.45	.40	.49	.21	.28											
IVE	.42	.33	.40	.28	.86	.86	.56	.47	.47	.48	.32	.53	.29	.32	.16	.86	.87	.32	.37	.48	.16	.16	.42										
IRF	.47	.46	.43	.36	.44	.58	.89	.85	.42	.54	.36	.47	.31	.40	.19	.43	.54	.34	.41	.52	.20	.18	.52	.59									
IMT	.41	.47	.49	.41	.39	.43	.56	.39	.56	.82	.84	.60	.40	.42	.23	.39	.40	.58	.67	.69	.19	.27	.49	.47	.54								
IVP	.35	.36	.35	.28	.29	.31	.43	.27	.43	.40	.38	.38	.89	.89	.42	.29	.34	.30	.27	.41	.32	.47	.39	.35	.40	.46							
CIT	.74	.71	.60	.48	.62	.59	.74	.68	.59	.75	.43	.62	.57	.53	.29	.62	.64	.54	.57	.69	.25	.30	.81	.70	.82	.70	.62						
IRC	.51	.52	.56	.45	.46	.55	.58	.79	.82	.60	.38	.60	.37	.39	.21	.46	.56	.42	.46	.59	.20	.23	.57	.59	.78	.59	.43	.79					
IMTA	.50	.50	.49	.44	.46	.51	.52	.45	.62	.90	.41	.92	.37	.40	.22	.45	.43	.66	.69	.78	.19	.24	.55	.55	.55	.78	.43	.75	.66				
INV	.49	.48	.50	.38	.68	.74	.79	.66	.56	.61	.62	.59	.59	.53	.27	.67	.70	.39	.46	.59	.24	.30	.55	.82	.84	.74	.63	.89	.76	.66			
ICG	.77	.72	.58	.48	.65	.59	.76	.73	.53	.63	.37	.56	.38	.44	.24	.64	.68	.46	.47	.60	.23	.25	.84	.72	.85	.60	.47	.96	.78	.66	.85		
ICC	.44	.47	.48	.39	.40	.42	.56	.37	.57	.69	.70	.56	.77	.78	.38	.40	.42	.51	.53	.63	.30	.43	.51	.47	.54	.84	.87	.77	.59	.69	.79	.61	
Media	9.9	10.1	9.9	10.0	10.0	10.0	10.0	9.9	10.0	9.7	10.0	9.7	10.0	10.0	10.0	9.9	9.7	9.8	9.8	9.9	10.0	10.0	99.9	100.0	99.7	99.1	100.1	99.6	99.8	98.6	99.9	99.7	99.6
Dt	3.0	2.9	2.7	3.1	2.9	2.8	3.1	2.7	2.9	2.8	3.0	3.2	3.0	3.0	2.8	3.0	2.5	2.8	2.7	2.7	2.6	2.8	14.3	13.7	14.7	14.0	15.2	14.3	13.2	15.2	14.5	13.8	15.2

Tabla A.2 Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas para las edades 8:0-9:11

Prueba/ punt. de procesamiento/ punt. compuesta	S	V	I	CO	C	PV	M	B	A	D	SD	LN	CL	BS	CA	Cs	Cp	Dd	Di	Dc	CAa	CAe	ICV	IVE	IRF	IMT	IVP	CIT	IRC	IMTA	INV	ICG	ICC
Semejanzas																							.62					.57				.56	
Vocabulario	.62																						.62					.60				.60	
Información	.63	.56																															
Comprensión	.49	.56	.49																														
Cubos	.35	.38	.34	.26																				.59				.58			.63	.53	
Puzles visuales	.35	.41	.32	.42	.59																			.59							.61		
Matrices	.45	.41	.29	.32	.48	.45																			.37			.60			.57	.58	
Balanzas	.25	.37	.32	.26	.35	.40	.37																		.37			.42	.34		.45	.44	
Aritmética	.51	.45	.49	.35	.42	.39	.48	.34																					.34				
Dígitos	.46	.43	.43	.29	.41	.37	.45	.29	.55																	.42		.57		.66			.46
Span de dibujos	.30	.28	.26	.23	.31	.35	.35	.28	.33	.42																.42					.43		.43
Letras y números	.44	.38	.43	.29	.35	.30	.42	.33	.57	.66	.41																			.66			
Claves	.19	.20	.24	.15	.37	.24	.28	.14	.27	.31	.22	.28															.48	.35			.35		.45
Búsqueda de símbolos	.21	.24	.22	.17	.41	.34	.32	.20	.38	.30	.32	.29	.48														.48						.49
Cancelación	.19	.15	.12	.17	.23	.15	.11	.22	.25	.16	.05	.07	.24	.26																			
Cs	.36	.38	.34	.26	.99	.59	.48	.35	.41	.40	.31	.35	.38	.42	.23																		
Cp	.46	.54	.44	.29	.97	.65	.47	.28	.50	.49	.20	.37	.41	.49	.23	.96																	
Dd	.48	.44	.41	.31	.35	.30	.33	.23	.43	.70	.33	.51	.19	.18	.11	.35	.49																
Di	.31	.27	.30	.13	.28	.29	.35	.19	.42	.81	.40	.50	.28	.26	.18	.28	.35	.39															
Dc	.31	.33	.35	.28	.35	.30	.39	.29	.48	.83	.30	.57	.29	.29	.10	.34	.37	.40	.51														
CAa	.19	.14	.12	.19	.21	.18	.13	.19	.24	.15	.02	.06	.25	.24	.91	.22	.16	.07	.15	.13													
CAe	.18	.15	.11	.11	.21	.12	.09	.24	.23	.16	.07	.07	.21	.25	.91	.21	.22	.13	.19	.07	.69												
ICV	.90	.90	.67	.59	.41	.43	.48	.34	.54	.50	.33	.46	.23	.26	.20	.42	.54	.51	.33	.36	.18	.19											
IVE	.39	.45	.37	.38	.89	.89	.52	.42	.45	.43	.37	.37	.34	.42	.21	.88	.90	.36	.32	.36	.22	.18	.47										
IRF	.41	.46	.37	.34	.50	.51	.81	.84	.49	.44	.38	.45	.24	.31	.20	.49	.44	.33	.31	.40	.20	.20	.49	.56									
IMT	.45	.41	.41	.32	.42	.42	.47	.34	.51	.84	.84	.64	.32	.36	.13	.41	.39	.61	.71	.67	.09	.13	.49	.47	.48								
IVP	.24	.26	.28	.19	.45	.34	.36	.20	.38	.37	.31	.35	.86	.86	.29	.46	.51	.22	.32	.35	.28	.26	.29	.44	.33	.40							
CIT	.70	.72	.59	.49	.71	.60	.73	.61	.64	.71	.46	.62	.51	.45	.28	.71	.73	.57	.52	.59	.27	.26	.79	.74	.80	.68	.57						
IRC	.46	.50	.49	.37	.47	.49	.52	.84	.79	.50	.37	.54	.25	.34	.28	.46	.46	.39	.36	.46	.26	.29	.53	.54	.83	.51	.35	.76					
IMTA	.48	.42	.45	.31	.41	.36	.46	.32	.60	.91	.45	.91	.31	.31	.11	.40	.44	.66	.71	.77	.10	.11	.51	.43	.47	.81	.37	.72	.56				
INV	.46	.50	.42	.39	.77	.76	.72	.65	.54	.54	.61	.52	.53	.50	.24	.76	.74	.41	.43	.46	.24	.22	.54	.86	.82	.68	.60	.89	.73	.57			
ICG	.72	.76	.58	.51	.71	.61	.75	.66	.60	.55	.41	.53	.32	.38	.25	.71	.73	.50	.38	.46	.24	.23	.82	.74	.85	.57	.41	.96	.78	.58	.86		
ICC	.41	.40	.40	.30	.52	.45	.49	.32	.53	.72	.70	.60	.70	.73	.25	.52	.51	.49	.62	.61	.23	.24	.46	.55	.48	.84	.83	.75	.51	.72	.76	.58	
Media	10.2	10.2	10.0	10.1	10.2	9.9	10.0	10.4	10.0	10.4	10.0	10.3	10.3	10.0	10.1	10.2	10.2	10.0	10.3	10.1	10.2	10.0	101.0	100.1	101.4	101.3	100.8	101.8	101.0	102.1	101.1	101.3	101.3
Dt	2.8	2.9	2.8	2.9	3.0	3.2	3.0	3.2	2.9	2.9	2.9	2.7	2.7	2.7	2.9	3.1	3.3	2.7	2.9	3.0	2.9	2.8	13.9	15.6	15.1	14.1	13.0	14.0	14.3	14.1	14.7	14.0	13.9

Tabla A.3 Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas para las edades 10:0-11:11

Prueba/ punt. de procesamiento/ punt. compuesta	S	V	I	CO	C	PV	M	B	A	D	SD	LN	CL	BS	CA	Cs	Cp	Dd	Di	Dc	CAa	CAe	ICV	IVE	IRF	IMT	IVP	CIT	IRC	IMTA	INV	ICG	ICC
Semejanzas																							.57					.60				.61	
Vocabulario	.57																						.57					.58				.57	
Información	.62	.59																															
Comprensión	.51	.48	.37																														
Cubos	.43	.36	.37	.34																				.64				.55			.60	.56	
Puzles visuales	.48	.40	.51	.46	.64																			.64							.67		
Matrices	.40	.36	.39	.40	.44	.50																			.49			.57			.58	.55	
Balanzas	.44	.45	.44	.36	.49	.52	.49																		.49			.59	.58		.54	.62	
Aritmética	.52	.46	.55	.37	.44	.50	.49	.58																					.58				
Dígitos	.45	.43	.45	.33	.37	.38	.49	.46	.55																	.44		.61		.72			.40
Span de dibujos	.29	.30	.32	.36	.31	.33	.40	.31	.31	.44																.44					.47		.46
Letras y números	.45	.47	.46	.50	.41	.49	.47	.48	.54	.72	.52																			.72			
Claves	.13	.16	.12	.31	.15	.23	.13	.04	.19	.25	.30	.27															.56	.20			.22		.50
Búsqueda de símbolos	.24	.15	.22	.36	.19	.33	.19	.24	.37	.24	.31	.34	.56														.56						.49
Cancelación	.23	.16	.18	.18	.18	.20	.18	.15	.23	.18	.15	.19	.50	.42																			
Cs	.45	.39	.40	.34	.98	.63	.46	.52	.45	.37	.32	.41	.15	.19	.19																		
Cp	.33	.30	.25	.30	.97	.56	.38	.48	.32	.43	.29	.32	.08	.10	.11	.95																	
Dd	.36	.42	.41	.28	.25	.28	.42	.31	.36	.77	.40	.61	.23	.18	.10	.26	.23																
Di	.36	.31	.36	.28	.32	.31	.38	.35	.46	.80	.41	.63	.19	.21	.16	.33	.32	.50															
Dc	.41	.35	.36	.31	.35	.37	.41	.46	.52	.83	.33	.58	.24	.25	.20	.35	.41	.46	.49														
CAa	.23	.17	.20	.15	.12	.16	.17	.16	.22	.17	.10	.18	.46	.39	.94	.14	.09	.11	.14	.18													
CAe	.22	.14	.15	.21	.23	.22	.18	.14	.23	.18	.17	.21	.48	.41	.93	.22	.13	.11	.16	.20	.76												
ICV	.88	.89	.69	.56	.45	.50	.43	.50	.55	.50	.33	.53	.17	.22	.22	.47	.36	.44	.38	.43	.23	.20											
IVE	.50	.42	.49	.44	.90	.91	.52	.55	.52	.42	.35	.49	.21	.29	.22	.89	.85	.30	.35	.39	.16	.25	.52										
IRF	.49	.48	.47	.44	.54	.59	.84	.88	.62	.55	.40	.55	.09	.25	.19	.57	.51	.42	.42	.50	.19	.19	.54	.62									
IMT	.43	.43	.46	.41	.40	.42	.53	.46	.51	.85	.84	.73	.32	.32	.19	.41	.39	.69	.71	.69	.16	.21	.49	.45	.56								
IVP	.21	.18	.19	.38	.19	.32	.18	.16	.32	.28	.35	.35	.88	.88	.51	.20	.11	.23	.23	.28	.48	.50	.22	.29	.20	.37							
CIT	.73	.72	.63	.58	.69	.67	.70	.73	.69	.73	.49	.69	.38	.38	.33	.71	.66	.58	.57	.65	.31	.33	.82	.75	.83	.72	.43						
IRC	.53	.51	.55	.41	.52	.57	.55	.90	.88	.57	.35	.57	.13	.34	.21	.55	.46	.38	.46	.54	.22	.21	.59	.61	.85	.54	.27	.80					
IMTA	.49	.48	.50	.45	.42	.47	.52	.51	.59	.93	.52	.92	.29	.31	.20	.43	.40	.75	.78	.76	.19	.21	.55	.49	.59	.86	.35	.77	.61				
INV	.53	.50	.53	.54	.74	.80	.73	.72	.62	.58	.64	.64	.44	.44	.32	.76	.71	.46	.47	.52	.28	.34	.58	.85	.83	.72	.50	.90	.75	.66			
ICG	.76	.74	.65	.55	.72	.67	.71	.78	.67	.58	.42	.60	.15	.27	.23	.75	.67	.47	.46	.53	.22	.24	.85	.77	.87	.59	.24	.96	.81	.64	.85		
ICC	.38	.36	.39	.48	.35	.44	.42	.37	.50	.67	.71	.65	.74	.73	.43	.36	.30	.55	.56	.57	.39	.43	.42	.44	.45	.82	.83	.69	.48	.72	.73	.49	
Media	10.2	9.9	10.0	10.0	10.1	10.2	10.0	10.2	10.0	10.2	10.2	10.1	10.0	10.0	10.2	10.0	10.2	10.2	10.2	10.0	10.2	10.3	100.2	100.9	100.6	101.2	100.0	100.7	100.5	100.9	100.9	100.6	100.8
Dt	3.1	3.3	3.0	2.7	3.0	3.1	3.0	3.3	3.1	3.1	3.0	2.7	3.0	3.0	3.2	3.0	3.5	2.8	2.9	3.3	3.1	3.3	15.5	15.6	15.7	14.8	15.1	14.8	16.6	15.1	15.0	15.2	14.8

Tabla A.4 Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas para las edades 12:0-13:11

Prueba/ punt. de procesamiento/ punt. compuesta	S	V	I	CO	C	PV	M	B	A	D	SD	LN	CL	BS	CA	Cs	Cp	Dd	Di	Dc	CAa	CAe	ICV	IVE	IRF	IMT	IVP	CIT	IRC	IMTA	INV	ICG	ICC
Semejanzas																							.68					.67				.66	
Vocabulario	.68																						.68					.61				.63	
Información	.64	.65																															
Comprensión	.55	.55	.48																														
Cubos	.43	.37	.46	.33																				.58	.			.56			.59	.54	
Puzles visuales	.37	.42	.39	.30	.58																			.58							.60		
Matrices	.45	.38	.42	.35	.43	.43																			.44			.58			.57	.55	
Balanzas	.40	.46	.43	.33	.46	.52	.44																		.44			.59	.50		.57	.57	
Aritmética	.50	.56	.61	.46	.54	.53	.45	.50																					.50				
Dígitos	.42	.38	.47	.31	.35	.41	.43	.49	.52																	.52		.55		.70			.47
Span de dibujos	.35	.36	.41	.36	.26	.31	.41	.35	.39	.52																.52					.47		.52
Letras y números	.45	.41	.48	.43	.29	.29	.38	.42	.57	.70	.53																			.70			
Claves	.36	.29	.29	.34	.37	.30	.31	.25	.37	.28	.35	.21															.62	.43			.44		.54
Búsqueda de símbolos	.34	.30	.32	.35	.34	.24	.28	.28	.39	.29	.33	.26	.62														.62						.54
Cancelación	.24	.27	.18	.28	.21	.20	.21	.13	.30	.17	.26	.15	.48	.34																			
Cs	.44	.36	.48	.34	.96	.55	.43	.42	.51	.34	.23	.28	.34	.31	.19																		
Cp	.40	.32	.44	.26	.94	.44	.46	.44	.48	.32	.24	.22	.31	.26	.09	.93																	
Dd	.36	.34	.38	.21	.28	.25	.34	.36	.40	.82	.43	.55	.20	.20	.11	.27	.27																
Di	.39	.28	.35	.25	.31	.35	.43	.41	.41	.82	.42	.55	.21	.20	.10	.33	.34	.57															
Dc	.32	.32	.44	.33	.35	.43	.33	.45	.52	.83	.47	.62	.31	.34	.21	.32	.30	.53	.52														
CAa	.18	.21	.11	.23	.18	.13	.17	.10	.27	.11	.24	.11	.46	.34	.92	.15	.06	.06	.04	.18													
CAe	.29	.31	.25	.34	.24	.26	.26	.16	.31	.23	.26	.19	.45	.30	.91	.24	.16	.18	.17	.23	.70												
ICV	.91	.92	.70	.60	.44	.43	.45	.47	.58	.44	.39	.47	.36	.35	.28	.44	.39	.38	.37	.35	.21	.32											
IVE	.45	.45	.48	.36	.88	.90	.49	.55	.60	.43	.32	.34	.38	.33	.23	.84	.78	.30	.37	.45	.18	.28	.49										
IRF	.50	.49	.50	.39	.52	.56	.84	.85	.56	.54	.44	.47	.33	.33	.20	.50	.53	.41	.49	.46	.15	.24	.54	.61									
IMT	.44	.42	.50	.38	.34	.40	.48	.47	.52	.86	.88	.69	.37	.36	.25	.32	.31	.71	.70	.73	.21	.28	.47	.42	.56								
IVP	.39	.33	.34	.39	.40	.31	.33	.30	.42	.32	.38	.26	.91	.89	.46	.36	.32	.23	.22	.36	.45	.42	.39	.40	.37	.41							
CIT	.78	.74	.69	.57	.69	.61	.70	.71	.70	.68	.53	.58	.59	.50	.35	.67	.64	.55	.58	.59	.29	.40	.83	.73	.82	.69	.61						
IRC	.52	.59	.61	.46	.57	.60	.51	.85	.88	.58	.43	.58	.37	.39	.25	.54	.52	.44	.47	.56	.21	.27	.61	.67	.80	.57	.42	.81					
IMTA	.48	.44	.52	.40	.36	.39	.45	.50	.60	.92	.58	.92	.27	.30	.17	.35	.29	.75	.75	.79	.12	.23	.50	.43	.55	.85	.32	.69	.63				
INV	.56	.54	.57	.47	.73	.75	.72	.71	.65	.59	.65	.50	.63	.51	.35	.69	.67	.43	.50	.56	.30	.38	.60	.83	.84	.71	.63	.91	.79	.60			
ICG	.80	.79	.70	.57	.70	.61	.71	.73	.67	.54	.45	.51	.42	.41	.28	.69	.67	.43	.47	.46	.22	.33	.87	.74	.84	.56	.46	.96	.81	.57	.86		
ICC	.49	.44	.50	.45	.43	.41	.48	.45	.56	.71	.76	.58	.75	.75	.41	.40	.37	.56	.56	.66	.38	.41	.51	.48	.54	.84	.83	.77	.59	.70	.79	.60	
Media	9.9	10.1	10.0	10.1	9.8	10.1	9.9	9.7	10.0	10.0	9.8	9.7	9.9	9.9	9.8	9.7	9.6	9.9	10.1	9.9	9.8	9.8	99.9	99.5	98.7	99.2	99.4	99.1	99.2	99.0	98.8	99.1	99.1
Dt	3.1	3.2	3.1	3.1	2.7	2.9	2.8	2.8	3.1	3.0	3.2	2.9	3.0	2.9	3.0	2.7	2.9	3.0	2.9	3.0	3.2	3.1	16.0	13.9	13.8	15.7	15.1	14.5	14.8	15.3	14.5	14.2	15.5

Tabla A.5 Intercorrelaciones de las puntuaciones de las pruebas, de las puntuaciones de procesamiento y de las puntuaciones compuestas para las edades 14:0-16:11

Prueba/ punt. de procesamiento/ punt. compuesta	S	V	I	CO	C	PV	M	B	A	D	SD	LN	CL	BS	CA	Cs	Cp	Dd	Di	Dc	CAa	CAe	ICV	IVE	IRF	IMT	IVP	CIT	IRC	IMTA	INV	ICG	ICC
Semejanzas																							.54					.65				.65	
Vocabulario	.54																						.54					.57				.58	
Información	.56	.63																															
Comprensión	.50	.51	.48																														
Cubos	.46	.40	.41	.36																				.58				.57			.59	.53	
Puzles visuales	.46	.43	.48	.39	.58																			.58							.65		
Matrices	.50	.46	.43	.44	.40	.49																			.56			.63			.59	.62	
Balanzas	.49	.42	.46	.40	.42	.56	.56																		.56			.59	.53		.60	.61	
Aritmética	.50	.45	.52	.42	.40	.43	.53	.53																					.53				
Dígitos	.48	.38	.42	.32	.41	.35	.48	.41	.57																	.52		.57		.65			.43
Span de dibujos	.34	.28	.31	.39	.36	.38	.37	.36	.43	.52																.52					.47		.43
Letras y números	.51	.43	.44	.37	.35	.40	.46	.42	.56	.65	.53																			.65			
Claves	.21	.20	.21	.24	.28	.21	.21	.16	.17	.23	.20	.27															.51	.29			.28		.42
Búsqueda de símbolos	.17	.24	.23	.29	.28	.21	.14	.13	.19	.22	.25	.29	.51														.51						.45
Cancelación	.27	.24	.24	.25	.22	.24	.23	.17	.20	.17	.21	.19	.40	.41																			
Cs	.43	.38	.37	.34	.95	.54	.37	.41	.37	.39	.32	.32	.27	.29	.19																		
Cp	.46	.45	.42	.41	.94	.55	.46	.44	.50	.42	.38	.42	.24	.25	.20	.91																	
Dd	.35	.28	.30	.24	.26	.22	.37	.29	.38	.78	.44	.55	.11	.13	.16	.24	.26																
Di	.39	.33	.37	.26	.34	.31	.39	.34	.52	.83	.40	.53	.22	.21	.12	.33	.34	.51															
Dc	.44	.31	.37	.28	.40	.34	.44	.41	.49	.82	.44	.58	.22	.19	.16	.36	.45	.50	.53														
CAa	.19	.21	.16	.21	.19	.21	.20	.16	.15	.12	.18	.13	.36	.36	.94	.17	.19	.12	.07	.12													
CAe	.33	.25	.31	.26	.23	.23	.25	.14	.25	.23	.22	.24	.39	.40	.91	.20	.21	.22	.16	.19	.74												
ICV	.87	.88	.68	.58	.49	.51	.55	.52	.54	.48	.36	.54	.23	.24	.29	.46	.51	.36	.41	.43	.22	.32											
IVE	.52	.47	.50	.42	.89	.89	.51	.55	.47	.43	.42	.42	.27	.28	.26	.84	.84	.27	.37	.42	.23	.26	.56										
IRF	.55	.49	.50	.47	.46	.60	.89	.88	.59	.50	.41	.49	.21	.14	.22	.44	.49	.37	.41	.48	.19	.22	.60	.59									
IMT	.46	.37	.41	.40	.44	.42	.48	.44	.57	.86	.88	.67	.24	.26	.22	.41	.44	.70	.70	.72	.17	.26	.47	.48	.52								
IVP	.21	.26	.25	.31	.32	.24	.20	.16	.21	.25	.26	.32	.87	.87	.47	.32	.29	.14	.24	.23	.41	.45	.27	.32	.20	.29							
CIT	.76	.70	.64	.57	.70	.64	.75	.72	.65	.70	.50	.64	.48	.35	.35	.67	.68	.50	.59	.63	.29	.37	.83	.76	.83	.68	.47						
IRC	.57	.49	.56	.46	.47	.57	.62	.88	.87	.56	.45	.55	.19	.18	.21	.44	.52	.38	.49	.51	.17	.22	.60	.58	.85	.57	.21	.78					
IMTA	.55	.45	.48	.39	.42	.41	.52	.46	.62	.91	.58	.91	.28	.28	.21	.39	.47	.74	.74	.77	.14	.27	.57	.47	.55	.84	.32	.74	.61				
INV	.59	.52	.55	.53	.73	.78	.74	.74	.60	.58	.64	.58	.50	.36	.35	.69	.69	.40	.48	.54	.31	.35	.63	.85	.83	.70	.49	.92	.77	.64			
ICG	.78	.74	.65	.58	.71	.67	.78	.77	.63	.56	.45	.57	.28	.25	.29	.67	.69	.40	.47	.53	.24	.31	.87	.77	.87	.57	.30	.96	.80	.63	.88		
ICC	.42	.38	.41	.43	.47	.40	.42	.37	.47	.69	.69	.61	.70	.72	.43	.45	.44	.50	.58	.59	.37	.44	.46	.49	.44	.79	.81	.72	.48	.72	.74	.54	
Media	10.0	9.7	10.1	9.9	9.9	9.9	10.0	9.9	9.9	9.8	10.0	10.0	10.1	10.0	10.0	9.9	9.9	9.6	9.9	10.1	10.0	10.1	99.3	99.3	99.7	99.5	100.3	99.4	99.5	99.6	99.6	99.4	100.0
Dt	2.8	3.0	3.2	3.0	3.0	2.9	3.0	2.9	2.7	2.9	3.0	2.9	3.1	3.1	3.0	3.1	2.6	2.9	3.0	2.8	3.1	3.0	13.8	14.7	15.3	14.8	15.2	14.4	14.2	14.6	14.8	14.5	14.6

Criterios de inclusión para los grupos especiales

Criterios de inclusión generales para los grupos especiales (muestras española y estadounidense)

Los niños fueron seleccionados para formar parte de la muestra de los grupos especiales si cumplían los siguientes criterios:

- español como lengua materna (muestra española);
- inglés como lengua materna (muestra estadounidense);
- capacidad para comunicar a un nivel acorde con la edad y el diagnóstico, y cierta capacidad para comunicar oralmente;
- audición y visión normales (corregidas);
- motricidad gruesa y fina normal (a excepción del déficit motriz leve que se da en los grupos de niños con discapacidad intelectual);
- ausencia de enfermedad o condición física que pudiera afectar al funcionamiento cognitivo o al rendimiento en las pruebas (a excepción de los trastornos asociados al grupo especial al que pertenece el niño);
- no haber sido diagnosticado con una afección neurológica (p. ej., trastorno convulsivo, epilepsia, encefalitis, cirugía cerebral, tumor cerebral) distinta a la pertinente al grupo especial correspondiente;
- no haber presentado un periodo de inconsciencia no relacionado con una intervención quirúrgica o superior a 20 minutos relacionado con una enfermedad;
- no haber sido diagnosticado con discapacidad intelectual (a excepción del grupo con discapacidad intelectual) ni con trastorno generalizado del desarrollo (a excepción de los grupos de niños con trastorno del espectro autista);
- no haber sido diagnosticado con un trastorno psiquiátrico (p. ej., trastornos psicóticos, trastornos del estado de ánimo) distinto a los establecidos como criterios de los grupos especiales;
- no estar internado en un hospital o en una institución mental o psiquiátrica, ni seguir ningún tratamiento (a excepción de los ámbitos en que se siguen tratamientos específicos para una afección determinada [p. ej., centros para los sujetos con discapacidad intelectual o trastorno del espectro autista]);
- no estar tomando en ese momento ninguna medicación que pudiera afectar al rendimiento del niño en la escala, a excepción de la adecuada a la afección que define al grupo especial, o a otras afecciones asociadas;
- cumplir los criterios correspondientes exclusivamente a uno de los grupos especiales (a excepción de los diagnósticos duales de trastorno por déficit de atención/hiperactividad o comportamiento disruptivo secundarios a la discapacidad intelectual); y

- no haber sido evaluado con la WISC-IV u otra medida de la aptitud cognitiva durante los últimos 6 meses antes de la fecha de aplicación de la escala.

Criterios de inclusión específicos para los grupos especiales (muestras española y estadounidense)

Alta capacidad intelectual

Criterios:

- tener entre 6 y 16 años;
- puntuar ≥ 2 desviaciones típicas por encima de la media en la Escala total de pruebas estandarizadas de aptitud intelectual aplicadas individualmente (p. ej., CI ≥ 130); y
- recibir atención especial para la alta capacidad intelectual.

Discapacidad intelectual

Criterios:

- tener entre 6 y 16 años;

Y

- cumplir los criterios del DSM-5 para el diagnóstico de discapacidad intelectual leve o moderada;

O

- puntuar entre 2 y 4 desviaciones típicas por debajo de la media en la Escala total de pruebas estandarizadas de aptitud intelectual aplicadas individualmente (p. ej., CI = 40-70).

Criterios de inclusión específicos para los grupos especiales (muestra estadounidense)

Capacidad intelectual límite

Criterios:

- tener entre 6 y 16 años;

Y

- cumplir los criterios del DSM-5 para la capacidad intelectual límite;

O

- puntuar entre 1 y 2 desviaciones típicas por debajo de la media (p. ej., CI de 71-84) en pruebas estandarizadas de aptitud intelectual aplicadas individualmente.

Trastorno específico del aprendizaje-lectura

Criterios:

- tener entre 6 y 16 años;
- cumplir los criterios del DSM-5 para el trastorno específico del aprendizaje con dificultades en la lectura;

Y

- mostrar un patrón de puntos fuertes y débiles en el rendimiento y en el procesamiento cognitivo, que sea consistente con el trastorno específico del aprendizaje-lectura;

O

- presentar una diferencia de ≥ 15 puntos entre la aptitud y el rendimiento lector en pruebas estandarizadas de aptitud cognitiva y rendimiento lector;

O

- algún documento en el que la escuela indique que el niño está recibiendo atención para el trastorno específico del aprendizaje-lectura.

Trastorno específico del aprendizaje-lectura y expresión escrita

Criterios:

- tener entre 6 y 16 años;
- cumplir los criterios del DSM-5 para el trastorno específico del aprendizaje con dificultades en la lectura y en la expresión escrita;

Y

- mostrar un patrón de puntos fuertes y débiles en el rendimiento y en el procesamiento cognitivo, que sea consistente con el trastorno específico del aprendizaje-lectura y expresión escrita;

O

- presentar una diferencia de ≥ 15 puntos entre la aptitud y el rendimiento lector y la aptitud y el rendimiento en expresión escrita en pruebas estandarizadas de aptitud cognitiva y rendimiento;

O

- algún documento en el que la escuela indique que el niño está recibiendo atención para el trastorno específico del aprendizaje-lectura y expresión escrita.

Trastorno específico del aprendizaje-matemáticas

Criterios:

- tener entre 6 y 16 años;
- cumplir los criterios del DSM-5 para el trastorno específico del aprendizaje con dificultad matemática;

Y

- mostrar un patrón de puntos fuertes y débiles en el rendimiento y en el procesamiento cognitivo, que sea consistente con el trastorno específico del aprendizaje-matemáticas;

O

- presentar una diferencia de ≥ 15 puntos entre la aptitud y el rendimiento en matemáticas en pruebas estandarizadas de aptitud y rendimiento en matemáticas;

O

- algún documento en el que la escuela indique que el niño está recibiendo atención para el trastorno específico del aprendizaje-matemáticas.

Trastorno por déficit de atención/hiperactividad

Criterios:

- tener entre 6 y 16 años;

- puntuar ≥ 80 en la Escala total de pruebas estandarizadas de aptitud intelectual aplicadas individualmente, o disponer de un informe psicológico en el que se estima un CI ≥ 80;
- cumplir los criterios del DSM-5 para el diagnóstico del trastorno por déficit de atención/hiperactividad;
- no haber tomado medicación psicoestimulante durante, como mínimo, 24 horas antes de la aplicación de la escala.

Comportamiento disruptivo

Criterios:

- tener entre 6 y 16 años;
- puntuar ≥ 70 en la escala Problemas de conducta, en su forma para los padres, de la *Behavior Assessment System for Children-Second Edition*;

O

- cumplir los criterios del DSM-5 para el diagnóstico del trastorno negativista desafiante o del trastorno de la conducta.

Traumatismo craneoencefálico

Criterios:

- tener entre 6 y 16 años;
- haber sido diagnosticado con un traumatismo craneoencefálico moderado o grave;
- disponer de un CI premórbido o premórbido estimado > 80:
- haber sufrido el traumatismo craneoencefálico entre 6 y 18 meses antes de la aplicación de la escala;
- no haber sido diagnosticado con una afección neurológica (p. ej., trastorno convulsivo, epilepsia o cirugía cerebral) antes de sufrir el traumatismo craneoencefálico; y
- la lesión no se debe a un tumor cerebral u otra enfermedad o afección (p. ej., infección bacteriana, meningitis, encefalitis).

Trastorno del espectro autista con deterioro del lenguaje acompañante

Criterios:

- tener entre 6 y 16 años;
- cumplir los criterios del DSM-5 para el diagnóstico del trastorno del espectro autista con deterioro del lenguaje acompañante;
- contar con destrezas lingüísticas adecuadas para realizar el test; y
- puntuar ≥ 60 en una medida estandarizada de aptitud intelectual aplicada individualmente, o disponer de un informe psicológico en el que se estima un CI ≥ 60.

Trastorno del espectro autista sin deterioro del lenguaje acompañante

Criterios:

- tener entre 6 y 16 años;
- cumplir los criterios del DSM-5 para el diagnóstico del trastorno del espectro autista sin deterioro del lenguaje acompañante; y
- puntuar ≥ 70 en una medida estandarizada de aptitud intelectual aplicada individualmente, o disponer de un informe psicológico en el que se estima un CI ≥ 70.

Anexo C

Equipo de expertos, revisores y consultores de la versión estadounidense y española

Equipo de expertos

Jacques Grégoire, PhD, Université Catholique de Louvain, Institut for Research in Psychology (IPSY), Louvain-la-Neuve, Bélgica

Antolin M. Llorente, PhD, ABPdN, Department of Psychiatry, College of Medicine, Penn State Hershey and Milton S. Hershey Medical Center, Hershey, Pennsylvania

Donald H. Saklofske, PhD, Department of Psychology, University of Western Ontario, London, Ontario, Canadá

Sally E. Shaywitz, MD, Audrey G. Ratner Professor in Learning Development; Co-Director, Yale Center for Dyslexia & Creativity (YCDC), Yale University School of Medicine, New Haven, Connecticut

Revisores y consultores

Lorraine Bridges, MA
Nelson Cowan, PhD
Dean Delis, PhD
Dennis Engi, PhD
Craig Frisby, PhD
Susan Gathercole, PhD
Stephen Hardy-Braz, PsyS
Josette Harris, PhD
Alan S. Kaufman, PhD
Nadeen L. Kaufman, PhD
Elizabeth Lichtenberger, PhD
James Livingston, PhD
Michelle Lurie, PsyD
Jennifer Mascolo, PsyD
Maria R. Muñoz
Nora Newcombe, PhD
Beth Rewalt, MS
Jim Rewalt, MS
Gustavo Rey, PhD
Matthew R. Reynolds, PhD
Marilee Ruebsamen, PhD
Joseph J. Ryan, PhD
Timothy Salthouse, PhD
Eric Saperstein
W. Joel Schneider, PhD
Linda Silverman, PhD
Mireille Simon
Lisa Suzuki, PhD
Marsha Vasserman, PsyD

Bibliografía

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. doi: 10.1109/TAC.1974.1100705

Akbar, M., Loomis, R., y Paul, R. (2013). The interplay of language on executive functions in children with ASD. *Research in Autism Spectrum Disorders, 7*, 494–501. doi: 10.1016/j.rasd.2012.09.001

Allen, D. N., Leany, B. D., Thaler, N. S., Cross, C., Sutton, G. P., y Mayfield, J. (2010). Memory and attention profiles in pediatric traumatic brain injury. *Archives of Clinical Neuropsychology, 25*(7), 618–633. doi: 10.1093/arclin/acq051

Allen, M. J., y Yen, W. M. (2002). *Introduction to measurement theory*. Long Grove, IL: Waveland Press.

American Association on Intellectual and Developmental Disabilities. (2010). *Intellectual disability: Definition, classification, and systems of supports* (11th ed.). Washington, DC: Author.

American Educational Research Association. (2006). Standards for reporting on empirical social science research in AERA publications. *Educational Researcher, 35*(6), 33–40.

American Educational Research Association, American Psychological Association, y National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5ª ed.). Arlington, VA: American Psychiatric Association.

American Psychiatric Association. (2014). *Manual diagnóstico y estadístico de los trastornos mentales* (5ª ed.). Madrid: Editorial Médica Panamericana.

Anastasi, A., y Urbina, S. (1997). *Psychological testing* (7ª ed.). Upper Saddle River, NJ: Prentice Hall.

Andersen, S. L. (2003). Trajectories of brain development: Point of vulnerability or window of opportunity? *Neuroscience and Biobehavioral Reviews, 27*(1–2), 3–18. doi: 10.1016/S0149-7634(03)00005-8

Anderson, V., Catroppa, C., Godfrey, C., y Rosenfeld, J. V. (2012). Intellectual ability 10 years after traumatic brain injury in infancy and childhood: What predicts outcome? *Journal of Neurotrauma, 29*(1), 143–153. doi: 10.1089/neu.2011.2012

Anderson, V., Godfrey, C., Rosenfeld, J. V., y Catroppa, C. (2012). Predictors of cognitive function and recovery 10 years after traumatic brain injury in young children. *Pediatrics, 129*(2), e254–e261. doi: 10.1542/peds.2011-0311

Archibald, L. M. D., y Gathercole, S. E. (2006). Short-term memory and working memory in specific language impairment. En T. P. Alloway y S. E. Gathercole (Eds.), *Working memory and neurodevelopmental disorders* (pp. 139–160). Nueva York, NY: Psychology Press.

Armour-Thomas, E., y Gopaul-McNicol, S.-A. (1997). Bio-ecological approach to cognitive assessment. *Cultural Diversity and Mental Health, 3*(2), 131–144.

Aukema, E. J., Caan, M. W. A., Oudhuis, N., Majoie, C. B. L. M., Vos, F. M., Reneman, L.,...Schouten-van Meeteren, A. Y. N. (2009). White matter fractional anisotropy correlates with speed of processing and motor speed in young childhood cancer survivors. *International Journal of Radiation Oncology • Biology • Physics, 74*(3), 837–843. doi: 10.1016/j.ijrobp.2008.08.060

Babikian, T., Satz, P., Zaucha, K., Light, R., Lewis, R. S., y Asarnow, R. F. (2011). The UCLA longitudinal study of neurocognitive outcomes following mild pediatric traumatic brain injury. *Journal of the International Neuropsychological Society, 17*(5), 886–895. doi: 10.1017/S1355617711000907

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423.

Baddeley, A. D. (2002). Is working memory still working? *European Psychologist, 7*(2), 85–97. doi: 10.1027//1016-9040.7.2.85

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63,* 1–29. doi: 10.1146/annurev-psych-120710-100422

Balluerka, N., Gómez, J., y Hidalgo, D. (2005). The controversy over null hypothesis significance testing revisited. *Methodology, 1*(2), 55–70.

Barbeau, E. B., Soulières, I., Dawson, M., Zeffiro, T. A., y Mottron, L. (2013). The level and nature of autistic intelligence III: Inspection time. *Journal of Abnormal Psychology, 122*(1), 295–301. doi: 10.1037/a0029984

Barbey, A. K., Colom, R., Solomon, J., Krueger, F., Forbes, C., y Grafman, J. (2012). An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain, 135*(4), 1154–1164. doi: 10.1093/brain/aws021

Barker, E. D., Tremblay, R. E., van Lier, P. A. C., Vitaro, F., Nagin, D. S., Assaad, J.-M., y Séguin, J. R. (2011). The neurocognition of conduct disorder behaviors: Specificity to physical aggression and theft after controlling for ADHD symptoms. *Aggressive Behavior, 37*(1), 63–72. doi: 10.1002/ab.20373

Barkley, R. A. (2007). School interventions for attention deficit hyperactivity disorder: Where to from here? *School Psychology Review, 36*(2), 279–286.

Baxendale, S., McGrath, K., y Thompson, P. J. (2014). Epilepsy & IQ: The clinical utility of the Wechsler Adult Intelligence Scale–Fourth Edition (WAIS–IV) indices in the neuropsychological assessment of people with epilepsy. *Journal of Clinical and Experimental Neuropsychology, 36*(2), 137–143. doi: 10.1080/13803395.2013.870535

Ben-David, B. M., Nguyen, L. L. T., y van Lieshout, P. H. H. M. (2011). Stroop effects in persons with traumatic brain injury: Selective attention, speed of processing, or color-naming? A meta-analysis. *Journal of the International Neuropsychological Society, 17,* 354–363. doi: 10.1017/S135561771000175X

Benson, N., Hulac, D. M., y Bernstein, J. D. (2013). An independent confirmatory factor analysis of the Wechsler Intelligence Scale for Children–Fourth Edition (WISC–IV) Integrated: What do the process approach subtests measure? *Psychological Assessment, 25*(3), 692–705. doi: 10.1037/a0032298

Benson, N., Hulac, D. M., y Kranzler, J. H. (2010). Independent examination of the Wechsler Adult Intelligence Scale–Fourth Edition (WAIS–IV): What does the WAIS–IV measure? *Psychological Assessment, 22*(1), 121–130. doi: 10.1037/a0017767

Benson, N., y Taub, G. E. (2013). Invariance of Woodcock-Johnson III scores for students with learning disorders and students without learning disorders. *School Psychology Quarterly, 28*(3), 256–272. doi: 10.1037/spq0000028

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246.

Bentler, P. M., y Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–606.

Beres, K. A., Kaufman, A. S., y Perlman, M. D. (2000). Assessment of child intelligence. En G. Goldstein y M. Hersen (Eds.), *Handbook of psychological assessment* (3ª ed., pp. 65–96). Kidlington, Oxford, United Kingdom: Elsevier Science.

Berl, M. M., Duke, E. S., Mayo, J., Rosenberger, L. R., Moore, E. N., VanMeter, J.,...Gaillard, W. D. (2010). Functional anatomy of listening and reading comprehension during development. *Brain and Language, 114*(2), 115–125. doi: 10.1016/j.bandl.2010.06.002

Berninger, V., y Richards, T. (2010). Inter-relationships among behavioral markers, genes, brain and treatment in dyslexia and dysgraphia. *Future Neurology, 5*(4), 597–617. doi: 10.2217/fnl.10.22

Betjemann, R. S., Johnson, E. P., Barnard, H., Boada, R., Filley, C. M., Filipek, P. A.,...Pennington, B. F. (2010). Genetic covariation between brain volumes and IQ, reading performance, and processing speed. *Behavior Genetics, 40*(2), 135–145. doi: 10.1007/s10519-009-9328-2

Binet, A., y Simon, T. (1905). Methodes nouvelles pour le diagnostic du niveau intellectual des anormaux [A new method for the diagnosis of intellectual level of abnormal persons]. *L'nnee Psychologique, 11,* 191–244.

Blalock, L. D., y McCabe, D. P. (2011). Proactive interference and practice effects in visuospatial working memory span task performance. *Memory, 19*(1), 83–91. doi: 10.1080/09658211.2010.537035

Bodin, D., Pardini, D. A., Burns, T. G., y Stevens, A. B. (2009). Higher order factor structure of the WISC–IV in a clinical neuropsychological sample. *Child Neuropsychology, 15,* 417–424. doi: 10.1080/09297040802603661

Bonifacci, P., y Snowling, M. J. (2008). Speed of processing and reading disability: A cross-linguistic investigation of dyslexia and borderline intellectual functioning. *Cognition, 107*, 999–1017. doi: 10.1016/j.cognition.2007.12.006

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience, 25*(1), 151–188. doi: 10.1146/annurev.neuro.25.112701.142946

Booth, J. R., Bebko, G., Burman, D. D., y Bitan, T. (2007). Children with reading disorder show modality independent brain-abnormalities during semantic tasks. *Neuropsychologia, 45*(4), 775–783. doi: 10.1016/j.neuropsychologia.2006.08.015

Borella, E., Carretti, B., y Pelegrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities, 43*(6), 541–552. doi: 10.1177/0022219410371676

Boucher, J., y Mayes, A. (2012). Memory in ASD: Have we been barking up the wrong tree? *Autism, 16*(6), 603–611. doi: 10.1177/1362361311417738

Boucher, J., Mayes, A., y Bigham, S. (2012). Memory in autistic spectrum disorder. *Psychological Bulletin, 138*(3), 458–496. doi: 10.1037/a0026869

Bowden, S. C., Weiss, L. G., Holdnack, J. A., y Lloyd, D. (2006). Age-related invariance of abilities measured with the Wechsler Adult Intelligence Scale–III. *Psychological Assessment, 18*(3), 334–339.

Boylan, K. (2014). The many faces of oppositional defiant disorder. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 23*(1), 8–9.

Braden, J. P., y Niebling, B. C. (2012). Using the joint test standards to evaluate the validity evidence for intelligence tests. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed., pp. 739–757). Nueva York, NY: Guilford Press.

Braver, T. S., Barch, D. M., Kelley, W. M., Buckner, R. L., Cohen, N. J., Miezin, F. M.,...Petersen, S. E. (2001). Direct comparison of prefrontal cortex regions engaged by working and long-term memory tasks. *NeuroImage, 14*, 48–59. doi: 10.1006/nimg.2001.0791

Bremner, D., McTaggart, B., Saklofske, D. H., y Janzen, T. (2011). WISC–IV GAI and CPI in psychoeducational assessment. *Canadian Journal of School Psychology, 26*(3), 209–219. doi: 10.1177/0829573511419090

Brown, T. E. (2001). *Brown attention-deficit disorder scales for children and adolescents.* San Antonio, TX: The Psychological Corporation.

Browne, M. W., y Cudeck, R. (1993). Alternative ways of assessing model fit. En K. A. Bollen y J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Bunge, S. A., Wendelken, C., Badre, D., y Wagner, A. D. (2005). Analogical reasoning and prefrontal cortex: Evidence for separable retrieval and integration mechanisms. *Cerebral Cortex, 15*, 239–249. doi: 10.1093/cercor/bhh126

Bunge, S. A., y Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Current Opinion in Neurobiology, 17*(2), 243–250. doi: 10.1016/j.conb.2007.02.005

Bunting, M. (2006). Proactive interference and item similarity in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(2), 183–196. doi: 10.1037/0278-7393.32.2.183

Burgaleta, M., MacDonald, P. A., Martínez, K., Román, F. J., Álvarez-Linera, J., González, A. R.,... Colom, R. (2013). Subcortical regional morphology correlates with fluid and spatial intelligence. *Human Brain Mapping. 35*(5), 1957–1968. doi 10.1002/hbm.22305

Burgess, G. C., y Braver, T. S. (2010). Neural mechanisms of interference control in working memory: Effects of interference expectancy and fluid intelligence. *PLoS ONE, 5*(9), e12861. doi: 10.1371/journal.pone.0012861

Burgess, G. C., Gray, J. R., Conway, A. R. A., y Braver, T. S. (2011). Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory span. *Journal of Experimental Psychology: General, 140*(4), 674–692. doi: 10.1037/a0024695

Byrne, B. M. (2001). Structural equation modeling: Perspectives on the present and the future. *International Journal of Testing, 1*(3 y 4), 327–334.

Calhoun, S. L., y Mayes, S. D. (2005). Processing speed in children with clinical disorders. *Psychology in the Schools, 42*(4), 333–343.

Campbell, D. T., y Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105.

Carlozzi, N. E., Grech, J., y Tulsky, D. S. (2013). Memory functioning in individuals with traumatic brain injury: An examination of the Wechsler Memory Scale–Fourth Edition (WMS–IV). *Journal of Clinical and Experimental Neuropsychology, 35*(9), 906–914. doi: 10.1080/13803395.2013.833178

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, England: Cambridge University Press.

Carroll, J. B. (2012). The three-stratum theory of cognitive abilities. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed., pp. 883–890). Nueva York, NY: Guilford Press.

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin, 38,* 592.

Cattell, R. B. (1957). *Personality and motivation structure and measurement.* Nueva York, NY: World Book.

Catts, H. W., Adlof, S. M., y Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*(2), 278–293. doi: 10.1044/1092-4388(2006/023)

Champod, A. S., y Petrides, M. (2010). Dissociation within the frontoparietal network in verbal working memory: A parametric functional magnetic resonance imaging study. *The Journal of Neuroscience, 30*(10), 3849–3856. doi: 10.1523/JNEUROSCI.0097-10.2010

Chen, H., Keith, T. Z., Weiss, L., Zhu, J., y Li, Y. (2010). Testing for multigroup invariance of second-order WISC–IV structure across China, Hong Kong, Macau, and Taiwan. *Personality and Individual Differences, 49,* 677–682. doi: 10.1016/j.paid.2010.06.004

Chiang, H.-L., Huang, L.-W., Gau, S. S.-F., y Shang, C.-Y. (2013). Associations of symptoms and subtypes of attention-deficit hyperactivity disorder with visuospatial planning ability in youth. *Research in Developmental Disabilities, 34*(9), 2986–2995. doi: 10.1016/j.ridd.2013.06.020

Chiang, M. C., McMahon, K. L., de Zubicaray, G. I., Martin, N. G., Hickie, I., Toga, A. W.,...Thompson, P. M. (2011). Genetics of white matter development: A DTI study of 705 twins and their siblings aged 12 to 29. *Neuroimage, 54*(3), 2308–2317. doi: 10.1016/j.neuroimage.2010.10.015

Chilosi, A. M., Brizzolara, D., Lami, L., Pizzoli, C., Gasperini, F., Pecini, C.,...y Zoccolotti, P. (2009). Reading and spelling disabilities in children with and without a history of early language delay: A neuropsychological and linguistic study. *Child Neuropsychology, 15,* 582–604. doi: 10.1080/09297040902927614

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., y Gabrieli, J. D. E. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage, 14,* 1136–1149. doi: 10.1006/nimg.2001.0922

Chuderski, A. (2013). When are fluid intelligence and working memory isomorphic and when are they not? *Intelligence, 41,* 244–262. doi: 10.1016/j.intell.2013.04.003

Cirino, P. T., Morris, M. K., y Morris, R. D. (2007). Semantic, executive, and visuospatial abilities in mathematical reasoning of referred college students. *Assessment, 14*(1), 94–104. doi: 10.1177/1073191106291487

Coalson, D. L., Raiford, S. E., Saklofske, D. H., y Weiss, L. G. (2010). WAIS–IV: Advances in the assessment of intelligence. En L. G. Weiss, D. H. Saklofske, D. L. Coalson, y S. E. Raiford (Eds.), *WAIS–IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 3–23). Amsterdam: Elsevier Academic Press.

Cohen, B. H. (1996). *Explaining psychological statistics.* Pacific Grove, CA: Brooks & Cole.

Cohen, J. (1957a). A factor-analytically based rationale for the Wechsler Adult Intelligence Scale. *Journal of Consulting Psychology, 21*(6), 451–457.

Cohen, J. (1957b). The factorial structure of the WAIS between early adulthood and old age. *Journal of Consulting Psychology, 21*(4), 283–290.

Cohen, J. (1959). The factorial structure of the WISC at ages 7–6, 10–6, and 13–6. *Journal of Consulting Psychology, 23,* 285–299.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2ª ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45,* 1304–1312.

Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin, 112*(1), 155–159.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist, 49,* 997–1003.

Colom, R., Burgaleta, M., Román, F. J., Karama, S., Álvarez-Linera, J., Abad, F. J.,...Haier, R. J. (2013). Neuroanatomic overlap between intelligence and cognitive factors: Morphometry methods provide support for the key role of the frontal lobes. *NeuroImage, 72,* 143–152. doi: 10.1016/j.neuroimage.2013.01.032

Colom, R., Haier, R. J., Head, K., Álvarez-Linera, J., Quiroga, M. Á., Shih, P. C., y Jung, R. E. (2009). Gray matter correlates of fluid, crystallized, and spatial intelligence: Testing the P-FIT model. *Intelligence, 37*(2), 124–135. doi: 10.1016/j.intell.2008.07.007

Colom, R., Jung, R. E., y Haier, R. J. (2007). General intelligence and memory span: Evidence for a common neuroanatomic framework. *Cognitive Neuropsychology, 24*(8), 867–878. doi: 10.1080/02643290701781557

Colom, R., Karama, S., Jung, R. E., y Haier, R. J. (2010). Human intelligence and brain networks. *Dialogues in Clinical Neuroscience, 12*(4), 489–501.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., y Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by *g*. *Intelligence, 32,* 277–296. doi: 10.1016/j.intell.2003.12.002

Colom, R., Stein, J. L., Rajagopalan, P., Martínez, K., Hermel, D., Wang, Y.,...Thompson, P. M. (2013). Hippocampal structure and human cognition: Key role of spatial processing and evidence supporting the efficiency hypothesis in females. *Intelligence, 41*(2), 129–140.

Compton, D. L., Fuchs, L. S., Fuchs, D., Lambert, W., y Hamlett, C. (2012). The cognitive and academic profiles of reading and mathematics learning disabilities. *Journal of Learning Disabilities, 45,* 79–95. doi: 10.1177/0022219410393012

Corbett, B. A., Constantine, L. J., Hendren, R., Rocke, D., y Ozonoff, S. (2009). Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder and typical development. *Psychiatry Research, 166*(2–3), 210–222. doi: 10.1016/j.psychres.2008.02.005

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*(2), 163–191.

Cowan, N. (1999). An embedded-processes model of working memory. En A. Miyake y P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Nueva York, NY: Cambridge University Press.

Cowan, N., y Alloway, T. (2009). Development of working memory in childhood. En M. L. Courage y N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 303–342). Nueva York, NY: Psychology Press.

Crocker, L., y Algina, J. (1986). *Introduction to classical and modern test theory.* Fort Worth, TX: Harcourt Brace Jovanovich College.

Cronbach, L. J., y Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

Crone, E. A., Wendelken, C., Donohue, S., van Leijenhorst, L., y Bunge, S. A. (2006). Neurocognitive development of the ability to manipulate information in working memory. *Proceedings of the National Academy of Sciences, 103*(24), 9315–9320. doi: 10.1073/pnas.0510088103

Crosbie, J., Arnold, P., Paterson, A., Swanson, J., Dupuis, A., Li, X.,...Schachar, R. J. (2013) Response inhibition and ADHD traits: Correlates and heritability in a community sample. *Journal of Abnormal Child Psychology, 41,* 497–507. doi: 10.1007/s10802-012-9693-9

Crowe, S. F. (2000). Does the letter number sequencing task measure anything more than digit span? *Assessment, 7*(2), 113–117.

Cureton, E. E., y Tukey, J. W. (1951). Smoothing frequency distributions, equating tests, and preparing norms. *American Psychologist, 6,* 404.

Cutting, L. E., Materek, A., Cole, C. A. S., Levine, T. M., y Mahone, E. M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Annals of Dyslexia, 59*(1), 34–54. doi: 10.1007/s11881-009-0022-0

D' Esposito, M., Aguirre, G. K., Zarahn, E., Ballard, D., Shin, R. K., y Lease, J. (1998). Functional MRI studies of spatial and nonspatial working memory. *Cognitive Brain Research, 7*(1), 1–13. doi: 10.1016/S0926-6410(98)00004-4

Dang, C.-P., Braeken, J., Colom, R., Ferrer, E., y Liu, C. (2014). Why is working memory related to intelligence? Different contributions from storage and processing, *Memory, 22*(4), 426–441, doi: 10.1080/09658211.2013.797471

Davis, F. B. (1959). Interpretation of differences among averages and individual test scores. *Journal of Educational Psychology, 50*(4), 162–170.

Dawson, M., Soulières, I., Gernsbacher, M. A., y Mottron, L. (2007). The level and nature of autistic intelligence. *Psychological Science, 18*(8), 657–662. doi: 10.1111/j.1467-9280.2007.01954.x

de Jonge, P., y de Jong, P. F. (1996). Working memory, intelligence and reading ability in children. *Personality and Individual Differences, 21*(6), 1007–1020.

De Monte, V. E., Geffen, G. M., May, C. R., y McFarland, K. (2010). Improved sensitivity of the rapid screen of mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 32*(1), 28–37. doi: 10.1080/13803390902806519

Deary, I. J., y Johnson, W. (2010). Intelligence and education: Causal perceptions drive analytic processes and therefore conclusions. *International Journal of Epidemiology, 39*(5), 1362–1369. doi: 10.1093/ije/dyq072

Deary, I. J., Penke, L., y Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience, 11*, 201–211. doi: 10.1038/nrn2793

Deary, I. J., Strand, S., Smith, P., y Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*(1), 13–21.

Dehaene, S. (2010). The massive impact of literacy on the brain and its consequences for education. En A. M. Battro, S. Dehaene, y W. J. Singer (Eds.), *Human neuroplasticity and education: The proceedings of the working group* (pp. 19–32). Vatican City: Pontifical Academy of Science.

Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., y Dehaene, S. (2008). How does early brain organization promote language acquisition in humans? *European Review, 16*(4), 399–411. doi: 10.1017/S1062798708000513

Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Mériaux, S., Roche, A., Sigman, M., y Dehaene, S. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proceedings of the National Academy of Sciences, 103*(38), 14240–14245. doi: 10.1073/pnas.0606302103

Demas, S., y Brogdon, J. L. (1997). Determining copyright status for preservation and access: Defining reasonable effort. *Library Resources & Technical Services, 41*(4), 323–334.

Devena, S. E., y Watkins, M. W. (2012). Diagnostic utility of WISC–IV general abilities index and cognitive proficiency index difference scores among children with ADHD. *Journal of Applied School Psychology, 28*(2), 133–154. doi: 10.1080/15377903.2012.669743

Dikmen, S. S., Machamer, J. E., Winn, H. R., y Temkin, N. R. (1995). Neuropsychological outcome at 1-year post head injury. *Neuropsychology, 9*(1), 80–90.

Dodonova, Y. A., y Dodonov, Y. S. (2012). Processing speed and intelligence as predictors of school achievement: Mediation or unique contribution? *Intelligence, 40*, 163–171. doi: 10.1016/j.intell.2012.01.003

Donders, J. (1997). Sensitivity of the WISC–III to injury severity in children with traumatic head injury. *Assessment, 4*, 107–109.

Donders, J., y Nesbit-Greene, K. (2004). Predictors of neuropsychological test performance after pediatric traumatic brain injury. *Assessment, 11*(4), 275–284. doi: 10.1177/1073191104268914

Donders, J., Tulsky, D. S., y Zhu, J. (2001). Criterion validity of new WAIS–III subtest scores after traumatic brain injury. *Journal of the International Neuropsychological Society, 7*, 892–898.

Dougherty, D. M., Dew, R. E., Mathias, C. W., Marsh, D. M., Addicott, M. A., y Barratt, E .S. (2007). Impulsive and premeditated subtypes of aggression in conduct disorder: Difference in time estimation. *Aggressive Behavior, 33*, 574–582. doi: 10 .1002/ab .20219

Drabick, D. A. G., Bubier, J., Chen, D., Price, J., y Lanza, H. I. (2011). Source-specific oppositional defiant disorder among inner-city children: Prospective prediction and moderation. *Journal of Clinical Child & Adolescent Psychology, 40*(1), 23–35. doi: 10.1080/15374416.2011.533401

Duan, X., Shi, J., y Zhou, D. (2010). Developmental changes in processing speed: Influence of accelerated education for gifted children. *Gifted Child Quarterly, 54*(2), 85–91.

Dubois, J., Hertz-Pannier, L., Cachia, A., Mangin, J. F., Le Bihan, D., y Dehaene-Lambertz, G. (2009). Structural asymmetries in the infant language and sensori-motor networks. *Cerebral Cortex, 19*(2), 414–423. doi: 10.1093/cercor/bhn097

Dudek, F. J. (1979). The continuing misinterpretation of the standard error of measurement. *Psychological Bulletin, 86*(2), 335–337.

Dumontheil, I., y Klingberg, T. (2012). Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later. *Cerebral Cortex, 22*(5), 1078–1085. doi: 10.1093/cercor/bhr175

Dunst, C. J., y Hamby, D. W. (2012). Guide for calculating and interpreting effect sizes and confidence intervals in intellectual and developmental disability research studies. *Journal of Intellectual and Developmental Disability, 37*(2), 89–99.

Eid, M., y Nussbeck, F. W. (2009). The multitrait multimethod matrix at 50! *Methodology, 5*(3), 71. doi: 10.1027/1614-2241.5.3.71

Elliott, C. D. (2007). *Differential ability scales* (2ª ed.). San Antonio, TX: Harcourt Assessment.

Englund, J. A., Decker, S. L., Allen, R. A., y Roberts, A. M. (2014). Common cognitive deficits in children with attention-deficit/hyperactivity disorder and autism: Working memory and visual-motor integration. *Journal of Psychoeducational Assessment, 32*(2), 95–106. doi: 10.1177/0734282913505074

Fiorello, C. A., Hale, J. B., y Snyder, L. E. (2006). Cognitive hypothesis testing and response to intervention for children with reading problems. *Psychology in the Schools, 43*(8), 835–853.

Fiorello, C. A., Hale, J. B., y Wycoff, K. L. (2012). Cognitive hypothesis testing: Linking tests results to the real world. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed., pp. 484–496). Nueva York, NY: Guilford Press.

Fitzpatrick, C., y Pagani, L. S. (2012). Toddler working memory skills predict kindergarten school readiness. *Intelligence, 40*(2), 205–212. doi: 10.1016/j.intell.2011.11.007

Flanagan, D. P., Alfonso, V. C., y Ortiz, S. O. (2012). The cross-battery assessment approach: An overview, historical perspective, and current directions. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (3ª ed., pp. 459–483). Nueva York, NY: The Guilford Press.

Flanagan, D. P., Alfonso, V. C., Ortiz, S. O., y Dynda, A. M. (2010). Integrating cognitive assessment in school neuropsychological evaluations. En D. C. Miller (Ed.), *Best practices in school neuropsychology: Guidelines for effective practice, assessment, and evidence-based intervention*, pp. 101–140. Hoboken, NJ: John Wiley & Sons.

Flanagan, D. P., Fiorello, C. A., y Ortiz, S. O. (2010). Enhancing practice through application of Cattell-Horn-Carroll theory and research: A "third method" approach to specific learning disability identification. *Psychology in the Schools, 47*(7), 739–760.

Flanagan, D. P., y Harrison, P. L. (Eds.). (2012). *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed.). Nueva York, NY: Guilford Press.

Flanagan, D. P., y Kaufman, A. S. (2009). *Essentials of WISC–IV assessment* (2ª ed.). Hoboken, NJ: John Wiley & Sons.

Flanagan, D. P., Ortiz, S. O., y Alfonso, V. C. (2007). *Essentials of cross-battery assessment* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., y Mascolo, J. T. (2006). *The achievement test desk reference: A guide to learning disability identification* (2ª ed.). Hoboken, NJ: John Wiley & Sons.

Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist, 54*(1), 5–20.

Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn effect.* Cambridge, England: Cambridge University Press.

Flynn, J. R., y Weiss, L. G. (2007). American IQ gains from 1932 to 2002: The WISC subtests and educational progress. *International Journal of Testing, 7*(2), 209–224.

Frangou, S., Chitins, X., y Williams, S. C. R. (2004). Mapping IQ and gray matter density in healthy young people. *NeuroImage, 23*(3), 800–805. doi: 10.1016/j.neuroimage.2004.05.027

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., y Bryant, J. D. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development, 81*(5), 1520–1533. doi: 10.1111/j.1467-8624.2010.01489.x

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M.,...Schatschneider, C. (2010). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities? *Developmental Psychology, 46*(6), 1731–1746. doi: 10.1037/a0020662

Gathercole, S. E., Pickering, S. J., Ambridge, B., y Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177–190. doi: 10.1037/0012-1649.40.2.177

Gau, S. S.-F., y Huang, W.-L. (2014). Rapid visual information processing as a cognitive endophenotype of attention deficit hyperactivity disorder. *Psychological Medicine, 44*, 435–446. doi: 10.1017/S0033291713000640

Gauger, L. M., Lombardino, L. J., y Leonard, C. M. (1997). Brain morphology in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*, 1272–1284. doi: 10.1044/jslhr.4006.1272

Geake, J. G. (2008). High abilities at fluid analogizing: A cognitive neuroscience construct of giftedness. *Roeper Review, 30*, 187–195. doi: 10.1080/02783190802201796

Geary, D. C. (2010). Mathematical disabilities: Reflections on cognitive, neuropsychological, and genetic components. *Learning and Individual Differences, 20*(2), 130. doi: 10.1016/j.lindif.2009.10.008

Geary, D. C. (2011a). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*(6), 1539–1552. doi: 10.1037/a0025510.

Geary, D. C. (2011b). Consequences, characteristics, and causes of mathematical learning disabilities and persistent low achievement in mathematics. *Journal of Developmental and Behavioral Pediatrics, 32*(3), 250–263. doi: 10.1097/DBP.0b013e318209edef.

Georgieff, M. K. (2007). Nutrition and the developing brain: Nutrient priorities and measurement. *The American Journal of Clinical Nutrition, 85*(Suppl), 614S–620S.

Giedd, J. N. (2008). The teen brain: Insights from neuroimaging. *Journal of Adolescent Health, 42*, 335–343. doi: 10.1016/j.jadohealth.2008.01.007

Giedd, J. N., Schmitt, J. E., y Neale, M. C. (2007). Structural brain magnetic resonance imaging of pediatric twins. *Human Brain Mapping, 28*, 474–481. doi: 10.1002/hbm.20403

Gilchrist, A., Green, J., Cox, A., Burton, D., Rutter, M., y Le Couteur, A. L. (2001). Development and current functioning in adolescents with Asperger syndrome: A comparative study. *Journal of Child Psychology and Psychiatry, 42*(2), 227–240.

Giofrè, D., Mammarella, I. C., y Cornoldi, C. (2013). The structure of working memory and how it relates to intelligence in children. *Intelligence, 41*, 396–406. doi: 10.1016/j.intell.2013.06.006

Glutting, J. J., McDermott, P. A., y Stanley, J. C. (1987). Resolving differences among methods of establishing confidence limits for test scores. *Educational and Psychological Measurement, 47*(3), 607–614.

Glutting, J. J., Watkins, M. W., Konold, T. R., y McDermott, P. A. (2006). Distinctions without a difference: The utility of observed versus latent factors from the WISC–IV in estimating reading and math achievement on the WIAT–II. *The Journal of Special Education, 40*(2), 103–114.

Goel, V. (2007). Anatomy of deductive reasoning. *TRENDS in Cognitive Sciences, 11*(10), 435–441. doi: 10.1016/j.tics.2007.09.003

Goel, V., Buchel, C., Frith, C., y Dolan, R. J. (2000). Dissociation of mechanisms underlying syllogistic reasoning. *NeuroImage, 12*, 504–514. doi: 10.1006/nimg.2000.0636

Goel, V., y Dolan, R. J. (2004). Differential involvement of left prefrontal cortexin inductive and deductive reasoning. *Cognition, 93*(3), B109–B121. doi: 10.1016/j.cognition.2004.03.001

Goldberg, R. F., Perfetti, C. A., Fiez, J. A., y Schneider, W. (2007). Selective retrieval of abstract semantic knowledge in left prefrontal cortex. *The Journal of Neuroscience, 27*(14), 3790–3798. doi: 10.1523/JNEUROSCI.2381-06.2007

Goldstein, G., y Beers, S. R. (Eds.). (2003). *Comprehensive handbook of psychological assessment, Vol. 1.: Intellectual and neuropsychological assessment.* Hoboken, NJ: John Wiley & Sons.

Goldstein, G., y Hersen, M. (2000). Historical perspectives. En G. Goldstein y M. Hersen (Eds.), *Handbook of psychological assessment* (3ª ed., pp. 3–14). Kidlington, Oxford, United Kingdom: Elsevier Science.

Gordon, S., Duff, S., Davidson, T., y Whitaker, S. (2010). Comparison of the WAIS–III and WISC–IV in 16-year-old special education students. *Journal of Applied Research in Intellectual Disabilities, 23*, 197–200. doi: 10.1111/j.1468-3148-2009-00538.x

Gorman, S., Barnes, M. A., Swank, P. R., Prasad, M., y Ewing-Cobbs, L. (2012). The effects of pediatric traumatic brain injury on verbal and visual-spatial working memory. *Journal of the International Neuropsychological Society, 18*, 29–38. doi: 10.1017/S1355617711001251

Gottfredson, L., y Saklofske, D. H. (2009). Intelligence: Foundations and issues in assessment. *Canadian Psychology, 50*(3), 183–195. doi: 10.1037/a0016641

Gottlieb, L., Zelko, F. A., Kim, D. S., y Nordli, D. R. (2012). Cognitive proficiency in pediatric epilepsy. *Epilepsy & Behavior, 23*, 146–151. doi: 10.1016/j.yebeh.2011.10.024

Goyal, M. S., y Raichle, M. E. (2013). Gene expression-based modeling of human cortical synaptic density. *Proceedings of the National Academy of Sciences, 110*(16), 6571–6576. doi: 10.1073/pnas.1303453110

Green, A. E., Kraemer, D. J. M., Fugelsang, J. A., Gray, J. R., y Dunbar, K. N. (2010). Connecting long distance: Semantic distance in analogical reasoning modulates frontopolar cortex activity. *Cerebral Cortex, 20*, 70–76. doi: 10.1093/cercor/bhp081

Green, R. E. A., Melo, B., Christensen, B., Ngo, L. A., Monette, G., y Bradbury, C. (2008). Measuring premorbid IQ in traumatic brain injury: An examination of the validity of the Wechsler Test of Adult Reading (WTAR). *Journal of Clinical and Experimental Neuropsychology, 30*, 163–172.

Grégoire, J., Coalson, D. L., y Zhu, J. (2011). Analysis of WAIS-IV index score scatter using significant deviation from the mean index score. *Assessment, 18*(2), 168–177. doi: 10.1177/1073191110386343

Groth-Marnat, G. (2009). *Handbook of psychological assessment* (5ª ed.). Nueva York, NY: John Wiley & Sons.

Gualtieri, C. T., y Johnson, L. G. (2006). Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Archives of Clinical Neuropsychology, 21*, 623–643.

Guilford, J. P. (1954). *Psychometric methods* (2ª ed.). Nueva York: McGraw-Hill.

Guilford, J. P., y Fruchter, B. (1978). *Fundamental statistics in psychology and education* (6ª ed.). Nueva York, NY: McGraw-Hill.

Gunning-Dixon, F. M., Brickman, A. M., Cheng, J. C., y Alexopoulos, G. S. (2009). Aging of cerebral white matter: A review of MRI findings. *International Journal of Geriatric Psychiatry, 24*(2), 109–117. doi: 10.1002/gps.2087

Haertel, E. H. (2006). Reliability. En R. L. Brennan (Ed.), *Educational measurement* (4ª ed., pp. 65–110). Westport, CT: Praeger.

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., y Alkire, M. T. (2005). The neuroanatomy of general intelligence: Sex matters. *NeuroImage, 25*, 320–327. doi: 10.1016/j.neuroimage.2004.11.019

Hale, J. B., y Fiorello, C. A. (2004). *School neuropsychology: A practitioner's handbook.* Nueva York, NY: Guilford Press.

Hale, J. B., Fiorello, C. A., y Brown, L. L. (2005). Determining medication treatment effects using teacher ratings and classroom observations of children with ADHD: Does neuropsychological impairment matter? *Educational & Child Psychology, 22*(2), 39–61.

Hale, J. B., Yim, M., Schneider, A. N., Wilcox, G., Henzel, J. N., y Dixon, S. G. (2012). Cognitive and neuropsychological assessment of attention-deficit/hyperactivity disorder: Redefining a disruptive behavior disorder. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed., pp. 687–707). Nueva York, NY: Guilford Press.

Hambleton, R. K. (1993). Principles and selected applications of item response theory. En R. L. Linn (Ed.), *Educational measurement* (3ª ed., pp. 147–220). Phoenix, AZ: Oryx Press.

Harris, J. G., y Llorente, A. M. (2005). Cultural considerations in the use of the Wechsler Intelligence Scale for Children–Fourth Edition (WISC–IV). En A. Prifitera, D. H. Saklofske, y L. G. Weiss (Eds.), *WISC–IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 381–413). Amsterdam: Elsevier Academic Press.

Harrison, A. G., DeLisle, M. M., y Parker, K. C. H. (2008). An investigation of the general abilities index in a group of diagnostically mixed patients. *Journal of Psychoeducational Assessment, 26*(3), 247–259. doi: 10.1177/0734282907304032

Hartshorne, J. K. (2008). Visual working memory capacity and proactive interference. *PLoS ONE, 3*(7), e2716. doi: 10.1371/journal.pone.0002716

Hiscock, M., Inch, R., y Gleason, A. (2002). Raven's Progressive Matrices performance in adults with traumatic brain injury. *Applied Neuropsychology, 9*, 129–138.

Holland, P. W., y Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. En H. Wainer y H. I. Braun (Eds.), *Test validity* (pp. 129–145). Hillsdale, NJ: Erlbaum.

Horn, J. L. (1985). Remodeling old models of intelligence. En B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurement, and applications* (pp. 267–300). Nueva York, NY: John Wiley & Sons.

Horn, J. L. (1988). Thinking about human abilities. En J. R. Nesselroade y R. B Cattell (Eds.), *Handbook of multivariate experimental psychology* (2ª ed., pp. 645–685). Nueva York, NY: Plenum.

Horn, J. L. (1991). Measurement of intellectual capabilities: A review of theory. En K. S. McGrew, J. K. Werder, y R. W. Woodcock (Eds.), *WJ–R technical manual* (pp. 197–232). Itasca, IL: Riverside.

Horn, J. L., y Blankson, A. N. (2012). Foundations for better understanding of cognitive abilities. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed., pp. 73–98). Nueva York, NY: Guilford Press.

Horn, J. L., y Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. En D. P. Flanagan, J. L. Genshaft, y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53–91). Nueva York, NY: Guilford Press.

Hornung, C., Brunner, M., Reuter, R. A. P., y Martin, R. (2011). Children's working memory: Its structure and relationship to fluid intelligence. *Intelligence, 39*, 210–221.

Horowitz-Kraus, T. (2014). Pinpointing the deficit in executive functions in adolescents with dyslexia performing the Wisconsin Card Sorting Test: An ERP study. *Journal of Learning Disabilities, 47*(3), 208–223. doi: 10.1177/0022219412453084

Hu, L., y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55.

Hubbard, R., y Lindsay, R. M. (2008). Why *p* values are not a useful measure of evidence in statistical significance testing. *Theory & Psychology, 18*(1), 69–88.

Hugdahl, K. (2011). Hemispheric asymmetry: Contributions from brain imaging. *WIREs Cognitive Science, 2,* 461–478. doi: 10.1002/wcs.122

Hunt, E., y Madhyastha, T. M. (2012). Cognitive demands of the workplace. *Journal of Neuroscience, Psychology, and Economics, 5*(1), 18–37. doi: 10.1037/a0026177

Hutchinson, E., Bavin, E., Efron, D., y Sciberras, E. (2012). A comparison of working memory profiles in school-aged children with specific language impairment, attention deficit/hyperactivity disorder, comorbid SLI and ADHD and their typically developing peers. *Child Neuropsychology, 18*(2), 190–207. doi: 10.1080/09297049.2011.601288

Individuals with Disabilities Education Improvement Act of 2004, Pub. L. No. 108–446, 118 Stat. 2647. (2004).

Instituto Nacional de Estadística (2011). *Censo de población y viviendas.* Recuperado de http://www.ine.es

Isen, J. (2010). A meta-analytic assessment of Wechsler's P > V sign in antisocial populations. *Clinical Psychology Review, 30,* 423–435. doi: 10.1016/j.cpr.2010.02.003

Jacobson, L. A., Ryan, M., Martin, R. B., Ewen, J., Mostofsky, S. H., Denckla, M. B., y Mahone, E. M. (2011). Working memory influences processing speed and reading fluency in ADHD. *Child Neuropsychology, 17*(3), 209–224. doi: 10.1080/09297049.2010.532204.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

Johnson, W., y Bouchard, T. J., Jr. (2005a). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence, 33,* 393–416. doi: 10.1016/j.intell.2004.12.002

Johnson, W., y Bouchard, T. J., Jr. (2005b). Constructive replication of the visual-perceptual-image rotation model in Thurstone's (1941) battery of 60 tests of mental ability. *Intelligence, 33,* 417–430. doi: 10.1016/j.intell.2004.12.001

Johnson, W., Bouchard, T. J., Jr., Krueger, R. F., McGue, M., y Gottesman, I. I. (2004). Just one *g*: Consistent results from three test batteries. *Intelligence, 32,* 95–107.

Johnson, W., Corley, J., Starr, J. M., y Deary, I. J. (2011). Psychological and physical health at age 70 in the Lothian birth cohort 1936: Links with early life IQ, SES, and current cognitive function and neighborhood environment. *Health Psychology, 30*(1), 1–11.

Johnson, W., Deary, I. J., y Iacono, W. G. (2009). Genetic and environmental transactions underlying educational attainment. *Intelligence, 37*(5), 466–478. doi: 10.1016/j.intell.2009.05.006

Johnson, W., te Nijenhuis, J., y Bouchard, T. J., Jr. (2007). Replication of the hierarchical visual-perceptual-image rotation model in de Wolff and Buiten's (1963) battery of 46 tests of mental ability. *Intelligence, 35,* 69–81.

Jolles, D. D., Kleibeuker, S. W., Rombouts, S. A. R. B., y Crone, E. A. (2011). Developmental differences in prefrontal activation during working memory maintenance and manipulation for different memory loads. *Developmental Science, 14*(4), 713–724. doi: 10.1111/j.1467-7687.2010.01016.x

Jöreskog, K. G., y Sörbom, D. (1993). *LISREL 8: User's reference guide.* Chicago, IL: Scientific Software International.

Joseph, R. M., Tager-Flusberg, H., y Lord, C. (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 43*(6), 807–821. doi: 10.1111/1469-7610.00092

Kamphaus, R. W. (1993). *Clinical assessment of children's intelligence: A handbook for professional practice.* Needham Heights, MA: Allyn & Bacon.

Kane, M. J., y Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9*(4), 637–671.

Kaplan, E. (1988). A process approach to neuropsychological assessment. En T. J. Boll y B. K. Bryant (Eds.), *Clinical neuropsychology and brain function: Research, measurement, and practice* (pp. 129–167). Washington, DC: American Psychological Association.

Kaplan, E., Fein, D., Morris, R., y Delis, D. C. (1991). *WAIS–R as a neuropsychological instrument.* San Antonio, TX: The Psychological Corporation.

Karama, S., Colom, R., Johnson, W., Deary, I. J., Haier, R., Waber, D. P.,...the Brain Development Group (2011). Cortical thickness correlates of specific cognitive performance accounted for by the general factor of intelligence in healthy children aged 6 to 18. *NeuroImage, 55*(4), 1443–1453. doi: 10.1016/j.neuroimage.2011.01.016

Kasper, L. J., Alderson, R. M., y Hudec, K. L. (2012). Moderators of working memory deficits in children with attention-deficit/hyperactivity disorder (ADHD): A meta-analytic review. *Clinical Psychology Review, 32,* 605–617. doi: 10.1016/j.cpr.2012.07.001

Kaufman, A. S. (1975). Factor analysis of the WISC–R at eleven age levels between 6 1/2 and 16 1/2 years. *Journal of Consulting and Clinical Psychology, 43*(2), 135–147.

Kaufman, A. S. (1994). *Intelligent testing with the WISC–III.* Nueva York, NY: Wiley

Kaufman, A. S. (2010). Foreword. En L. G. Weiss, D. H. Saklofske, D. L. Coalson, y S. E. Raiford (Eds.), *WAIS–IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. xiii–xxi). Amsterdam: Elsevier Academic Press.

Kaufman, A. S., Flanagan, D. P., Alfonso, V. C., y Mascolo, J. T. (2006). Test review: Wechsler Intelligence Scale for Children, Fourth Edition (WISC–IV). *Journal of Psychoeducational Assessment, 24,* 278–295.

Kaufman, A. S., y Kaufman, N. L. (2004). *Kaufman assessment battery for children* (2ª ed.). Bloomington, MN: NCS Pearson.

Kaufman, A. S., y Lichtenberger, E. O. (2006). *Assessing adolescent and adult intelligence* (3ª ed.). Hoboken, NJ: John Wiley & Sons.

Kaufman, A. S., Lichtenberger, E. O., y McLean, J. (2001). Two- and three-factor solutions of the WAIS–III. *Assessment, 8*(3), 267–280.

Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., y McGrew, K. S. (2012). Are cognitive *g* and academic *g* one and the same *g*? An exploration on the Woodcock-Johnson and Kaufman tests. *Intelligence, 40,* 123–138. doi: 10.1016/j.intell.2012.01.009

Kavale, K. A., Holdnack, J. A., y Mostert, M. P. (2005). Responsiveness to intervention and the identification of specific learning disability: A critique and alternative proposal. *Learning Disability Quarterly, 28,* 2–16.

Keenan, H. T., y Bratton, S. L. (2006). Epidemiology and outcomes of pediatric traumatic brain injury. *Developmental Neuroscience, 28*(4–5), 256–263. doi: 10.1159/000094152

Keith, T. Z. (1985). Questioning the K–ABC: What does it measure? *School Psychology Review, 14*(1), 19–20.

Keith, T. Z. (1990). Confirmatory and hierarchical confirmatory analysis of the Differential Ability Scales. *Journal of Psychoeducational Assessment, 8*(3), 391–405.

Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., y Kranzler, J. H. (2006). Higher order, multisample, confirmatory factor analysis of the Wechsler Intelligence Scale for Children–Fourth Edition: What does it measure? *School Psychology Review, 35*(1), 108–127.

Kibby, M. Y., y Cohen, M. J. (2008). Memory functioning in children with reading disabilities and/or attention deficit/hyperactivity disorder: A clinical investigation of their working memory and long-term memory functioning. *Child Neuropsychology, 14*(6), 525–546. doi: 10.1080/09297040701821752.

Kirk, R. E. (2001). Promoting good statistical practices: Some suggestions. *Educational and Psychological Measurement, 61,* 213–218.

Kirschen, M. P., Chen, S. H. A., Schraedley-Desmond, P., y Desmond, J. E. (2005). Load- and practice-dependent increases in cerebro-cerebellar activation in verbal working memory: An fMRI study. *NeuroImage, 24,* 462–472. doi: 10.1016/j.neuroimage.2004.08.036

Klingberg, T. (2006). Development of a superior frontal–intraparietal network for visuo-spatial working memory. *Neuropsychologia, 44*(11), 2171–2177. doi: 10.1016/j.neuropsychologia.2005.11.019

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences, 14,* 317–324. doi: 10.1016/j.tics.2010.05.002

Klinger, L. G., O'Kelley, S. E., Mussey, J. L., Goldstein, S., y DeVries, M. (2012). Assessment of intellectual functioning in autism spectrum disorder. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed., pp. 670–686). Nueva York, NY: Guilford Press.

Knauff, M., Mulack, T., Kassubek, J., Salih, H. R., y Greenlee, M. W. (2002). Spatial imagery in deductive reasoning: A functional MRI study. *Cognitive Brain Research, 13,* 203–212. doi: 10.1016/S0926-6410(01)00116-1

Knaus, T. A., Silver, A. M., Dominick, K. C., Schuring, M. D., Shaffer, N., Lindgren, K. A.,...Tager-Flusberg, H. (2009). Age-related changes in the anatomy of language regions in autism spectrum disorder. *Brain Imaging and Behavior, 3*(1), 51–63. doi 10.1007/s11682-008-9048-x

Korkman, M., Kirk, U., y Kemp, S. (2007). *NEPSY–II.* Bloomington, MN: Pearson.

Kotz, S. A., Cappa, S. F., von Cramon, D. Y., y Friederici, A. D. (2002). Modulation of the lexical–semantic network by auditory semantic priming: An event-related functional MRI study. *NeuroImage, 17,* 1761–1772. doi: 10.1006/nimg.2002.1316

Koziol, L. F., y Budding, D. E. (2009). *Subcortical structures and cognition: Implications for neuropsychological assessment.* Nueva York, NY: Springer.

Koziol, L. F., Budding, D. E., y Chidekel, D. (2010). Adaptation, expertise, and giftedness: Towards an understanding of cortical, subcortical, and cerebellar network contributions. *Cerebellum, 9,* 499–529. doi: 10.1007/s12311-010-0192-7

Krawczyk, D. C., McClelland, M. M., y Donovan, C. M. (2011). A hierarchy for relational reasoning in the prefrontal cortex. *Cortex, 47*(5), 588–597. doi: 10.1016/j.cortex.2010.04.008

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron, 67,* 713–727. doi: 10.1016/j.neuron.2010.08.038

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., y Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society, 363,* 979–1000. doi: 10.1098/rstb.2007.2154

Kuncel, N. R., Ones, D. S., y Sackett, P. R. (2010). Individual differences as predictors of work, educational, and broad life outcomes. *Personality and Individual Differences, 49,* 331–336. doi: 10.1016/j.paid.2010.03.042

Kundu, B., Sutterer, D. W., Emrich, S. M., y Postle, B. R. (2013). Strengthened effective connectivity underlies transfer of working memory training to tests of short-term memory and attention. *The Journal of Neuroscience, 33*(20), 8705–8715. doi: 10.1523/JNEUROSCI.5565-12.2013

Kuwajima, M., y Sawaguchi, T. (2010). Similar prefrontal cortical activities between general fluid intelligence and visuospatial working memory tasks in preschool children as revealed by optical topography. *Experimental Brain Research, 206,* 381–397. doi: 10.1007/s00221-010-2415-z

Lakin, J. M., y Lohman, D. F. (2011). The predictive accuracy of verbal, quantitative, and nonverbal reasoning tests: Consequences for talent identification and program diversity. *Journal for the Education of the Gifted, 34*(4), 595–623. doi: 10.1177/016235321103400404

Lange, N., Froimowitz, M. P., Bigler, E. D., Lainhart, J. E., y Brain Development Cooperative Group. (2010). Associations between IQ, total and regional brain volumes, and demography in a large normative sample of healthy children and adolescents. *Developmental Neuropsychology, 35*(3), 296–317. doi: 10.1080/87565641003696833

Langeslag, S. J. E., Schmidt, M., Ghassabian, A., Jaddoe, V. W., Hofman, A., van der Lugt, A.,...White, T. J. H. (2013). Functional connectivity between parietal and frontal brain regions and intelligence in young children: The generation R study. *Human Brain Mapping, 34*(12), 3299–3307. doi: 10.1002/hbm.22143

Lederbogen, F., Kirsch, P., Haddad, L., Streit, F., Tost, H., Schuch, P.,...Meyer-Lindenberg, A. (2011). City living and urban upbringing affect neural social stress processing in humans. *Nature, 474*(7352), 498–501. doi: 10.1038/nature10190

Lee, K. H., Choi, Y. Y., Gray, J. R., Cho, S. H., Chae, J.-H., Lee, S., y Kim, K. (2006). Neural correlates of superior intelligence: Stronger recruitment of posterior parietal cortex. *NeuroImage, 29,* 578–586. doi: 10.1016/j.neuroimage.2005.07.036

Lenroot, R. K., y Giedd, J. N. (2006). Brain development in children and adolescents: Insights from anatomical magnetic resonance imaging. *Neuroscience and Biobehavioral Reviews, 30,* 718–729. doi: 10.1016/j.neubiorev.2006.06.001

Lezak, M. D., Howieson, D. B., y Loring, D. W. (con Hannay, H. J., y Fischer, J. S.). (2004). *Neuropsychological assessment* (4ª ed.). Nueva York, NY: Oxford University Press.

Li, H., Rosenthal, R., y Rubin, D. B. (1996). Reliability of measurement in psychology: From Spearman-Brown to maximal reliability. *Psychological Methods, 1*(1), 98–107.

Li, L., y Liu, J. (2013). The effect of pediatric traumatic brain injury on behavioral outcomes: A systematic review. *Developmental Medicine & Child Neurology, 55*(1), 37–45. doi: 10.1111/j.1469-8749.2012.04414.x

Li, Y., y Geary, D. C. (2013). Developmental gains in visuospatial memory predict gains in mathematics achievement. *PLoS ONE, 8*(7), e70160. doi: 10.1371/journal.pone.0070160

Lichtenberger, E. O., y Kaufman, A. S. (2013). *Essentials of WAIS–IV assessment* (2ª ed.). Hoboken, NJ: John Wiley & Sons.

Lichtenstein, R. (2013). Psychoeducational reports that matter: A consumer-responsive approach, part 2. *Communiqué, 42*(4), pp. 1, 10–13.

Lichtenstein, R. (2014). Psychoeducational reports that matter: A consumer-responsive approach, part 3. *Communiqué, 42*(6), pp. 1, 30–32.

Likert, R. y Quasha, W. H. (1995). *Revised Minnesota paper form board test manual* (2ª ed.). San Antonio, TX: The Psychological Corporation.

Lipinski, J., Simmering, V. R., Johnson, J. S., y Spencer, J. P. (2010). The role of experience in location estimation: Target distributions shift location memory biases. *Cognition, 115*(1), 147–153. doi: 10.1016/j.cognition.2009.12.008

Lipszyc, J., y Schachar, R. (2010). Inhibitory control and psychopathology: A meta-analysis of studies using the stop signal task. *Journal of the International Neuropsychological Society, 16*(6), 1064–1076. doi: 10.1017/S1355617710000895

Litt, R. A., y Nation, K. (2014). The nature and specificity of paired associate learning deficits in children with dyslexia. *Journal of Memory and Language, 71,* 71–88. doi: 10.1016/j.jml.2013.10.005

Lohman, D. F. (2009). Identifying academically talented students: Some general principles, two specific procedures. En L. V. Shavinina (Ed.), *International handbook on giftedness, part one* (pp. 971–997). The Netherlands: Springer. doi: 10.1007/978-1-4020-6162-2_49

Longman, R. S. (2004). Values for comparison of WAIS–III index scores with overall means. *Psychological Assessment, 16*(3), 323–325. doi: 10.1037/1040-3590.16.3.323

Low, L. K., y Cheng, H.-J. (2006). Axon pruning: An essential step underlying the developmental plasticity of neuronal connections. *Philosophical Transactions of the Royal Society B: Biological Sciences, 361,* 1531–1544. doi: 10.1098/rstb.2006.1883

MacDonald, M. C., Almor, A., Henderson, V. W., Kempler, D., y Andersen, E. S. (2001). Assessing working memory and language comprehension in Alzheimer's disease. *Brain and Language, 78,* 17–42.

Magnusson, D. (1967). *Test theory.* Reading, MA: Addison-Wesley.

Makovski, T., y Jiang, Y. V. (2008). Proactive interference from items previously stored in visual working memory. *Memory & Cognition, 36*(1), 43–52.

Maricle, D. E., y Avirett, E. (2012). The role of cognitive and intelligence tests in the assessment of executive functions. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed., pp. 820–838). Nueva York, NY: Guilford Press.

Maroof, D. A. (2012). *Statistical methods in neuropsychology: Common procedures made comprehensible.* Nueva York, NY: Springer.

Martinussen, R., Hayden, J., Hogg-Johnson, S., y Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*(4), 377–384.

Marvel, C. L., y Desmond, J. E. (2010). The contributions of cerebro-cerebellar circuitry to executive verbal working memory. *Cortex, 46*(7), 880–895. doi: 10.1016/j.cortex.2009.08.017

Matarazzo, J. D., Daniel, M. H., Prifitera, A., y Herman, D. O. (1988). Inter-subtest scatter in the WAIS–R standardization sample. *Journal of Clinical Psychology, 44,* 940–950.

Matarazzo, J. D., y Herman, D. O. (1985). Clinical uses of the WAIS–R: Base rates of differences between VIQ and PIQ in the WAIS–R standardization sample. En B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurement, and applications* (pp. 899–932). Nueva York, NY: John Wiley & Sons.

Mathias, J. L., y Wheaton, P. (2007). Changes in attention and information-processing speed following severe traumatic brain injury: A meta-analytic review. *Neuropsychology, 21*(2), 212–223.

Mattson, A. J., y Levin, H. S. (1990). Frontal lobe dysfunction following closed head injury: A review of the literature. *Journal of Nervous and Mental Disease, 178*(5), 282–291.

Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T., y Saxena, S. (2011). Prevalence of intellectual disability: A meta-analysis of population-based studies. *Research in Developmental Disabilities, 32,* 419–436. doi: 10.1016/j.ridd.2010.12.018

Mayes, S. D., y Calhoun, S. L. (2004). Similarities and differences in Wechsler Intelligence Scale for Children–Third Edition (WISC–III) profiles: Support for subtest analysis in clinical referrals. *The Clinical Neuropsychologist, 18,* 559–572.

Mayes, S. D., y Calhoun, S. L. (2007). Learning, attention, writing, and processing speed in typical children and children with ADHD, autism, anxiety, depression, and oppositional-defiant disorder. *Child Neuropsychology, 13,* 469–493. doi: 10.1080/09297040601112773

Mayes, S. D., y Calhoun, S. L. (2008). WISC–IV and WIAT–II profiles in children with high-functioning autism. *Journal of Autism and Developmental Disorders, 38,* 429–439. doi: 10.1007/s10803-007-0410-4

Mayes, S. D., Calhoun, S. L., Chase, G. A., Mink, D. M., y Stagg, R. E. (2009). ADHD subtypes and co-occurring anxiety, depression, and oppositional-defiant disorder: Differences in Gordon diagnostic system and Wechsler working memory and processing speed index scores. *Journal of Attention Disorders, 12*(6), 540–550. doi: 10.1177/1087054708320402

Mayes, S. D., Calhoun, S. L., Mayes, R. D., y Molitoris, S. (2012). Autism and ADHD: Overlapping and discriminating symptoms. *Research in Autism Spectrum Disorders, 6,* 277–285. doi: 10.1016/j.rasd.2011.05.009

McCloskey, G., Whitaker, J., Murphy, R., y Rogers, J. (2012). Intellectual, cognitive, and neuropsychological assessment in three-tier service delivery systems in schools. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (3ª ed., pp. 852–881). Nueva York, NY: Guilford Press.

McCrea, M., Janecek, J. K., Powell, M. R., y Hammeke, T. A. (2014). Traumatic brain injury and the postconcussion syndrome. En M. W. Parsons, T. A. Hammeke, y P. J. Synder (Eds.), *Clinical neuropsychology: A pocket handbook for assessment* (3ª ed., pp. 208–236). Washington, DC: American Psychological Association.

McGraw, K. O., y Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1,* 30–46.

McGrew, K. S., y Wendling, B. J. (2010). Cattell–Horn–Carroll cognitive-achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools, 47*(7), 651–675. doi: 10.1002/pits.20497

McNab, F., y Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience, 11*(1), 103–107. doi: 10.1038/nn2024

Messbauer, V. C. S., y de Jong, P. F. (2003). Word, nonword, and visual paired associate learning in Dutch dyslexic children. *Journal of Experimental Child Psychology, 84,* 77–96. doi: 10.1016/S0022-0965(02)00179-0

Metin, B., Roeyers, H., Wiersema, J. R., van der Meere, J. J., Thompson, M., y Sonuga-Barke, E. (2013). ADHD performance reflects inefficient but not impulsive information processing: A diffusion model analysis. *Neuropsychology, 27*(2), 193–200.

Meyer, M. L., Salimpoor, V. N., Wu, S. S., Geary, D. C., y Menon, V. (2010). Differential contribution of specific working memory components to mathematics achievement in 2nd and 3rd graders. *Learning and Individual Differences, 20,* 101–109. doi: 10.1016/j.lindif.2009.08.004

Michalczyk, K., Malstädt, N., Worgt, M., Könen, T., y Hasselhorn, M. (2013). Age differences and measurement invariance of working memory in 5- to 12-year-old children. *European Journal of Psychological Assessment, 29*(3), 220–229. doi: 10.1027/1015-5759/a000149

Miller, D. C. (2010). *Best practices in school neuropsychology: Guidelines for effective practice, assessment, and evidence-based intervention.* Hoboken, NJ: John Wiley & Sons.

Miller, D. C. (2013). *Essentials of school neuropsychological assessment* (2ª ed.). Hoboken, NJ: John Wiley & Sons.

Miller, D. C., y Maricle, D. E. (2012). The emergence of neuropsychological constructs into tests of intelligence and cognitive abilities. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed., pp. 800–819). Nueva York, NY: Guilford Press.

Monzalvo, K., y Dehaene-Lambertz, G. (2013). How reading acquisition changes children's spoken language network. *Brain & Language, 12,* 356–365. doi: 10.1016/j.bandl.2013.10.009

Morcillo, C., Duarte, C. S., Sala, R., Wang, S., Lejuez, C. W., Kerridge, B. T., y Blanco, C. (2012). Conduct disorder and adult psychiatric diagnoses: Associations and gender differences in the U.S. adult population. *Journal of Psychiatric Research, 46*(3), 323–330. doi: 10.1016/j.jpsychires.2011.10.012

Morris, P. E., y Fritz, C. O. (2013). Effect sizes in memory research. *Memory, 21*(7), 832–842. doi: 10.1080/09658211.2013.763984

Munro, J. (2013). High ability learning and brain processes: How neuroscience can help us to understand how gifted and talented students learn and the implications for teaching. En *ACER Research Conference 13. How the brain learns: What lessons are there for teaching?* (pp. 103–110). Camberwell, VIC 3124 Australia: Australian Council for Educational Research.

Naglieri, J. A., y Paolitto, A. W. (2005). Ipsative comparisons of WISC–IV index scores. *Applied Neuropsychology, 12*(4), 208–211.

Närhi, V., Lehto-Salo, P., Ahonen, T., y Marttunen, M. (2010). Neuropsychological subgroups of adolescents with conduct disorder. *Scandinavian Journal of Psychology, 51*(3), 278–284. doi: 10.1111/j.1467-9450.2009.00767.x

Nelson, J. M., Canivez, G. L., y Watkins, M. W. (2013). Structural and incremental validity of the Wechsler Adult Intelligence Scale–Fourth Edition with a clinical sample. *Psychological Assessment, 25*(2), 618–630. doi: 10.1037/a0032086

Newman, J. B., Reesman, J. H., Vaughan, C. G., y Gioia, G. A. (2013). Assessment of processing speed in children with mild TBI: A "first look" at the validity of pediatric ImPACT. *The Clinical Neuropsychologist, 27*(5), 779–793. doi: 10.1080/13854046.2013.789552

Niileksela, C. R., y Reynolds, M. R. (2014). Global, broad, or specific cognitive differences? Using a MIMIC model to examine differences in CHC abilities in children with learning disabilities. *Journal of Learning Disabilities, 47*(3), 224–236. doi: 10.1177/0022219412453172

Nock, M. K., Kazdin, A. E., Hiripi, E., y Kessler, R. C. (2007). Lifetime prevalence, correlates, and persistence of oppositional defiant disorder: Results from the national comorbidity survey replication. *Journal of Child Psychology and Psychiatry, 48*(7), 703–713. doi: 10.1111/j.1469-7610.2007.01733.x

Nunes, M. M., Honjo, R. S., Dutra, R. L., Amaral, V. A. S., Oh, H. K.,...Teixeira, M. C. T. V. (2012). Assessment of intellectual and visuo-spatial abilities in children and adults with Williams syndrome. *Universitas Psychologica, 12*(2), 581–589. doi: 10.11144/Javieriana.upsy12-2.aiva

Nunnally, J., y Bernstein, I. H. (1994). *Psychometric theory* (3ª ed.). Nueva York: McGraw-Hill.

Oliveras-Rentas, R. E., Kenworthy, L., Roberson, R., Mattin, A., y Wallace, G. (2012). WISC-IV profile in high-functioning autism spectrum disorders: Impaired processing speed is associated with increased autism communication symptoms and decreased adaptive communication abilities. *Journal of Autism and Developmental Disorders, 42,* 655–664. doi: 10.1007/s10803-011-1289-7

Ortiz, S. O., Ochoa, S. H., y Dynda, A. M. (2012). Testing with culturally and linguistically diverse populations: Moving beyond the verbal–performance dichotomy into evidence-based practice. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed., pp. 526–552). Nueva York, NY: Guilford Press.

Osmon, D. C., Smerz, J. M., Braun, M. M., y Plambeck, E. (2006). Processing abilities associated with math skills in adult learning disability. *Journal of Clinical and Experimental Neuropsychology, 28,* 84–95. doi: 10.1080/13803390490918129

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554 –566. doi: 10.1037/0022-0663.98.3.554

Owen, A. M., McMillan, K. M., Laird, A. R., y Bullmore, E. (2005). N-Back working memory paradigm: A meta-analysis of normative functional neuroimaging studies paradigm. *Human Brain Mapping, 25,* 46–59. doi: 10.1002/hbm.20131

Palva, J. M., Monto, S., Kulashekhar, S., y Palva, S. (2010). Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *Proceedings of the National Academy of Sciences, 107*(16), 7580–7585. doi: 10.1073/pnas.0913113107

Parsons, L. M., y Osherson, D. (2001). New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning. *Cerebral Cortex, 11*(10), 954–965.

Passolunghi, M. C. (2006). Working memory and arithmetic learning disability. En T. P. Alloway y S. E. Gathercole (Eds.), *Working memory and neurodevelopmental disorders* (pp. 113–138). Nueva York, NY: Psychology Press.

Petersen, I. T., Bates, J. E., D'Onofrio, B. M., Coyne, C. A., Lansford, J. E., Dodge, K. A.,...Van Hulle, C. A. (2013). Language ability predicts the development of behavior problems in children. *Journal of Abnormal Psychology, 122*(2), 542–557. doi: 10.1037/a0031963

Pfeiffer, S. I. (2013). *Serving the gifted: Evidence-based clinical and psychoeducational practice.* Nueva York, NY: Routledge.

Pickering, S. J. (2006). Working memory in dyslexia. En T. P. Alloway y S. E. Gathercole (Eds.), *Working memory and neurodevelopmental disorders* (pp. 7–40). Nueva York, NY: Psychology Press.

Pimperton, H., y Nation, K. (2014). Poor comprehenders in the classroom: Teacher ratings of behavior in children with poor reading comprehension and its relationship with individual differences in working memory. *Journal of Learning Disabilities, 47*(3), 199–207. doi: 10.1177/0022219412454172

Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., y Rohde, L. A. (2007). The worldwide prevalence of ADHD: A systematic review and metaregression analysis. *American Journal of Psychiatry, 164*(6), 942–948.

Posthuma, D., Baaré, W. F. C., Hulshoff Pol, H. E., Kahn, R. S., Boomsma, D. I., y De Geus, E. J. C. (2003). Genetic correlations between brain volumes and the WAIS-III dimensions of verbal comprehension, working memory, perceptual organization, and processing speed. *Twin Research, 6*(2), 131–139.

Prabhakaran, V., Smith, J. A. L., Desmond, J. E., Glover, G. H., y Gabrieli, J. D. E. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices test. *Cognitive Psychology, 33*, 43–63.

Prieto, G. y Muñiz, J. (2000). Un modelo para evaluar la calidad de los tests utilizados en Español. *Papeles del psicólogo, 77*, 65-71.

Prifitera, A., Saklofske, D. H., y Weiss, L. G. (2008). *WISC–IV clinical assessment and intervention* (2ª ed.). Amsterdam: Elsevier Academic Press.

Prigatano, G. P., Gray, J. A., y Gale, S. D. (2008). Individual case analysis of processing speed difficulties in children with and without traumatic brain injury. *The Clinical Neuropsychologist, 22*(4), 603–619. doi: 10.1080/13854040701440501

Rackley, C., Allen, D. N., Fuhrman, L. J., y Mayfield, J. (2012). Generalizability of WISC-IV index and subtest score profiles in children with traumatic brain injury. *Child Neuropsychology, 18*(5), 512–519. doi: 10.1080/09297049.2011.628308

Raghubar, K., Cirino, P., Barnes, M., Ewing-Cobbs, L., Fletcher, J., y Fuchs, L. (2009). Errors in multi-digit arithmetic and behavioral inattention in children with math difficulties. *Journal of Learning Disabilities, 42*(4), 356–371. doi: 10.1177/0022219409335211

Raiford, S. E., y Coalson, D. L. (2014). *Essentials of WPPSI™–IV assessment.* Hoboken, NJ: John Wiley & Sons.

Raiford, S. E., Weiss, L. G., Rolfhus, E., y Coalson, D. (2005). *General ability index* [WISC–IV Technical Report No. 4]. Recuperado de http://www.pearsonassessments.com/NR/rdonlyres/1439CDFE-6980-435F-93DA-05888C7CC082/0/80720_WISCIV_Hr_r4.pdf

Raizada, R. D. S., Richards, T. L., Meltzoff, A., y Kuhl, P. K. (2008). Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. *NeuroImage, 40*(3), 1392–1401. doi: 10.1016/j.neuroimage.2008.01.021

Recla, M., Bardoni, A., Galbiati, S., Pastore, V., Dominici, C., Tavano, A.,...Strazzer, S. (2013). Cognitive and adaptive functioning after severe TBI in school-aged children. *Brain Injury, 27*(7–8), 862–871. doi: 10.3109/02699052.2013.775499

Reeve, C. L., y Charles, J. E. (2008). Survey of opinions on the primacy of *g* and social consequences of ability testing: A comparison of expert and non-expert views. *Intelligence, 36*, 681–688.

Reiss, A. L., Abrams, M. T., Singer, H. S., Ross, J. L., y Denckla, M. B. (1996). Brain development, gender and IQ in children: A volumetric imaging study. *Brain, 119*, 1763–1774.

Reynolds, C. R. (1997). Forward and backward memory span should not be combined for clinical analysis. *Archives of Clinical Neuropsychology, 12*(1), 29–40.

Reynolds, C. R., y Kamphaus, R. W. (2004). *Behavior assessment system for children* (2ª ed.). Bloomington, MN: Pearson.

Reynolds, M. R., Keith, T. Z., Flanagan, D. P., y Alfonso, V. C. (2013). A cross-battery, reference variable, confirmatory factor analytic investigation of the CHC taxonomy. *Journal of School Psychology, 51*, 535–555.

Rhodes, S. M., Park, J., Seth, S., y Coghill, D. R. (2012). A comprehensive investigation of memory impairment in attention deficit hyperactivity disorder and oppositional defiant disorder. *Journal of Child Psychology and Psychiatry, 53*(2), 128–137. doi: 10.1111/j.1469-7610.2011.02436.x

Rice, D., y Barone, S., Jr. (2000). Critical periods of vulnerability for the developing nervous system: Evidence from humans and animal models. *Environmental Health Perspectives, 108*(Supplement 3), 511–533.

Rimm, S., Gilman, B., y Silverman, L. (2008). Alternative assessments with gifted and talented students. En J. L. VanTassel-Baska (Ed.), *Nontraditional applications of traditional testing* (pp. 175–202). Waco, TX: Prufrock Press.

Robertson, K. F., Smeets, S., Lubinski, D., y Benbow, C. P. (2010). Beyond the threshold hypothesis: Even among the gifted and top math/science graduate students, cognitive abilities, vocational interests, and lifestyle preferences matter for career choice, performance, and persistence. *Current Directions in Psychological Science, 19*(6), 346–351. doi: 10.1177/0963721410391442

Roid, G. H. (2003). *Stanford-Binet intelligence scales, fifth edition, technical manual.* Itasca, IL: Riverside.

Rosch, K. S., Dirlikov, B., y Mostofsky, S. H. (2013). Increased intrasubject variability in boys with ADHD across tests of motor and cognitive control. *Journal of Abnormal Child Psychology, 41*(3), 485–495.

Rowe, E. W., Kingsley, J. M., y Thompson, D. F. (2010). Predictive ability of the general ability index (GAI) versus the full scale IQ among gifted referrals. *School Psychology Quarterly, 25*(2), 119–128. doi: 10.1037/a0020148

Rowe, E. W., Miller, C., Ebenstein, L. A., y Thompson, D. F. (2012). Cognitive predictors of reading and math achievement among gifted referrals. *School Psychology Quarterly, 27*(3), 144–153. doi: 10.1037/a0029941

Rubia, K. (2011). "Cool" inferior frontostriatal dysfunction in attention-deficit/hyperactivity disorder versus "hot" ventromedial orbitofrontal-limbic dysfunction in conduct disorder: a review. *Biological Psychiatry, 69*(12), e69–e87. doi: 10.1016/j.biopsych.2010.09.023

Rudy, H. L., y Levinson, E. M. (2008). Best practices in the multidisciplinary assessment of emotional disturbances: A primer for counselors. *Journal of Counseling & Development, 86,* 494–504.

Ruggles, S., Alexander, J. T., Genadek, K., Goeken, R., Schroeder, M. B., y Sobek, M. (2010). *Integrated public use microdata series: Version 5.0* [Machine-readable database]. Minneapolis, MN: University of Minnesota.

Saarinen, S., Fontell, T., Vuontela, V., Carlson, S., y Aronen, E. T. (2014). Visuospatial working memory in 7- to 12-year-old children with disruptive behavior disorders. *Child Psychiatry & Human Development.* doi: 10.1007/s10578-014-0449-3

Salthouse, T. A. (2004). Localizing age-related individual differences in a hierarchical structure. *Intelligence, 32,* 541–561.

Sanders, S., McIntosh, D. E., Dunham, M., Rothlisberg, B. A., y Finch, H. (2007). Joint confirmatory factor analysis of the Differential Ability Scales and the Woodcock-Johnson Tests of Cognitive Abilities–Third Edition. *Psychology in the Schools, 44*(2), 119–138. doi: 10.1002/pits.20211

Sattler, J. M. (2008a). *Assessment of children: Cognitive foundations* (5ª ed.). San Diego, CA: Author.

Sattler, J. M. (2008b). *Resource guide to accompany assessment of children: Cognitive foundations* (5ª ed.). San Diego, CA: Author.

Sattler, J. M., y Ryan, J. J. (2009). *Assessment with the WAIS–IV.* La Mesa, CA: Author.

Scherf, K. S., Sweeney, J. A., y Luna, B. (2006). Brain basis of developmental change in visuospatial working memory. *Journal of Cognitive Neuroscience, 18*(7), 1045–1058.

Schinka, J. A., Vanderploeg, R. D., y Curtiss, G. (1997). WISC–III subtest scatter as a function of highest subtest scaled score. *Psychological Assessment, 9*(2), 83–88.

Schmidt, E. A., Homeyer, L. E., y Walker, J. L. (2009). Counselor preparation: Predictors of success on the counselor preparation comprehensive examination. *Counselor Education & Supervision, 48,* 226–238.

Schmidt, F. L. (2014). A general theoretical integrative model of individual differences in interests, abilities, personality traits, and academic and occupational achievement: A commentary on four recent articles. *Perspectives on Psychological Science, 9*(2), 211–218. doi: 10.1177/1745691613518074

Schmithorst, V. J. (2009). Developmental sex differences in the relation of neuroanatomical connectivity to intelligence. *Intelligence, 37*(2), 164–173. doi: 10.1016/j.intell.2008.07.001

Schmithorst, V. J., y Holland, S. K. (2006). Functional MRI evidence for disparate developmental processes underlying intelligence in boys and girls. *NeuroImage, 31,* 1366–1379. doi: 10.1016/j.neuroimage.2006.01.010

Schmithorst, V. J., Holland, S. K., y Dardzinski, B. J. (2008). Developmental differences in white matter architecture between boys and girls. *Human Brain Mapping, 29*(6), 696–710. doi: 10.1002/hbm.20431

Schmitt, J. E., Lenroot, R. K., Wallace, G. L., Ordaz, S., Taylor, K. N., Kabani, N.,...Giedd, J. N. (2008). Identification of genetically mediated cortical networks: A multivariate study of pediatric twins and siblings. *Cerebral Cortex, 18*(8), 1737–1747. doi: 10.1093/cercor/bhm211

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment, 29*(4), 304–321. doi: 10.1177/0734282911406653

Schneider, W., Schumann-Hengsteler, R., y Sodian, B. (Eds.). (2005). *Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind.* Mahwah, NJ: Lawrence Erlbaum Associates.

Schneider, W. J., y McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed., pp. 99–144). Nueva York, NY: Guilford Press.

Schoemaker, K., Bunte, T., Wiebe, S. A., Espy, K. A., Deković, M., y Matthys, W. (2012). Executive function deficits in preschool children with ADHD and DBD. *Journal of Child Psychology and Psychiatry, 53*(2), 111–119. doi: 10.1111/j.1469-7610.2011.02468.x

Schuchardt, K., Bockmann, A.-K., Bornemann, G., y Maehler, C. (2013). Working memory functioning in children with learning disorders and specific language impairment. *Topics in Language Disorders, 33*(4), 298–312. doi: 10.1097/01.TLD.0000437943.41140.36

Schumacker, R., y Lomax, R. (2004). *A beginner's guide to structural equation modeling* (2ª ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464. doi: 10.1214/aos/1176344136

Séguin, J. R., Parent, S., Tremblay, R. E., y Zelazo, P. D. (2009). Different neurocognitive functions regulating physical aggression and hyperactivity in early childhood. *Journal of Child Psychology and Psychiatry, 50*(6), 679–687. doi: 10.1111/j.1469-7610.2008.02030.x

Seidler, J., Vondráček, J., y Saxl, I. (2000). The life and work of Zbyněk Šidák (1933–1999). *Applications of Mathematics, 45*(5), 321–336.

Shanahan, M. A., Pennington, B. F., Yerys, B. E., Scott, A., Boada, R., Willcutt, E. G.,...DeFries, J. C. (2006). Processing speed deficits in attention deficit/hyperactivity disorder and reading disability. *Journal of Abnormal Child Psychology, 34*, 585–602. doi: 10.1007/s10802-006-9037-8

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N.,...Giedd, J. (2006). Intellectual ability and cortical development in children and adolescents. *Nature, 440*, 676–679. doi: 10.1038/nature04513

Shrout, P., y Fleiss, J. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420–428.

Silver, N. C., y Dunlap, W. P. (1987). Averaging correlation coefficients: Should Fisher's *z* transformation be used? *Journal of Applied Psychology, 72*(1), 146–148.

Skogan, A. H., Zeiner, P., Egeland, J., Rohrer-Baumgartner, N., Urnes, A.-G., Reichborn-Kjennerud, T., y Aase, H. (2013). Inhibition and working memory in young preschool children with symptoms of ADHD and/or oppositional-defiant disorder. *Child Neuropsychology, 20*(5), 607–624.

Soulières, I., Dawson, M., Gernsbacher, M. A., y Mottron, L. (2011). The level and nature of autistic intelligence II: What about Asperger syndrome? *PloS ONE, 6*(9), e25372. doi: 10.1371/journal.pone.0025372

Sparrow, S. S., Cicchetti, D. V., y Balla, D. A. (2005). *Vineland adaptive behavior scales* (2ª ed.). Bloomington, MN: Pearson.

Spearman, C. (1904). "General intelligence:" Objectively determined and measured. *American Journal of Psychology, 15*, 201–293.

Spruill, J. (1998). Assessment of mental retardation with the WISC–III. En A. Prifitera y D. H. Saklofske (Eds.), *WISC–III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 73–91). San Diego, CA: Academic Press.

Stanley, J. C. (1971). Reliability. En R. L. Thorndike (Ed.), *Educational measurement* (2ª ed., pp. 356–442). Washington, DC: American Council on Education.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*(2), 173–180.

Sternberg, R. J. (Ed.). (2000). *Handbook of intelligence.* Nueva York, NY: Cambridge University Press.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences.* Mahwah, NJ: Erlbaum.

Stevenson, J. L. (2011). *Autistic cognition: Effects of test domain and reasoning level* (Doctoral dissertation). Recuperado de ProQuest Dissertations and Theses. (UMI No. 3486771)

Strube, M. J. (1988). Some comments of the use of magnitude-of-effect estimates. *Journal of Counseling Psychology, 35*(3), 342–345.

Sweetland, J. D., Reina, J. M., y Tatti, A. F. (2006). WISC–III verbal/performance discrepancies among a sample of gifted children. *Gifted Child Quarterly, 50*(1), 7–10.

Szmalec, A., Verbruggen, F., Vandierendonck, A., y Kemps, E. (2011). Control of interference during working memory updating. *Journal of Experimental Psychology: Human Perception and Performance, 37*(1), 137–151. doi: 10.1037/a0020365

Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y., Komuro, N.,...Kawashima, R. (2010). Training of working memory impacts structural connectivity. *The Journal of Neuroscience, 30*(9), 3297–3303. doi: 10.1523/JNEUROSCI.4611-09.2010

Takeuchi, H., Taki, Y., Hashizume, H., Sassa, Y., Nagase, T., Nouchi, R., y Kawashima, R. (2011). Effects of training of processing speed on neural systems. *The Journal of Neuroscience, 31*(34), 12139–12148. doi: 10.1523/JNEUROSCI.2948-11.2011

Terman, L. M. (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon Intelligence Scale.* Oxford, England: Houghton Mifflin.

Thaler, N. S., Bello, D. T., y Etcoff, L. M. (2012). WISC-IV profiles are associated with differences in symptomatology and outcome in children with ADHD. *Journal of Attention Disorders, 17*(4), 291–301. doi: 10.1177/1087054711428806

Theiling, J., y Petermann, F. (2014). Neuropsychological profiles on the WAIS–IV of ADHD adults. *Journal of Attention Disorders.* Advance online publication. doi: 10.1177/1087054713518241

Thompson, B. (2000). Ten commandments of structural equation modeling. En L. G. Grimm y P. R. Yarnold (Eds.), *Reading and understanding MORE multivariate statistics* (pp. 261–283). Washington, DC: American Psychological Association.

Thorndike, R. L. (1936). Factor analysis of social and abstract intelligence. *Journal of Educational Psychology, 27*(3), 231–233.

Till, C., Ghassemi, R., Aubert-Broche, B., Kerbrat, A., Collins, D. L., Narayanan, S.,...Banwell, B. L. (2011). MRI correlates of cognitive impairment in childhood-onset multiple sclerosis. *Neuropsychology, 25*(3), 319–332. doi: 10.1037/a0022051

Tolar, T. D., Fuchs, L., Cirino, P. T., Fuchs, D., Hamlett, C. L., y Fletcher, J. M. (2012). Predicting development of mathematical word problem solving across the intermediate grades. *Journal of Educational Psychology, 104*(4), 1083–1093. doi: 10.1037/a0029020

Towse, J., y Cowan, N. (2005). Working memory and its relevance for cognitive development. En W. Schneider, R. Schumann-Hengsteler, y B. Sodian (Eds.), *Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind* (pp. 9–37). Mahwah, NJ: Lawrence Erlbaum Associates.

Travers, B. G., Bigler, E. D., Tromp, D. P. M., Adluru, N., Froehlich, A. L., Ennis, C.,...Lainhart, J. E. (2014). Longitudinal processing speed impairments in males with autism and the effects of white matter microstructure. *Neuropsychologia 53,* 137–145. doi: 10.1016/j.neuropsychologia.2013.11.008

Tremont, G., Mittenberg, W., y Miller, L. J. (1999). Acute intellectual effects of pediatric head trauma. *Child Neuropsychology, 5,* 104–114.

Tucker, L. R., y Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*(1), 1–10.

Turken, A. U., Whitfield-Gabrieli, S., Bammer, R., Baldo, J., Dronkers, N. F., y Gabrieli, J. D. E. (2008). Cognitive processing speed and the structure of white matter pathways: Convergent evidence from normal variation and lesion studies. *NeuroImage, 42*(2), 1032–1044. doi: 10.1016/j.neuroimage.2008.03.057

Valdés, A. A., Vera, J. A., y Carlos, E. A. (2013). Variables que diferencian a estudiantes de bachillerato con y sin aptitudes intelectuales sobresalientes. Revista Electrónica de Investigación Educativa, 15(3), 85–97. Recuperado de http://redie.uabc.mx/vol15no3/contenido-valdesverac.html

Van der Molen, M. J., Henry, L. A., y Van Luit, J. E. H. (2014). Working memory development in children with mild to borderline intellectual disabilities. *Journal of Intellectual Disability Research, 58*(7), 637–650. doi: 10.1111/jir.12061

Van der Molen, M. J., Van Luit, J. E. H., Van der Molen, M. W., y Jongmans, M. J. (2010). Everyday memory and working memory in adolescents with mild intellectual disability. *American Journal of Intellectual and Developmental Disabilities, 115*(3), 207–217. doi: 10.1352/1944-7558-115.3.207

van Lier, P. A. C., van der Ende, J., Koot, H. M., y Verhulst, F. C. (2007). Which better predicts conduct problems? The relationship of trajectories of conduct problems with ODD and ADHD symptoms from childhood into adolescence. *Journal of Child Psychology and Psychiatry, 48*(6), 601–608. doi: 10.1111/j.1469-7610.2006.01724.x

Vugs, B., Cuperus, J., Hendriks, M., y Verhoeven, L. (2013). Visuospatial working memory in specific language impairment: A meta-analysis. *Research in Developmental Disabilities, 34,* 2586–2597. doi: 10.1016/j.ridd.2013.05.014

Vukovic, R. K., Fuchs, L. S., Geary, D. C., Jordan, N. C., Gersten, R., y Siegler, R. S. (2014). Sources of individual differences in children's understanding of fractions. *Child Development, 85*(4), 1461–1476. doi: 10.1111/cdev.12218

Wakkinen, H. B. (2008). *Maximizing resources to gain information about clients: Profile analysis, configural frequency analysis, and the WISC–IV.* Recuperado de ProQuest Dissertations and Theses. (UMI No. 3322469)

Wang, H., Jing, J., Jin, Y., Liu, B.-Y., y Wang, X. (2013). Relation of intelligence structural to clinical symptoms in children with Asperger syndrome. *Chinese Mental Health Journal, 27*(6), 440–444.

Wang, S., y Gathercole, S. E. (2013). Working memory deficits in children with reading difficulties: Memory span and dual task coordination. *Journal of Experimental Child Psychology, 115*(1), 188–197. doi: 10.1016/j.jecp.2012.11.015

Ward, L. C., Bergman, M. A., y Hebert, K. R. (2011). WAIS–IV subtest covariance structure: Conceptual and statistical considerations. *Psychological Assessment, 24*(2), 328–340. doi: 10.1037/a0025614

Ward, L. C., Ryan, J. J., y Axelrod, B. N. (2000). Confirmatory factor analyses of the WAIS–III standardization data. *Psychological Assessment, 12*(3), 341–345.

Wasserman, J. D. (2012). A history of intelligence assessment: The unfinished tapestry. En D. P. Flanagan y P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3ª ed., pp. 3–55). Nueva York, NY: Guilford Press.

Watkins, M. W. (2010). Structure of the Wechsler Intelligence Scale for Children–Fourth Edition among a national sample of referred students. *Psychological Assessment, 22*(4), 782–787.

Watkins, M. W., Wilson, S. M., Kotz, K. M., Carbone, M. C., y Babula, T. (2006). Factor structure of the Wechsler Intelligence Scale for Children—Fourth Edition among referred students. *Educational and Psychological Measurement, 66,* 975–983. doi: 10.1177/0013164406288168

Wechsler, D. (1939). *Wechsler-Bellevue intelligence scale.* Nueva York, NY: The Psychological Corporation.

Wechsler, D. (1944). *The measurement of adult intelligence* (3ª ed.). Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1945). A standardized memory scale for clinical use. *The Journal of Psychology, 19,* 87–95.

Wechsler, D. (1949). *Wechsler intelligence scale for children.* Nueva York, NY: The Psychological Corporation.

Wechsler, D. (1955). *Wechsler adult intelligence scale.* Nueva York, NY: The Psychological Corporation.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4ª ed.). Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1967). *Wechsler preschool and primary scale of intelligence.* Nueva York, NY: The Psychological Corporation.

Wechsler, D. (1974a). *Manual for the Wechsler intelligence scale for children–revised.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1974b). *Escala de inteligencia de Wechsler para niños.* Madrid: TEA Ediciones.

Wechsler, D. (1975). Intelligence defined and undefined: A relativistic appraisal. *American Psychologist, 30,* 135–139.

Wechsler, D. (1991). *Wechsler intelligence scale for children* (3ª ed.). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1993). *Escala de inteligencia de Wechsler para niños-Revisada.* Madrid: TEA Ediciones.

Wechsler, D. (1997). *Wechsler adult intelligence scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2002). *Wechsler preschool and primary scale of intelligence* (3ª ed.). San Antonio, TX: Pearson.

Wechsler, D. (2003). *Wechsler intelligence scale for children* (4ª ed.). San Antonio, TX: Pearson.

Wechsler, D. (2008). *Wechsler adult intelligence scale* (4ª ed.). Bloomington, MN: Pearson.

Wechsler, D. (2009a). *Wechsler memory scale* (4ª ed.). Bloomington, MN: Pearson.

Wechsler, D. (2009b). *Escala de inteligencia de Wechsler para preescolar y primaria-III.* Madrid: TEA Ediciones.

Wechsler, D. (2010 [ed. orig. TEA ediciones, 1999]). *Escala de inteligencia de Wechsler para adultos-III.* Madrid: TEA Ediciones.

Wechsler, D. (2011 [ed. orig. TEA, 2005]). *Escala de inteligencia de Wechsler para niños-IV.* Madrid: Pearson Educación.

Wechsler, D. (2012a). *Wechsler preschool and primary scale of intelligence* (4ª ed.). Bloomington, MN: Pearson.

Wechsler, D. (2012b). *Escala de inteligencia de Wechsler para adultos-IV.* Madrid: Pearson Educación.

Wechsler, D. (2013). *Escala de memoria de Wechsler-IV.* Madrid: Pearson Educación.

Wechsler, D. (2014a). *Wechsler intelligence scale for children* (5ª ed.). San Antonio, TX: Pearson.

Wechsler, D. (2014b). *Escala de inteligencia de Wechsler para preescolar y primaria-IV.* Madrid: Pearson Educación.

Weiss, L. G., Chen, H., Harris, J. G., Holdnack, J. A., y Saklofske, D. H. (2010). WAIS–IV use in societal context. En L. G. Weiss, D. H. Saklofske, D. L. Coalson, y S. E. Raiford (Eds.), *WAIS–IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 97–139). Amsterdam: Elsevier Academic Press.

Weiss, L. G., y Gabel, A. D. (2008). *Using the cognitive proficiency index in psychoeducational assessment* [WISC–IV Technical Report No. 6]. Recuperado de http://www.pearsonclinical.com.au/userfiles/WISCIVTechReport6_CPI.pdf

Weiss, L. G., Keith, T. Z., Zhu, J., y Chen, H. (2013a). WAIS-IV clinical validation of the four- and five-factor interpretive approaches [Special edition]. *Journal of Psychoeducational Assessment, 31*(2), 94–113. doi: 10.1177/0734282913478030

Weiss, L. G., Keith, T. Z., Zhu, J., y Chen, H. (2013b). WISC-IV and clinical validation of the four- and five-factor interpretive approaches [Special edition]. *Journal of Psychoeducational Assessment, 31*(2), 114–131. doi: 10.1177/0734282913478032

Weiss, L. G., Saklofske, D. H., Prifitera, A., y Holdnack, J. A. (2006). *WISC–IV advanced clinical interpretation.* Amsterdam: Elsevier Academic Press.

Wendelken, C., Nakhabenko, D., Donohue, S. E., Carter, C. S., y Bunge, S. A. (2008). "Brain is to thought as stomach is to ??": Investigating the role of rostrolateral prefrontal cortex in relational reasoning. *Journal of Cognitive Neuroscience, 20*(4), 682–693. doi: 10.1162/jocn.2008.20055

Werheid, K., Hoppe, C., Thöne, A., Müller, U., Müngersdorf, M., y von Cramon, D. Y. (2002). The adaptive digit ordering test: Clinical application, reliability, and validity of a verbal working memory test. *Archives of Clinical Neuropsychology, 17,* 547–565.

Wilde, E. A., McCauley, S. R., Chu, Z., Hunter, J. V., Bigler, E. D., Yallampalli, R.,...Levin, H. S. (2009). Diffusion tensor imaging of hemispheric asymmetries in the developing brain. *Journal of Clinical and Experimental Neuropsychology, 31*(2), 205–218. doi: 10.1080/13803390802098118

Wilkins, C., Rolfhus, E., Weiss, L., y Zhu, J. J. (2005, April). *A new method for calibrating translated tests with small sample sizes.* Paper presented at the 2005 annual meeting of the American Educational Research Association, Montreal, Canadá.

Willcutt, E. G., Petrill, S. A., Wu, S., Boada, R., DeFries, J. C., Olson, R. K., y Pennington, B. F. (2013). Comorbidity between reading disability and math disability: Concurrent psychopathology, functional impairment, and neuropsychological functioning. *Journal of Learning Disabilities, 46*(6), 500–516. doi: 10.1177/0022219413477476

Willerman, L., Schultz, R., Rutledge, J. N., y Bigler, E. D. (1991). In vivo brain size and intelligence. *Intelligence, 15,* 223–228.

Wrulich, M., Brunner, M., Stadler, G., Schalke, D., Keller, U., y Martin, R. (2014). Forty years on: Childhood intelligence predicts health in middle adulthood. *Health Psychology, 33*(3), 292–296. doi: 10.1037/a0030727

Zayat, M., Kalb, L., y Wodka, E. L. (2011). Brief report: Performance pattern differences between children with autism spectrum disorders and attention deficit-hyperactivity disorder on measures of verbal intelligence. *Journal of Autism and Developmental Disorders, 41*(12), 1743–1747. doi: 10.1007/s10803-011-1207-z

Zheng, X., Flynn, L. J., y Swanson, H. L. (2013). Experimental intervention studies on word problem solving and math disabilities: A selective analysis of the literature. *Learning Disability Quarterly, 36*(2), 97–111. doi: 10.1177/0731948712444277

Zhu, J., y Chen, H. (2011). Utility of inferential norming with smaller sample sizes. *Journal of Psychoeducational Assessment, 29*(6), 570–580.

Zhu, J., y Chen, H. (2013). Clinical utility of Cancellation on the WISC-IV. *Journal of Psychoeducational Assessment, 31*(6), 527–537. doi: 10.1177/0734282913480865

Zieman, S. F. X., Jr. (2010). *Performance analysis on the WISC–IV working memory and processing speed index among ADHD subtypes.* (Tesis doctoral). Recuperado de WorldCat. (Acceso núm. 526695555)

WISC-V

ESCALA DE INTELIGENCIA
DE WECHSLER PARA NIÑOS-V

ISBN 978-84-9035-302-8

PEARSON PsychCorp Pearson Clinical & Talent Assessment www.pearsonclinical.es