

Uso del WISC-III en Chile: análisis de fortalezas y limitaciones según la percepción de los psicólogos

Use of WISC III in Chile: Analysis of strengths and limitations according to the perception of psychologists

Marcela Alejandra Rodríguez Cancino & María Beatriz Vizcarra Larrañaga

Universidad de La Frontera, Temuco, Chile

El estudio buscó conocer la percepción de los psicólogos chilenos acerca de las fortalezas y limitaciones de la versión chilena de la Escala de Inteligencia de Wechsler para Niños (WISC-III). Se administró una encuesta anónima online a 160 psicólogos, de todo el país, aunque predominantemente de la IX región. Entre las principales fortalezas de la escala se destacan: (i) que entrega un perfil completo de habilidades cognitivas; y (ii) utiliza material atractivo y dinámico para los niños. Entre las limitaciones señalan: presencia de ítems desactualizados; y posibles sesgos en función de nivel socioeconómico y/o procedencia rural de los niños. Se encontraron relaciones estadísticamente significativas entre la acreditación en el uso de la escala y la completitud de los informes. Los resultados se discuten en torno a los desafíos de la formación en evaluación psicológica en Chile, la acreditación en el uso de test, y las buenas prácticas en evaluación psicológica.

Palabras clave: WISC III, formación en evaluación psicológica, buenas prácticas en evaluación psicológica.

The purpose of the present study was to explore psychologist's perceptions about the strengths, weaknesses, and challenges when using the Wechsler Intelligence Scale for Children (WISC-III, Chilean version). An online survey was conducted with the voluntary participation of 160 psychologists from all over the country, though most of them were from the Araucanía region. Participants highlight as strengths of the instrument the fact that it provides a complete profile of cognitive skills, and its attractiveness for children. Among the limitations, they pointed to the presence of outdated items, and to potential biases as a function of socioeconomic level and/or the rural origin of children. Results are discussed in terms of the challenges posed by adequate training in intellectual assessment, the need for accreditation in the use of tests, and best practices in psychological assessment, in general.

Keywords: WISC-III, training in psychological assessment, best practices in psychological assessment.

Este estudio se realizó para recoger datos sobre el uso de la escala WISC III, previamente a la implementación de un Proyecto de Investigación en la Universidad de La Frontera, código DIUFRO INI17-006, titulado "Exploración propiedades psicométricas de la Escala de Inteligencia para niños (WISC V), en una muestra de escolares rurales de la Araucanía". Contamos con certificado de aprobación de comité de ética de este proyecto DIUFRO.

Contacto: M. A. Rodríguez Cancino. Departamento de Psicología, Universidad de La Frontera, Calle Montevideo 0830, Temuco, Chile. Correo electrónico: marcela.rodriguez@ufrontera.cl

Cómo citar: Rodríguez Cancino, M. A. & Vizcarra Larrañaga, M. B. (2018). Uso del WISC-III en Chile: análisis de fortalezas y limitaciones según la percepción de los psicólogos. *Revista de Psicología*, 27(2), 1-13.
<http://dx.doi.org/10.5354/0719-0581.2019.52316>

Introducción

Los test son una herramienta básica del quehacer del psicólogo y su uso abarca los distintos ámbitos del ejercicio profesional: educativo, clínico, organizacional y jurídico (Hernández, Ponsoda, Muñiz, Prieto, & Elosua, 2016). Los test, en conjunto con otras fuentes de información, cumplen la función de orientar al profesional, en el proceso de toma de decisiones, el cual tiene un impacto significativo en la vida de las personas y las organizaciones (Elosua & Muñiz, 2013; Fonseca-Pedrero & Muñiz, 2017).

La correcta utilización de los test requiere que estos cuenten con adecuadas propiedades psicométricas, y que sean administrados por profesionales competentes y calificados para su uso, solo así podrán configurarse en herramientas fiables de apoyo al quehacer profesional (Elosua, 2017; Fonseca-Pedrero & Muñiz, 2017).

Las orientaciones emanadas desde organismos internacionales, como la International Test Commission (ITC) sobre evaluación psicológica y uso de test, ofrecen claras directrices en torno a las buenas prácticas y su principal objetivo es mejorar el uso de las pruebas (ITC, 2001, 2014). Dentro de estas orientaciones se destaca la importancia del uso ético de los test. Esto implica (a) resguardar el respeto por los derechos y la dignidad de los evaluados; (b) la competencia y formación de los evaluadores; (c) el cuidado de los materiales del test y de sus resultados; y (d) una práctica rigurosa, vale decir, la elección cuidadosa de las pruebas y la revisión de sus propiedades psicométricas, una administración, puntuación, análisis de resultados y comunicación de los resultados, adecuadas, así como el análisis de posibles sesgos, entre otros (ITC, 2001; 2014; Muñiz, Hernández, & Ponsoda, 2015).

Las directrices sobre las buenas prácticas en evaluación psicológica también sugieren la necesidad de evaluar constantemente la validez de los instrumentos en diferentes grupos, con el fin de reconocer su pertinencia y su aplicabilidad, descartando la presencia de sesgos hacia algún grupo en particular (ITC 2001, 2014). Asimismo, las orientaciones para la intervención deben basarse en datos o resultados de test con baremos actualizados en los últimos diez años y adaptados a la región en la cual se están aplicando (Asociación Argentina de Estudio e Investigación en

Psicodiagnóstico, 1999; American Psychological Association [APA], 2010).

La manera en que se aplica un test psicológico también puede impactar sobre sus propiedades psicométricas y, por ende, en su nivel de utilidad. En este sentido, es fundamental que el psicólogo sea riguroso en la forma de entregar las instrucciones, tenga experiencia en la aplicación de instrumentos y sea capaz de reconocer el valor de seguir exhaustivamente los procedimientos de administración estandarizados, así como también cuidar el establecimiento de una relación o rapport apropiado (Muñiz, Elosua, & Hambleton, 2013).

Para que un psicólogo pueda desarrollar su trabajo en el marco de las buenas prácticas, es recomendable que cuente con información relativa a las fortalezas y debilidades de los test, y la pertinencia de estos a la población en la que cumple sus funciones (APA, 2000).

En Chile, tal como evidencian Vinet y González R. (2013), existe una escasa reglamentación en torno al desempeño profesional del psicólogo, así como también sobre la pertinencia de los instrumentos de evaluación. La carencia de instrumentos psicológicos adecuados para la población local es un problema crítico, según las autoras, y se relaciona, por una parte, con la ausencia de una industria editorial dedicada al desarrollo y/o adaptación de instrumentos, y, por otra, con la falta de recursos públicos destinados a la investigación psicométrica. Un desafío importante para la psicología chilena, junto con disponer de mayores recursos para la elaboración y adaptación de instrumentos, es contar con instituciones o grupos de expertos que evalúen los instrumentos de uso masivo, de acuerdo con estándares internacionales tal como existen en Brasil, Estados Unidos y España (American Educational Research Association, APA, & National Council on Measurement in Education, 2014; Consejo General de la Psicología, 2017; Wechsler, 2013).

Este último punto es particularmente relevante en la evaluación del funcionamiento cognitivo infantil, ya que este es un punto de partida para la toma de decisiones que afectan la trayectoria académica de niños y niñas. La evaluación del funcionamiento cognitivo se refiere, específicamente, a la apreciación de la actividad cerebral en áreas tales como la memoria, concentración, velocidad de procesamiento,

lenguaje y capacidades de razonamiento. Este tipo de evaluación permite identificar a los estudiantes con talento académico, contribuir al diagnóstico de problemas de aprendizaje y discapacidad intelectual, y detectar alteraciones incipientes en el desarrollo cognitivo. Lo anterior resulta crítico si se considera que el déficit en el despliegue de algunas de las funciones cognitivas en los primeros años de escolaridad puede tener un gran impacto transcurridos varios años después, concretamente en problemas de aprendizaje derivados de dislexias, discalculias y disgrafías (Urzúa, Ramos, Alday, & Alquinta, 2010).

Para la evaluación del funcionamiento cognitivo a nivel internacional actualmente existen diversos instrumentos entre los que se destacan la Batería Kaufman (K-ABC); la Escala Stanford-Binet, cuarta edición; la Batería Woodcock-Johnson Cognitive Battery, cuarta edición (W-J IV); la Escala de Habilidad Diferencial, segunda edición (DAS-II); y el Sistema de Evaluación Cognitiva (CAS), las cuales concentran el mayor volumen de publicaciones científicas (Elliot, 2007; Kaufman & Kaufman, 1983; Laurent, Swerdlik, & Ryburn, 1992; Naglieri, 1999; Schrank, McGrew, & Mather, 2014).

Sin embargo, son las Escalas de Inteligencia de Wechsler, los instrumentos más usados, a nivel mundial, para la evaluación intelectual de niños y adultos, en los distintos contextos de desempeño profesional (clínico, educacional, laboral, etc.). Su origen se remonta al año 1939, en que el psicólogo clínico David Wechsler, del Hospital Bellevue, desarrolló la escala Wechsler-Bellevu (1939) para medir desempeño intelectual en adultos, basado en la concepción de la inteligencia como “la habilidad agregada o global de un individuo para pensar racionalmente, actuar con propósito y adaptarse al medio” (Matarazzo, 1972, p. 79). Hasta esa fecha, la mayoría de los test para este segmento etario solo consistían en adaptaciones de escalas infantiles, con escasa validez aparente. Desde entonces, se han desarrollado tres escalas derivadas de la original: la Escala de Inteligencia de Wechsler para Adultos (WAIS), que cuenta con cuatro revisiones: WAIS, 1955; WAIS- R, 1981; WAIS-III, 1997; WAIS-IV, 2008 (Weiss, Saklofske, Coalson, & Raiford, 2010); la Escala de Inteligencia de Wechsler para Niños en Edad Escolar (WISC), que cuenta con cinco revisiones; y la Escala para Preescolares (WPPSI), creada en el año 1967, y

que tras sucesivas revisiones (WPPSI-R, WPPSI III) publicó su cuarta edición (WPPSI-IV) en 2012. Esta última explora el desarrollo cognitivo en niños de entre dos años seis meses y siete años de edad (Watkins & Beaujean, 2014).

Si bien la Escala de Inteligencia de Wechsler para niños WISC, fue publicada en 1949, recién en el año 1974 se realizó la adaptación y estandarización de una nueva versión, denominada WISC-R. Esta versión revisada conservó esencialmente la estructura original de dos escalas, una verbal y otra manipulativa. En esta estandarización se modificaron o eliminaron ítems ambiguos, obsoletos, inadecuados o poco discriminativos. En 1991, el autor y sus colaboradores realizaron una segunda revisión y actualización, dando origen a la Escala WISC-III. Esta escala, si bien mantuvo las características fundamentales de WISC-R, contaba con materiales más atractivos, y procedimientos de administración actualizados, que incluyeron un mayor énfasis en la velocidad de procesamiento junto con datos normativos actualizados (Wechsler, 1991). En sus distintas versiones, el test ha sido adaptado en países tales como Francia, Inglaterra, Suecia, Colombia, Eslovenia, Japón y Vietnam, entre otros (Georgas, Weiss, Van de Vijver, & Saklofske, 2003; Gregoire, 2001; Decker, Hale, & Flanagan, 2013; Sotelo-Dynega & Dixon, 2014).

En Chile, la Escala WISC fue adaptada en 1962, la versión revisada WISC-R fue adaptada y estandarizada entre 1975 y 1980 con una muestra de niños de la Región Metropolitana, convirtiéndose en la prueba más usada para la evaluación cognitiva de niños y adolescentes por más de tres décadas. Dada su obsolescencia, en el año 2007 el Equipo del Centro de Desarrollo de Tecnologías de Inclusión, de la Pontificia Universidad Católica de Chile (Cedeti-UC) inició el proceso de adaptación y estandarización de la Escala WISC-III en una muestra de 1.924 niños de entre 6 y 16 años de edad, de ambos sexos, seleccionados aleatoriamente con base en un muestreo estratificado por regiones. Las características del estudio, tanto respecto del tamaño muestral como la selección de la muestra, constituyó un importante avance respecto a procesos previos de adaptación y/o estandarización (Ramírez & Rosas, 2007). En 2003 aparece la cuarta versión (WISC -IV) en Estados Unidos, la cual nunca fue adaptada y

estandarizada para Chile, y finalmente en 2014 aparece la última versión el WISC V, la cual fue recientemente publicada en nuestro país, luego de un exhaustivo proceso de adaptación y estandarización (Rascón, 2018).

Si bien el uso de esta escala como procedimiento estandarizado permite el acceso a una gran cantidad de información sobre el funcionamiento cognitivo, la interpretación de sus resultados puede nutrirse cualitativamente con el uso de procedimientos específicos tales como el registro de la observación de la conducta y actitud el niño ante el test, y el examen de límites. El primero se refiere al registro de la conducta no verbal, lenguaje, atención, estado de ánimo, conducta motora y apariencia personal del evaluado. Por su parte, el examen de límites, se refiere a la readministración de las subpruebas, entregando al niño/a una serie de ayudas adicionales, o más tiempo para la ejecución de la tarea, obviando las instrucciones regulares. Este procedimiento se realiza al finalizar la aplicación estandarizada del instrumento y sus resultados en ningún caso sustituyen los obtenidos en la aplicación original, sino que se utilizan como un insumo cualitativo sobre las estrategias de resolución de problemas utilizadas por el evaluado. Ambos procedimientos proveen al psicólogo de información cualitativa altamente relevante para la determinación de un diagnóstico en el área clínica y/o educacional (Sattler, 2010).

A pesar de los numerosos estudios a nivel internacional respecto del funcionamiento psicométrico del WISC-III (Cianci, Orsini, Hulbert, & Pezuti, 2013; Gregoire, 2001, Rotzika et al., 2009), y su uso por los profesionales, en Chile no existen estudios en relación con su pertinencia cultural, la percepción de los profesionales usuarios, o sus prácticas evaluativas con este instrumento. Esto es espe-

cialmente relevante si se considera que, actualmente, en el país se está incorporando una nueva versión de la escala, por lo que acceder a las limitaciones e insuficiencias respecto del WISC-III podría orientar a los profesionales en un mejor uso de la nueva versión. Por otra parte, el impacto de estudios en esta área no solo podría beneficiar las prácticas evaluativas del psicólogo en el ámbito profesional, sino que también podría ser un elemento a considerar en la formación de los psicólogos, a nivel de pre y posgrado. Por tanto, y considerando los antecedentes expuestos, el presente estudio tuvo como principal objetivo explorar el uso que los profesionales psicólogos hacen de la Escala WISC-III en Chile, así como la apreciación de estos, respecto de las fortalezas y limitaciones de este instrumento en la evaluación del funcionamiento cognitivo de niños, niñas y adolescentes. Este estudio pretende, además, ser un aporte a la discusión sobre las buenas prácticas del uso de test, la evaluación psicológica en Chile y la importancia de la formación y especialización en esta área.

Método

Diseño

Se utilizó un diseño descriptivo, correlacional, de tipo transversal de alcance exploratorio.

Participantes

La muestra fue de tipo no probabilística de autoselección, compuesta por 160 psicólogas y psicólogos. El 76% de los participantes fueron mujeres. Al momento de responder la encuesta, la mayoría (69,5%) se desempeñaban laboralmente en regiones del centro y sur del país, 44% trabajaba en la Región de la Araucanía y 18% en la Región Metropolitana (ver tabla 1).

Tabla 1
Características de la muestra

Variables	N	%
Hombre	38	23,8
Mujer	122	76,3
Grado académico		
Licenciado	128	77
Magíster	32	20
Doctorado	3	2
Posdoctorado	2	1,3
Universidad de egreso		
Consejo de rectores (CRUCH)	76	47
Privadas	84	53
Tiempo de egreso		
5 años o menos	99	62
6 a 11 años	59	36,8
Más de 11	2	1,2
Procedencia		
Zona Centro Norte (Regiones de Tarapacá, Coquimbo, Valparaíso)	20	12,5
Región Metropolitana	29	18
Zona Centro Sur (regiones de O'Higgins, Maule, Biobío, La Araucanía, Los Ríos, Los Lagos, Aysén)	111	69,5

Instrumento

Para conocer la percepción de los psicólogos chilenos acerca de las fortalezas y limitaciones de la versión chilena de la Escala de Wechsler para Niños (WISC-III) y de sus prácticas como usuarios de la misma, se creó una encuesta ad hoc. El diseño del instrumento consideró tres fases. En una primera fase se invitó a participar en un grupo focal a los coordinadores de Proyectos de Integración Escolar (PIE) de la Araucanía, y a los psicólogos que se desempeñaban en dichos proyectos de integración. En este grupo focal se abordó su experiencia con el uso de WISC-III, y las ventajas y desventajas y/o limitaciones que observaban en su utilización dentro de PIE. En una segunda fase, y considerando los datos de la primera, se elaboró una primera versión del instrumento, la que fue revisada por jueces expertos, quienes analizaron el contenido y formato. Por último, en la tercera fase se incorporaron las sugerencias de los jueces y se creó la versión definitiva del instrumento.

La encuesta estaba compuesta de tres secciones. La primera sección indaga respecto de las características sociodemográficas de los participantes, tales como sexo, universidad de titulación y año de egreso, grado académico, acreditación en el uso de

la escala e institución en que obtuvo la certificación. Además, se consultó si ha trabajado y/o lo hace actualmente en PIE y la región en que se desempeña laboralmente. La segunda sección preguntaba acerca de la experiencia de los participantes en la administración de la Escala WISC-III, considerando años de aplicación, número aproximado de administraciones, contexto laboral en que la utiliza y la cantidad de tiempo que dedica. En esta sección se indaga también sobre el conocimiento y realización de procedimientos específicos tales como aplicación en formas abreviadas, examen de límites y/o registro de la conducta observada de los niños durante la aplicación, solicitándoles emitir un juicio acerca de la utilidad de estos procedimientos para el diagnóstico. Asimismo, se consultó si ha requerido y/o utilizado los servicios de un traductor, en el caso de tener que administrar la escala a niños que tienen una primera lengua distinta al español (p.e., mapudungun, quechua, aymara). Al final de esta sección se preguntaba acerca de qué aspectos incluían en el informe. La última sección, de respuesta abierta, contenía preguntas acerca de las fortalezas y limitaciones que los participantes visualizan de la escala, y/o dificultades en la administración considerando cada una de las subpruebas.

El instrumento fue adaptado para su aplicación online utilizando la herramienta de formularios de Google y fue aplicado en una muestra piloto de ocho profesionales para asegurar usabilidad.

Procedimiento

Se envió una invitación a psicólogos, vía correo electrónico y redes sociales, para responder la encuesta dentro de un plazo de 2 meses. La invitación explicitaba el propósito y relevancia del estudio y se aseguraba el anonimato de las respuestas. La encuesta podía ser contestada por los participantes desde sus espacios personales, sin asistencia de las investigadoras y sin restricción de tiempo.

Con respecto a los resguardos éticos, antes que los participantes contestaran la encuesta, se les presentó un consentimiento informado, en el que se les explicó el propósito del estudio, tiempo de respuesta de la encuesta, el derecho a retirarse en cualquier momento de la investigación sin tener que dar ningún tipo de explicación y sin que ello signifique ningún perjuicio. Además, se explicitó que la información obtenida sería tratada confidencialmente y se utilizaría solo en el marco de este estudio con fines científicos y académicos. Finalmente se entregaron los datos de contacto de las investigadoras, si es que el/la participante necesitara mayor información.

Las respuestas de los participantes fueron recopiladas directamente en la plataforma web y compiladas en una base de datos. Una vez concluido el proceso de recolección de datos se procedió a revisar las respuestas en términos de valores fuera de rango, tasa de repuestas omitidas, etc.

Análisis de datos

Para el análisis de los datos se utilizó el software estadístico SPSS versión 21. Se realizaron análisis descriptivos para obtener porcentajes, frecuencias, medias y desviación estándar.

Por otra parte, se realizaron análisis estadísticos

bivariados para explorar la existencia de diferencias y correlaciones entre características sociodemográficas de los participantes y aspectos específicos del uso de la escala. Para estos análisis, y como paso preliminar, se contrastó si la variable “Compleitud del informe” se ajustaba a una distribución normal, a fin de determinar el uso de pruebas paramétricas o no paramétricas para el análisis de los datos. Para esto, se escogió la prueba estadística de Kolmogorov-Smirnov de un grupo. Debido a que los datos de esta variable no presentaron una distribución normal, se procedió a la utilización de la prueba estadística U de Mann-Whitney y la de Kruskal Wallis para explorar diferencias según variables de agrupación y el coeficiente de correlación rho de Spearman, para examinar correlaciones entre las variables de interés.

Resultados

En relación con las características de la muestra, el 52% de los participantes contaba con cuatro o menos años de ejercicio profesional, el 31% entre cinco y diez años, y el 17% más de diez años. Por otra parte, el 64% se encuentra acreditado en el uso de WISC-III. De este grupo, 40% ha sido acreditado por Cedeti-UC, 35% por la Universidad de La Frontera, y el 25% restante se divide en diversos organismos de capacitación. Un alto porcentaje (76%) de los participantes ha trabajado en PIE y 52% trabaja actualmente en ellos. En cuanto al uso de la escala, la mayor parte de los encuestados declara utilizarla en instituciones educacionales. Respecto de los años de experiencia en la administración de la escala, el 62% de los participantes cuenta con menos de cuatro años, mientras que el 37% tiene entre cinco y diez años. En cuanto a la cantidad de aplicaciones del instrumento, el 39,4% declara haberlo administrado en más de 100 ocasiones, 44% entre 20 y 100, y 16,6 % de los participantes, en menos de 20.

Tabla 2
Contexto profesional de uso de la escala

Contexto	% de respuestas
Psicología educacional	87
Psicología clínica	36,3
Psicología jurídica	5

Con relación al tiempo que se utiliza para su evaluación a la administración del instrumento, la mayor parte de los participantes (78,2%) señala dedicar en promedio entre 1,5 y dos horas; 11,5%, más de dos horas; y 10%, una hora. En cuanto a la utilización de procedimientos específicos durante la administración, es relevante destacar que solo 12% aplica formas abreviadas de la escala en el caso de niños que presentan necesidades educativas especiales (NEE) y/o cuando no existe sospecha de discapacidad intelectual.

Frente a la pregunta sobre el procedimiento de examen de límites, solo 35% de los participantes reporta conocerlo, y la mitad de ellos (17%) señala utilizarlo, considerándolo esencial para el diagnóstico. Por otra parte, 92,5% de quienes responden declaran registrar la conducta del niño durante la

aplicación de la escala, y 82% considera esta información como altamente relevante para el diagnóstico. El 99% de los participantes señala que no ha requerido y, por tanto, no ha utilizado el servicio de traductor, en el caso de evaluar niños que no hablen el español como primera lengua.

En lo que concierne a los inconvenientes al administrar la escala (tabla 3), la mayoría de los psicólogos (53,8%) reporta haber experimentado dificultades en niños con discapacidad cognitiva, mientras que otros aluden a diversas características.

Al indagar sobre los elementos que usualmente se incluyen en el informe, la casi totalidad reporta incorporar coeficiente intelectual (CI) total, CI verbal y CI de ejecución. Otros descriptores aparecen en menor proporción (tabla 4).

Tabla 3

Dificultades en la administración de WISC III, según características de los niños evaluados

Características	% de respuestas
Discapacidad cognitiva	53,8
Procedencia rural	48,1
Nivel socioeconómico bajo	35
Discapacidades sensoriales	30
Provenientes de pueblos originarios	25

Tabla 4

Porcentaje de psicólogos que incluyen descriptores en los informes

Descriptores	% de respuestas
CI total	99
CI verbal	96,3
CI de ejecución	97
Índice de comprensión verbal	84,4
Índice de organización perceptual	80,6
Índice de ausencia de distractibilidad	80,6
Índice de velocidad de procesamiento	80
Fortalezas y debilidades de perfil cognitivo	75,6
Análisis de funciones cognitivas según Kaufman	61,3
Percentil	60
Intervalo de confianza	53,1

Sobre las fortalezas de la escala (tabla 5), los participantes destacan que esta entrega un perfil exhaustivo de habilidades cognitivas, que incluye una descripción de fortalezas y debilidades intrasujeto y, en segundo término, que es un instrumento de evaluación dinámico y entretenido para

los niños, con material atractivo. Además, se menciona la facilidad de administración, corrección y obtención de puntajes y, con menor frecuencia, la entrega de información cualitativa que permite diseñar intervenciones pertinentes. Los aspectos psicométricos de la prueba incluyendo la existencia de normas nacionales son escasamente mencionados.

Tabla 5
Fortalezas de la escala

Categorías	% de respuestas
Entrega perfil de habilidades, detecta fortalezas y debilidades intrasujeto	34
Pruebas dinámicas y entretenidas, material atractivo	15
Facilidad de administración, corrección y obtención de puntajes, funcionalidad y claridad del protocolo	16,6
Permite evaluación cualitativa, estrategias de resolución de problemas, información permite diseñar intervención	14
Existencia de normas para comparar, entrega información confiable	13
Facilita evaluación de trastornos específicos (TDA, Asperger) o detección problemas visuales o auditivos	3,7
Otros: buen predictor rendimiento escolar, intervalo de confianza	3,7

Con respecto a las debilidades del instrumento (tabla 6), si bien algunos de los profesionales encuestados no reportan dificultad (ni en sus estímulos, material o administración), alrededor del 33% critica la desactualización y sesgos de los ítems en varias de las subpruebas, señalando que, de acuerdo con su experiencia, afectan mayormente a los niños de procedencia rural y de nivel socioeconómico bajo. También se menciona con alta frecuencia la

excesiva extensión de la prueba, su falta de discriminación para evaluar distintos niveles de discapacidad intelectual y la rigidez en su formato que no facilita la evaluación de los niños con necesidades educativas especiales. Otros participantes critican que la escala solo considere la inteligencia lógica matemática, descuidando otras formas de inteligencia.

Tabla 6
Debilidades de la escala

Categorías	% de respuestas
Desactualización de los ítems	33
Patrón de bajo desempeño en niños de nivel socioeducativo bajo, rurales, mapuches e inmigrantes	24
Muy extensa	13
Discrimina poco en niños con discapacidad cognitiva	5
Rigidez de formato, no se ajusta a niños con NEE	5

Sobre las dificultades observadas en las diversas subpruebas de WISC-III, los psicólogos refieren que en “Completación de figuras” se observa la presencia de imágenes desactualizadas (por ejemplo, ítem teléfono o ampollita), o poco familiares y/o desconocidas para los niños (por ejemplo, ítem termómetro). También mencionan la presencia de imágenes confusas o que poseen distractores relevantes (por ejemplo, en el ítem “guitarra” a la que le faltan las clavijas, pero por su dibujo parece faltar una cuerda).

En la subprueba de “Información”, los participantes cuestionan la presencia de preguntas descontextualizadas para la realidad cultural de niños rurales, de niveles socioeconómicos bajos u originarios de pueblo mapuche (por ejemplo, ítem de “océanos” en el cual los niños tienden a nombrar ríos o lagos, o el ítem “¿Quién es Gabriel García Márquez?” o “¿Cuál es la capital de Grecia?”). También mencionan que muchas preguntas son excesivamente dependientes de la escolarización.

En la subprueba de “Analogías” reportan dificultades en el ítem “Televisión-Teléfono”, ya que las posibilidades de respuestas correctas se encuentran desactualizadas, considerando que hoy en día muchas de las funciones de un televisor, como ver videos, también se pueden realizar en un teléfono, pero no están consignadas como respuestas correctas en el manual. En esta subprueba también refieren dificultades con el ítem “Cuadro-poema”, señalando que suele ser desconocido o poco familiar para niños rurales, inmigrantes o de bajo nivel sociocultural.

En la subprueba “Vocabulario”, los participantes destacan la presencia de palabras desconocidas para niños de procedencia rural, niveles socioeconómicos bajos o pertenecientes a pueblos originarios (por ejemplo: “mímica”, “reforma”, “aflicción”, “recluir”, “inminente” o “jactarse”). Agregan que ante el ítem “unánime” muchos niños responden “monitos japoneses”.

En la subprueba de “Comprensión”, al igual que en las subpruebas verbales anteriormente señaladas, se refiere la presencia de ítems desactualizados (por ejemplo, el ítem de la “Guía telefónica”) o poco pertinentes para los niños que viven en contextos rurales o de bajo nivel socioeconómico (ítems “cinturón de seguridad”, “patentes de autos”, “libertad de expresión” y “ventajas de senadores y diputados”). Agregan que han observado dificultades en la comprensión del ítem “Luces”, debido a su formulación confusa.

Por otra parte, en las subpruebas no verbales, basadas principalmente en estímulos perceptuales o que requieren manipulación (“Completación de figuras”, “Claves”, “Ordenamiento de historias”, “Cubos”, “Ensamblaje de objetos” y “Búsqueda de símbolos”), los encuestados no reportan dificultades en su administración, salvo en el caso de los niños que presentan sensorialidades diferentes.

En cuanto a los análisis bivariados, debido a que la variable “Complejidad del informe” no presentó una distribución semejante, de acuerdo con los resultados de la prueba de Kolmogorov-Smirnov ($Z = 2,14$; $p < ,001$), se procedió a la utilización de pruebas no paramétricas para el análisis de los datos.

Se examinó la existencia de diferencias estadísticamente significativas en el nivel de “Complejidad del informe”, según las variables universidad de egreso, acreditación en el uso de la escala, estar tra-

bajando o haber trabajado en proyectos de integración escolar. De acuerdo con los resultados obtenidos, la única variable que presenta diferencias es la acreditación en el uso de la escala ($U = 2410,00$; $p = 0,046$), evidenciando que quienes están acreditados elaboran informes más completos ($Md = 9$) que quienes no lo están ($Md = 8$).

La exploración de diferencias en el nivel de Complejidad del informe según el grado académico del participante, arroja que no existen diferencias significativas entre quienes poseen el grado de licenciado, magíster, doctorado o posdoctorado en psicología ($\chi^2(3, N = 159) = 3,94, p = ,268$).

Al relacionar el nivel Complejidad del informe y las variables años de experiencia en la administración de la escala y la cantidad de administraciones de esta, los resultados obtenidos evidencian que no existen asociaciones estadísticamente significativas según años de experiencia en la administración de la escala ($r_s(152) = -0,05$; $p > ,05$) y cantidad de administraciones realizadas ($r_s(157) = -0,03$; $p > ,05$).

Finalmente, la exploración de la existencia de relaciones entre las variables años de experiencia en la administración de la escala, cantidad de administraciones de esta y tiempo utilizado en su aplicación, sugiere una relación directa entre años de experiencia y cantidad de administraciones ($r_s(149) = 0,55$; $p < ,001$), una relación inversa entre años de experiencia y cantidad de tiempo utilizado en la administración ($r_s(148) = -0,26$; $p < ,001$), y entre cantidad de administraciones y tiempo utilizado ($r_s(154) = -0,37$; $p < ,001$). Esto indicaría que, a más años de experiencia del profesional, más cantidad de aplicaciones ha realizado y menos tiempo demora en la aplicación.

Discusión

En cuanto a las características de la muestra, la mayor parte de los profesionales encuestados son mujeres, lo que es congruente con la distribución por género de los estudiantes que ingresan a la carrera de psicología (Consejo Nacional de Educación, 2016), y con las áreas de desempeño profesional donde las psicólogas tienden a concentrarse: áreas clínica y educacional (Makrinov, Scharager, & Molina, 2005).

Dado la limitada trayectoria profesional (casi dos tercios de los psicólogos cuentan con menos de

cinco años de experiencia), la mayoría no ha cursado formación de posgrado con posterioridad al grado académico de licenciado en psicología. Es relevante destacar, sin embargo, que dos tercios de los encuestados se encuentran acreditados en el uso de la escala, lo que sugiere que este grupo de profesionales privilegia la formación profesional que les permite responder a las exigencias del contexto laboral por sobre la formación académica. En este sentido, es destacable que, pese a su limitada trayectoria profesional, han acumulado amplia experiencia en el uso de la prueba, lo que refleja que la evaluación intelectual, particularmente en contextos educacionales, constituye un espacio laboral relevante para los psicólogos recién egresados.

Con relación a las prácticas evaluativas, los tiempos de administración de la escala declarados por los psicólogos (más de una hora y media en su aplicación) son mayores que los propuestos por el autor (Wechsler, 1991), quien sugiere un tiempo aproximado de 65 a 80 minutos, pero coinciden con el estudio de Ryan, Glass, y Brown (2007), quienes encontraron que los tiempos utilizados varían en función de la edad y nivel escolar de los examinados, así como la cantidad de subpruebas aplicadas. Este mayor tiempo utilizado sugiere rigurosidad y exhaustividad de las evaluaciones, lo que se condice con la relevancia de las decisiones que se toman a partir de los resultados.

Respecto de los contenidos que se incluyen en los informes, un altísimo porcentaje de los participantes reporta los valores del CI total, CI verbal y CI de ejecución. Esto es esperable, considerando que es la información más básica que entrega la prueba y, posiblemente, la más demandada desde sus espacios laborales. Una menor proporción (75%), aunque igualmente mayoritaria, refiere incluir el perfil de fortalezas y debilidades del evaluado, de lo que se infiere que un cuarto de los encuestados, a pesar de dedicar 90 minutos, o más, en la evaluación, desaprovecha información relevante que permitiría diseñar intervenciones más pertinentes.

En lo que concierne a procedimientos específicos de la administración de la Escala WISC-III, destaca especialmente el valor que los profesionales otorgan al registro de la conducta del niño durante la aplicación, información altamente relevante en la determinación de un diagnóstico de funcionamiento cognitivo (Sattler, 2010). Sin embargo, un procedimiento relevante como el examen

de límites es poco conocido y, posiblemente, poco valorado, dado que la mitad de quienes declaran conocerlo no lo utilizan. Considerando el valor de este procedimiento, tanto para evaluar la influencia de variables extrañas en el desempeño del evaluado como para acceder a las estrategias de procesamiento de la información y resolución de problemas, es lamentable que, a pesar de la cantidad de tiempo utilizado en la administración, no se acceda a esa información (Sattler, 2010). Cabe preguntarse si estos contenidos están disponibles en los planes de estudios de las carreras de psicología, considerando su relevancia para un diagnóstico integral de los evaluados, y de no ser así, la necesidad de incorporarlos como contenido mínimo.

Sobre las fortalezas de la escala, alrededor de un tercio de la muestra valora que a través de ella es posible acceder a un completo y exhaustivo perfil de las habilidades cognitivas de los evaluados, además de considerarla atractiva y/o entretenida para los niños/as. Llama la atención que la existencia de normas nacionales y las evidencias de confiabilidad y validez de la prueba son mencionada por un bajo porcentaje de los usuarios, obteniendo mayor frecuencia aspectos más funcionales como la facilidad de administración, corrección y obtención de puntajes, funcionalidad y claridad del protocolo. Lo anterior nuevamente lleva a preguntarse sobre la inclusión de contenidos psicométricos en la formación de los psicólogos, que permita a los profesionales evaluar la calidad de los instrumentos que utilizan.

Las debilidades más frecuentemente mencionadas, como la presencia de elementos desactualizados, son esperables considerando la cada vez más rápida obsolescencia de los instrumentos, atendida la velocidad de los cambios tecnológicos y culturales (ITC, 2013). Por otra parte, la percepción de los profesionales de los sesgos hacia grupos específicos, como niños de sectores rurales o vulnerables económicamente, son una crítica frecuente a las escalas Wechsler en cuanto a la carga cultural que presentan, especialmente en las subpruebas verbales, que favorecen a niños de nivel educacional más alto (Cianci et al., 2013; Eilertsen et al., 2016; Fuica et al., 2014). Este último aspecto se constituye en un desafío para los psicólogos en la evaluación de esta población y sugiere la necesidad de utilizar otros instrumentos y/o fuentes de información para compensar estas falencias. Además, se hace necesario desarrollar investigación a nivel nacional que

explore la existencia de posibles sesgos en los ítems de instrumentos como este, respondiendo no solo a una necesidad profesional particular, sino a imperativos éticos y orientaciones internacionales sobre el uso de test. En este sentido, es alentador el trabajo de estandarización de la quinta versión de la escala (Kaufman, Raiford, & Coalson, 2016; Wechsler, 2014), realizado en Chile, que se espera supere las falencias de la actual versión en uso (Rascón, 2018).

Finalmente, la relación entre estar certificado y la elaboración de informes más completos sugiere la pertinencia de este tipo de formación para el buen desempeño de los profesionales. En esta línea la obligatoriedad de la certificación en el uso de instrumentos específicos por parte de organismos competentes que den garantías de calidad e imparcialidad puede ser una respuesta al desafío de instalar en lo/as profesionales, competencias para la aplicación, corrección e interpretación de instrumentos psicológicos. Esto es particularmente relevante si se considera que la aplicación de test es considerada una de las competencias menos relevantes en la formación de psicólogos, de acuerdo con el estudio de Cabrera, Larraín, Moretti, Arteaga, y Energici (2010) en una muestra de estudiantes, profesionales y profesores de psicología de la Región Metropolitana.

Lo anterior, sin embargo, no exime a las universidades de la responsabilidad de formar en evaluación psicológica, entendida esta como una compleja tarea conceptual de resolución de problemas y toma de decisiones, que requiere de la integración de conocimientos teóricos, psicométricos y éticos para realizar una evaluación rigurosa y culturalmente pertinente.

En cuanto a las limitaciones de este estudio, es importante señalar el bajo tamaño muestral y la escasa proporcionalidad en la distribución geográfica de los psicólogos que contestaron la encuesta, lo que se asocia al tipo de muestreo utilizado. Respecto de la mayor concentración en la muestra de psicólogos desempeñándose en el ámbito educacional, si bien puede sesgar los resultados, por otra parte, son estos quienes tienen más experiencia en el uso de la batería, y mayor cercanía con las implicancias que esa evaluación tiene en la trayectoria educativa de los niños y niñas.

Considerando estas limitaciones, se espera que se desarrollen nuevos estudios que exploren la percepción de los psicólogos sobre la quinta versión de

la escala de Wechsler en Chile o de otros instrumentos de evaluación psicológica de uso masivo en el país, con el fin de nutrir el campo de la evaluación psicológica nacional y aportar al desarrollo de prácticas evaluativas rigurosas, éticas y sensibles culturalmente.

Referencias

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, District of Columbia: American Educational Research Association.
- American Psychological Association. (2000). *Report of the task force on test user qualifications*. Recuperado de <https://bit.ly/2R4Q2fi>
- American Psychological Association. (2010). *Enmiendas 2010*. Recuperado de <https://bit.ly/2sq4esr>
- Asociación Argentina de Estudio e Investigación en Psicodiagnóstico. (1999). *Código de ética del psicodiagnóstico*. Recuperado de <https://bit.ly/2Hp9y6Q>
- Cabrera, P., Larraín, A., Moretti, R., Arteaga, M., & Energici, M. A. (2010). La formación en psicología desde una perspectiva de competencias. Una contribución para el mejoramiento de la formación universitaria en Chile. *Calidad en la Educación, 33*, 183-221. <http://dx.doi.org/10.31619/caledu.n33.143>
- Cianci, L., Orsini, A., Hulbert, S., & Pezuti, L. (2013). The influence of parents' education in the Italian standardization sample of the WISC-III. *Learning & Individual Differences, 28*, 47-53. <http://dx.doi.org/10.1016/j.lindif.2013.09.009>
- Consejo General de la Psicología de España. (2017). *Evaluación tests editados en España. Consejo General de la Psicología de España*. Recuperado de <https://bit.ly/2T3PjNf>
- Consejo Nacional de Educación. (2016). *Resumen Ejecutivo*. Recuperado de <https://bit.ly/2R5FT1T>
- Decker, S. L., Hale, J. B., & Flanagan, D. P. (2013). Professional practice issues in the assessment of cognitive functioning for educational applications. *Psychology in the Schools, 50*(3), 300-314. <https://doi.org/10.1002/pits.21675>
- Eilertsen, T., Thorsen, A. L., Holm, S. E. H., Be, T., Sorensen, L., & Undersold, A. J. (2016). Parental socioeconomic status and child intellectual functioning in a Norwegian sample. *Scandinavian Journal of Psychology, 57*(5), 399-405. <http://dx.doi.org/10.1111/sjop.12324>

- Elliot, C. D. (2007). *Differential Ability Scale: Introductory and technical handbook* (2nd ed.). San Antonio, Texas: Harcourt Assessment.
- Elosua, P. (2017). Avances, proyectos y retos internacionales ligados al uso de tests en Psicología. *Estudios de Psicología*, 34(2), 201-210.
<http://doi.org/czxx>
- Elosua, P. & Muñiz, J. (2013). Proyectos españoles para una mejora en el uso de los tests. *Psiencia. Revista Latinoamericana de Ciencia Psicológica*, 5(2), 139-143.
 Recuperado de <https://bit.ly/2FRNiQC>
- Fonseca-Pedrero, E. & Muñiz, J. (2017). Quinta evaluación de tests editados en España: mirando hacia atrás, construyendo el futuro. *Papeles del Psicólogo*, 38(3), 161-168.
<http://doi.org/czxm>
- Fuica, P., Lira, J., Alvarado K., Araneda, C., Lillo, G., Miranda, R., Tenorio, M., & Pérez-Salas, C. P. (2014). Habilidades cognitivas, contexto rural y urbano: comparación de perfiles WAIS-IV en jóvenes. *Terapia Psicológica*, 32(2), 143-152.
<http://doi.org/czxn>
- Georgas, J., Weiss, L. G., van de Vijver, F. J. R., & Saklofske, D. H. (Eds.). (2003). *Culture and children's intelligence: Cross-cultural analysis of the WISC-III*. San Diego, California: Academic Press.
<http://doi.org/czxp>
- Gregoire, J. (2001). Factor Structure of the French Adaptation of the WISC-III: Three or Four Factors? *International Journal of Testing*, 1(3-4), 271-281.
<https://doi.org/10.1080/15305058.2001.9669475>
- Hernández, A., Ponsoda, V., Muñiz, J., Prieto, G., & Elosua, P. (2016). Revisión del modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 37(3), 192-197.
 Recuperado de <https://bit.ly/2AVSob6>
- International Test Commission. (2001). International guidelines for test use. *International Journal of Testing*, 1(2), 93-114.
http://dx.doi.org/10.1207/S15327574IJT0102_1
- International Test Commission. (2013). *Guidelines for practitioner use of test revisions, obsolete tests, and test disposal*.
 Recuperado de <http://www.intestcom.org>
- International Test Commission. (2014). ITC guidelines on quality control in scoring, test analysis, and reporting of test scores. *International Journal of Testing*, 14(3), 195-217.
<http://dx.doi.org/10.1080/15305058.2014.918040>
- Kaufman, A. S. & Kaufman, N. L. (1983). *K-ABC administration and scoring manual*. Circle Pines, Minnesota: American Guidance Service.
- Kaufman, A. Raiford, S. E., & Coalson, D. L. (2016). *Intelligence testing with WISC-V*. Hoboken, New Jersey: John Wiley.
- Laurent, J., Swerdlik, M., & Ryburn, M. (1992). Review of validity research on the Stanford-Binet Intelligence Scale: Fourth edition. *Psychological Assessment*, 4(1), 102-112.
<http://dx.doi.org/10.1037/1040-3590.4.1.102>
- Makrinov, N., Scharager, J., & Molina, L. (2005). Situación actual de una muestra de psicólogos egresados de la Pontificia Universidad Católica de Chile. *Psykhé*, 14(1), 69-77.
<http://doi.org/dmh5g4>
- Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Oxford, England: Williams & Wilkins.
- Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). Direcciones para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25(2), 151-157.
<http://dx.doi.org/10.7334/psicothema2013.24>
- Muñiz, J., Hernández, A., Ponsoda, V. (2015). Nuevas directrices sobre el uso de los tests: investigación, control de calidad y seguridad. *Papeles del Psicólogo*, 36(3), 161-173.
 Recuperado de <https://bit.ly/2DvRSCx>
- Naglieri, J. A. (1999). *Essentials of CAS assessment*. New York, New York: Wiley.
- Ramírez, V. & Rosas, R. (2007). Estandarización del WISC-III en Chile: descripción del test, estructura factorial y consistencia interna de las escalas. *Psykhé*, 16(1), 91-109.
<http://doi.org/c8dect>
- Rascón, A. (2018, 20 de agosto). *CEDETi UC presenta la Escala Wechsler de Inteligencia para Niños, Quinta Edición (WISC-V)*. Centro UC Tecnologías de Inclusión.
 Recuperado de <https://bit.ly/2W8tw94>
- Rotzika, V., Vlassopoulos, M., Legaki, L., Sini, A., Rogakou, E., Sakellariou, K., ... & Anagnostopoulos, D. C. (2009). The WISC-III profile in Greek children with learning disabilities: Different language, similar difficulties. *International Journal of Testing*, 9(3), 271-282.
<https://doi.org/10.1080/15305050903106891>
- Ryan, J. J., Glass, L. A., & Brown, C. N. (2007). Administration time estimates for Wechsler Intelligence Scale for Children-IV subtests, composites, and short forms. *Journal of Clinical Psychology*, 63(4), 309-318.
<https://doi.org/10.1002/jclp.20343>
- Sattler, J. M. (2010). *Evaluación infantil. Fundamentos cognitivos* (5ta ed.). Ciudad de México, México: Manual Moderno.
- Schrank, F. A., McGrew, K. S., & Mather, N. (2014). *Woodcock-Johnson IV tests of cognitive abilities*. Rolling Meadows, Illinois: Riverside.
- Sotelo-Dynega, M. & Dixon, S. G. (2014). Cognitive assessment practices: A survey of school psychologists. *Psychology in the Schools*, 51(10),

1031-1045.

<https://doi.org/10.1002/pits.21802>

- Urzúa, A., Ramos, M., Alday, C., & Alquinta, A. (2010). Madurez neuropsicológica en preescolares: propiedades psicométricas del test CUMANIN. *Terapia Psicológica, 28*(1), 13-25.
- Vinet, E. V. & González R., N. (2013). Desarrollos actuales y desafíos futuros en la evaluación psicológica en Chile. *Psiencia. Revista Latinoamericana de ciencia Psicológica, 134-138*.
<http://dx.doi.org/10.5872/psiencia.v5i2.117>
- Watkins, M. W. & Beaujean, A. A. (2014). Bifactor structure of the Wechsler Preschool and Primary Scale of Intelligence-fourth edition. *School Psychology Quarterly, 29*(1), 52-63.
<https://psycnet.apa.org/doi/10.1037/spq0000038>
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—Third Edition manual*. San Antonio, Texas: The Psychological Corporation.
- Wechsler, D. (2014). *Wechsler Intelligence Scale for Children—fifth edition*. San Antonio, Texas: The Psychological Corporation.
- Wechsler, S. M. (2013). El movimiento para el desarrollo de los tests en Brasil. Una experiencia pionera. *Psiencia. Revista Latinoamericana de Ciencia Psicológica, 5*(2), 130-133.
Recuperado de <https://bit.ly/2UacNjN>
- Weiss, L. G., Saklofske, D. H., Coalson, D. L., & Raiford, S. E. (2010). *WAIS-IV clinical use and interpretation*. San Diego, California: Academic Press.

Fecha de recepción: 2 de febrero de 2018

Fecha de aceptación: 23 de octubre de 2018