

INTERPRETACIÓN DEL WISC-IV. PUNTUACIONES COMPUESTAS Y MODELOS CHC

WISC-IV SCORING AND CHC-BASED INTERPRETATIVE APPROACHES

María Elena Brenlla

Centro de Investigaciones en Psicología y Psicopedagogía (CIPP), Argentina

Universidad Católica Argentina

Resumen: Los objetivos de este artículo fueron valorar si la adaptación argentina de la Escala de Inteligencia de Wechsler para Niños, cuarta edición WISC-IV (Taborda, Brenlla y Barbenza, 2011), evalúa los mismos constructos en distintos grupos de edad, analizar si la estructura de puntuaciones compuestas de cuatro dominios específicos y un factor general postulada por los autores del WISC-IV se replica en nuestro medio y evaluar dos modelos alternativos basados en la teoría de Cattell-Horn-Carroll (CHC). Se llevaron a cabo estudios de Análisis Factorial Confirmatorio de las puntuaciones obtenidas ($n = 1437$). Los resultados indican que el WISC-IV evalúa los mismos constructos en distintas edades, que tanto la estructura de puntuaciones compuestas propuesta en el WISC-IV como uno de los modelos CHC son significativos para interpretar las puntuaciones del test y que la varianza explicada es similar en ambos modelos. El WISC-IV evalúa habilidad cristalizada, procesamiento visual, razonamiento fluido, memoria de corto plazo y operativa y procesamiento visual. En particular, se ha notado que el razonamiento fluido y el procesamiento visual pueden evaluarse por separado o combinados como razonamiento perceptivo.

Palabras Clave: *inteligencia, niños, interpretación, puntuaciones compuestas*

Abstract: The validity of WISC-IV current four-factor scoring structure and the Cattell-Horn-Carroll (CHC) theory-based models of the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) we investigated by the application of confirmatory factor analysis of scores from the Argentine WISC-IV standardized sample ($n = 1437$). Results reveal that the WISC-IV measures the same constructs across age, the resulting interpretation could be applied to children with various age levels. Both the four-factor structure and the alternative CHC-based model were supported. Variance explained was similar across models. The WISC-IV measures crystallized ability, visual processing, fluid reasoning, working memory and processing speed. In particular, either separating fluid reasoning and visual processing, or combining them as the Perceptual Reasoning Index provides meaningful explanation.

Keywords: *intelligence, children, interpretation, composite scores*

Introducción

Es conocida la importancia que la psicometría tiene para la psicología. En la psicología experimental, proveyendo de la construcción de procedimientos –escalas- para la medición de estímulos; en la psicología diferencial, proponiendo modelos para la cuantificación de las diferencias individuales de los sujetos en rasgos, atributos y/o conductas (Martínez Arias, 1996). Un aspecto esencial es la construcción de tests que posibiliten la medición y la estimación válida y confiable de esas diferencias individuales ya que esto tiene gran impacto tanto en el ámbito de la

investigación cuanto en la psicología aplicada, en particular en la evaluación psicológica (EP).

Entre los tests de los que se vale la EP se destacan los psicométricos caracterizados por refinamientos metodológicos y tecnológicos que ayudaron enormemente a la precisión y validez de las medidas con ellos realizados. Esto permitió, por una parte, una evaluación más ajustada de los rasgos, atributos o características que median y, por otra, profundizar y ampliar el campo de la investigación de variables psicológicas. Un subgrupo son los tests de ejecución máxima que evalúan atributos cognitivos a través de tareas de rendimiento.

Mediante ellos se ha logrado no solo contribuir a un mejor diagnóstico de trastornos relacionados –tales como el retraso y el talento intelectuales o los trastornos de aprendizaje– sino también han posibilitado inferencias teóricas sustantivas a partir del análisis e interpretación de métodos factoriales. Esto ha dado lugar a lo que se conoce como el enfoque psicométrico para la explicación de la inteligencia (Cohen y Swerdlick, 2000).

Fue Spearman (1904, 1927) quien estableció, en los albores del siglo XX, una teoría de la inteligencia y un método para probarla –el análisis factorial– cuyos postulados básicos han mostrado consistencia y robustez hasta la actualidad. Spearman pensó una teoría que se caracteriza por su parsimonia: esto es, con pocas variables se explica un gran dominio de conocimiento. Esencialmente, consideraba que en cualquier habilidad intelectual pueden distinguirse –al menos formalmente– un factor *general (g)* de la inteligencia y otro *específico*. El factor general se caracteriza por mantenerse constante en cada individuo aunque puede variar dentro de márgenes muy amplios de un sujeto a otro. En cambio, los específicos, varían no solo entre los diversos individuos sino también en el mismo sujeto conforme a qué aptitud está siendo examinada. Con el tiempo, comprendió que entre algunas habilidades específicas –como la de memorizar sílabas y números– era posible establecer grupos extensos de tareas próximamente relacionadas. Spearman llamó factores de *grupo* a aquellas características que están presentes en más de una habilidad pero no en todas y que expresan alguna afinidad entre las integrantes de un conjunto. Gran parte del debate sobre la evaluación intelectual durante los últimos 60 años se ha centrado en la existencia de este aspecto global subyacente de la inteligencia que influye en el desempeño del individuo a través de dominios cognitivos (Gustafsson & Undheim, 1996; Jensen, 1998).

Unas décadas más tarde que Spearman, Cattell (1941, 1957) introdujo la noción que la inteligencia estaba compuesta por dos factores, la inteligencia *fluida* y la *cristalizada*. La primera relacionada con la actuación intelectual ante problemas novedosos y, la segunda, con el conocimiento adquirido. Luego, su discípulo y colega Horn amplió esta teoría e incluyó factores tales como la percepción visual, la

memoria de corto plazo y el almacenamiento y recuperación a largo plazo, entre otros (Horn, 1985, 1988, 1991; Horn & Noll, 1997).

Después de varias décadas de investigación, las ideas de Cattell y Horn en cuanto a dominios más específicos de la inteligencia han recibido un apoyo empírico sustancial. Carroll (1993, 1997), en su ya clásica revisión, comparó los resultados de distintos análisis factoriales de las medidas de habilidades cognitivas existentes y concluyó que la evidencia de un factor general de inteligencia era abrumadora. Al mismo tiempo, indicó que la inteligencia tiene una estructura jerárquica y que se compone de habilidades específicas que parecen agruparse en dominios de habilidad cognitiva de nivel superior (Carroll, 1993; Keight, 1990).

La confluencia de la teoría de la inteligencia fluida y cristalizada de Cattell y Horn y de la concepción jerárquica de Carroll se plasmó en lo que hoy conocemos como la teoría CHC. Ésta versa acerca del contenido y la estructura de la inteligencia y sostiene que las habilidades intelectuales se organizan en tres estratos, un estrato próximo (*narrow strata*), un estrato amplio (*broad strata*) y un estrato general. La CHC fue revisada y ampliada sobre la base de los estudios de McGrew y Flanagan (1998) y McGrew (2009), en la actualidad, hay consenso en considerar que el primer estrato está constituido por 70 componentes que se subsumen en habilidades amplias del segundo estrato, a saber, inteligencia cristalizada, razonamiento fluido, procesamientos visual y auditivo, memoria de corto plazo, recuperación a largo plazo, velocidad de procesamiento, conocimiento cuantitativo y habilidades para la lectura y la escritura. Según la revisión de McGrew (2009) y de Schneider y McGrew (2012), en este estrato deberían incluirse también la velocidad para la reacción y la decisión, la velocidad psicomotora, el conocimiento de dominios específicos y las habilidades táctiles, cinestésicas, olfatorias y psicomotoras. De esta manera, el estrato II incluye 16 habilidades. Éstas, a su vez, se condensan en el tercer estrato que, en términos psicométricos, se corresponde con *g*.

En forma paralela a esta línea de investigaciones y también inspirado en las concepciones de Spearman, Wechsler (1939) había definido a la inteligencia como “la capacidad global y agregada de un individuo para actuar con una finalidad, pensar racionalmente y re-

lacionarse adecuadamente con el entorno” (p. 3). Sobre la base de estos supuestos, creó sus mundialmente conocidos tests de inteligencia (*Wechsler Adult Intelligence Scale –WAIS–*; *Wechsler Intelligence Scale for Children –WISC–*; *Wechsler Preschool and Primary Scale of Intelligence –WPPSI–*). Para el diseño de sus tests, Wechsler seleccionó y desarrolló medidas destinadas a evaluar los aspectos cognitivos de la inteligencia que consideró más relevantes: comprensión verbal, razonamiento abstracto, organización perceptual, razonamiento cuantitativo, memoria y velocidad de procesamiento. Las teorías y las medidas de inteligencia más modernas han confirmado que todas esas áreas reflejan aspectos importantes de la habilidad cognitiva (Carroll, 1993, 1997; Horn, 1991).

Algunos investigadores han deducido que Wechsler consideraba que la estructura de la inteligencia tenía dos factores y que por eso dividió sus tests en tareas verbales y de ejecución (Wechsler, 2003). Aunque este agrupamiento fue de naturaleza racional y no psicométrica dio lugar, en las versiones anteriores del Test de Inteligencia para Niños de Wechsler (WISC, WISC-R, WISC-III), a tres Cocientes Intelectuales (CI): Verbal (CIT), de Ejecución (CIE) y Total (CIT).

No obstante, en la revisión más reciente –WISC-IV– se eliminaron los CI verbal y de ejecución, se incrementó la medición de dominios de funcionamiento cognitivo más específicos –velocidad de procesamiento (VP), memoria operativa (MO), comprensión verbal (CV) y razonamiento perceptivo (RP)– y se mantuvo la evaluación de la inteligencia global a través de un cociente intelectual total.

Estos cambios en el WISC-IV, que son de un tenor muy importante, obedecen indudablemente al *aggiornamento* del test a los hallazgos de la teoría CHC habida cuenta que, en términos estrictos, no hay una teoría formal que sustente a las escalas Wechsler más que la asunción general de los postulados de Spearman. Por otro lado, no hay que olvidar que la CHC ha tenido la suficiente consistencia como para servir de fundamento a otros tests de inteligencia o de habilidades cognitivas como el *Stanford-Binet Fifth Edition* (Roid, 2003) o el *Woodcock-Johnson Test of Cognitive Abilities III* (Woodcock, McGrew & Mather, 2001).

Ya la evidencia obtenida de investigaciones anteriores con el WISC-III indicaba que éste medía cuatro dominios cognitivos: Comprensión Verbal, Organización Perceptiva, Ausencia de Distractibilidad y Velocidad de Procesamiento (Blaha & Wallbrown, 1996; Sattler, 2001; Wechsler, 1991). Debido a que el WISC-IV retiene diez subtests de la edición anterior y agrega cinco subtests nuevos diseñados para medir constructos similares, se esperaba que el WISC-IV midiera los mismos cuatro dominios cognitivos.

Para probar esta hipótesis, los autores del WISC-IV llevaron a cabo estudios factoriales exploratorios y confirmatorios. Basándose en la teoría, la investigación y los resultados de los análisis factoriales exploratorios pudo evaluarse el modelo factorial propuesto para el WISC-IV utilizando modelos de ecuación estructural (Jöreskog & Sörbom, 1993). El análisis factorial confirmatorio está diseñado para testear la estructura interna especificada por los investigadores. En este sentido, es diferente del exploratorio ya que la agrupación de subtests se realiza *a priori* y no por medio de un algoritmo. Las relaciones específicas entre las variables observadas (por ejemplo los subtests) y las variables latentes (por ejemplo, Comprensión Verbal) se indican previamente y ese modelo se testea para determinar si las relaciones entre las variables observadas y las latentes confirman la estructura sobre la que se hizo la hipótesis. En el estudio norteamericano con el WISC-IV se testearon diferentes modelos de entre uno y cinco factores y se compararon entre sí para evaluar su estructura factorial y las evidencias de validez interna. Todos los modelos asumieron que los factores pueden tener correlación y que los errores no tenían correlación entre sí.

Los resultados reportados por los autores del WISC-IV (Wechsler, 2003) sugieren que los modelos de 4 y 5 factores se ajustan bien a los datos. Sin embargo, los referidos al modelo 5 no indican una mejora sustancial sobre el modelo 4 y tienen el inconveniente de que plantean un factor compuesto por un solo subtest (Aritmética). De acuerdo a ello, la hipótesis de cuatro dominios bien diferenciados de la inteligencia resultó más consistente y recibió un sustento muy satisfactorio luego de realizados dichos análisis (Wechsler, 2003).

La Figura 1 muestra la estructura interna definitiva del WISC-IV.

Factor general	Comprensión Verbal	Semejanzas Vocabulario Comprensión Información Adivinanzas
	Razonamiento Perceptivo	Construcción con cubos Conceptos Matrices Completamiento de figuras
	Memoria Operativa	Retención de dígitos Letras y Números Aritmética
	Velocidad de Procesamiento	Claves Búsqueda de símbolos Animales

Figura 1. Estructura del WISC-IV (Wechsler, 2003)

No obstante, Keith, Fine, Taub, Reynolds & Kranzler (2006) han objetado esta estructura de cuatro dominios específicos en el WISC-IV. En su artículo *Análisis de primer orden, con muestras múltiples, análisis factorial confirmatorio del WISC-IV: qué es lo que se evalúa?*¹ reseñan un estudio realizado a partir de la muestra norteamericana cuyos objetivos fueron analizar si el test evalúa los mismos constructos en los distintos rangos de edad y contribuir a la explicación de la naturaleza de ellos.

Los resultados que obtuvieron indicaron que si bien el WISC-IV evaluaba los mismos atributos en las distintas edades, no obstante, la estructura de puntuaciones del test se ajustaba mejor a un modelo alternativo basado en la teoría CHC que al de los cuatro dominios específicos postulado por los autores del WISC-IV (Comprensión verbal; Razonamiento perceptivo; Memoria operativa y Velocidad de procesamiento).

El modelo que propusieron Keith et al. (2006) sugiere que el WISC-IV mide aspectos de la inteligencia cristalizada, del procesamiento visual, del razonamiento fluido, de la memoria de corto plazo y de la velocidad de procesamiento. Por esta razón, recomiendan que los usuarios del WISC-IV reagrupen la interpretación de los subtests de Razonamiento perceptivo y de Aritmética tal como lo muestra la Figura 2. En ésta puede notarse que la novedad del modelo de Keith consiste en reubicar en distintos dominios a los subtests de *Completamiento de figuras*, *Aritmética*, *Búsqueda de símbolos* y *Matrices*, ya que estas pruebas son las que, en el análisis

factorial, presentaron cargas compartidas en esos distintos factores.

Unos años más tarde, Chen, Keith, Chen y Chang (2009), replicaron la investigación de Keith et al. (2006) a partir de los datos de la muestra taiwanesa del WISC-IV. Los resultados coincidieron en cuanto que el test evalúa los mismos constructos a través de las distintas edades pero difirieron claramente respecto de la supremacía del modelo CHC y aún de la estructura de subtests que la componen. De acuerdo a los hallazgos de Chen et al. (2009) ambos modelos –el de la estructura de cuatro factores del WISC-IV y el modelo alternativo CHC– recibieron un sustento empírico satisfactorio. La varianza explicada fue similar en ambos y tanto uno como otro proveyeron de explicaciones significativas acerca de la organización interna del test. No obstante, la estructura factorial según criterios CHC no coincidió con la obtenida por Keith. En la Figura 2 se muestra la estructura obtenida por Chen et al. (2009). Las diferencias más importantes son en cuanto al razonamiento fluido, la inteligencia cristalizada y el procesamiento visual. Concomitantemente, los subtests *Completamiento de figuras* (CF), *Aritmética* (Ar), *Búsqueda de símbolos* (BS), *Semejanzas* (S) y *Matrices* (M) presentaron cargas compartidas en distintos factores. Por ejemplo, según el estudio de Keith et al. (2006) M tiene cargas compartidas en procesamiento visual y razonamiento fluido; CF, en procesamiento visual e inteligencia cristalizada y Ar, en memoria de trabajo y razonamiento fluido. Por el contrario, en el estudio de Chen et al. (2009), Matrices carga solo en el segundo, CF solo en procesamiento visual y Ar en memoria de trabajo e inteligencia cristalizada pero no en razonamiento fluido.

En consecuencia, el análisis de las puntuaciones sobre la base del modelo CHC y sobre la estructura de cuatro factores del WISC-IV (CIT e Índices CV, RP, MO y VP) son de interés para analizar e interpretar los resultados del test en ambos estudios aunque sus resultados difieren de los hallados por Keith et al. (2006) no solo en cuanto a la supremacía de la teoría CHC para explicar las puntuaciones del WISC-IV sino también a la estructura de su modelo CHC.

Como es lógico, este tipo de resultados divergentes interpelan las evidencias de la validez de la estructura interna de las pun-

1 (Higher Order, Multisample, Confirmatory Factor Analysis of the Wechsler Intelligence Scale for Children – Fourth Edition: What Does It measure?)

		Keith et al. (2006)	Chen et al. (2009)
Factor general	Inteligencia cristalizada	Semejanzas Vocabulario Comprensión Información Adivinanzas Completamiento de figuras	Semejanzas Vocabulario Comprensión Información Aritmética (Adivinanzas)
	Procesamiento visual	Construcción con cubos Matrices Completamiento de figuras Búsqueda de símbolos	Construcción con cubos Completamiento de figuras Búsqueda de símbolos
	Razonamiento fluido	Conceptos Matrices Aritmética	Conceptos Matrices Semejanzas
	Memoria Operativa	Retención de dígitos Letras y Números	Retención de dígitos Letras y Números Aritmética
	Velocidad de Procesamiento	Claves Búsqueda de símbolos Animales	Claves Búsqueda de símbolos Animales

Figura 2. Estructura del WISC-IV según la teoría CHC (Keith et al., 2006 y Chen et al., 2009)

tuaciones del test. Y, con ello, la calidad de las inferencias psicológicas que pueden hacerse a partir de sus resultados. No debe olvidarse que muchos tests, como es el caso del WISC-IV, brindan información que puede ser relevante para la toma de decisión en psicología aplicada (clínica, educacional, forense, entre otras). Su administración permite al psicólogo evaluar la habilidad intelectual de niños y adolescentes y es uno de los elementos de juicio que se tienen en cuenta para indicar la educación especial, la pertinencia de una intervención psicopedagógica o la derivación a otras especialidades como la neurología o la psicología clínica (Wechsler, 2004). Por ende, resulta crucial contar con estudios que avalen que la estructura interna del test responde a una concepción clara de la inteligencia y no a una construcción *pseudoética* (Marín, 1986; Triandis et al., 1980), en la que se aceptan como válidas y universales las inferencias realizadas en otros contextos sin el debido análisis en la cultura en la que se quiere adaptar un test.

En vista de estos antecedentes y de la importancia que tiene el estudio de los supuestos teóricos sobre los que se fundamenta un test, los objetivos de este artículo son (a) valorar si en la adaptación argentina del WISC-IV (Taborda, Brenlla y Barbenza, 2011) se evalúan los mismos constructos en distintos grupos de edad,

tal como ocurre en la versión norteamericana, (b) analizar si la estructura de puntuaciones compuestas de cuatro dominios específicos y un factor general postulada por los autores del WISC-IV se replica en nuestro medio y (c) poner a prueba los modelos alternativos basados en la CHC propuestos por Keith et al. (2006) y Chen et al. (2009).

Método

Tipo de estudio y diseño

Se trata de un estudio de naturaleza inferencial con diseño transversal que se sirve de métodos estadísticos específicos para el estudio de las evidencias de validez interna del WISC-IV en el que se comparan distintos modelos para la explicación de su estructura interna. Se utilizaron modelos de ecuaciones estructurales para analizar los objetivos propuestos. A continuación se describen los modelos operativos utilizados en este estudio.

Grupos de edad

Los datos se reagruparon en cuatro categorías de edad: 6-7 años / 8 a 10 años / 11 a 13 años / 14 a 16 años.

Modelo de Puntuaciones Compuestas del WISC-IV

Para el agrupamiento de los subtests, véase la Figura 1.

Modelos de Cattell-Horn-Carroll (CHC).

Modelo de Keith et al. (2006) y de Chen et al. (2009): Para el agrupamiento de cada uno de los modelos, véase la Figura 2.

Muestra

El estudio para la tipificación del WISC-IV fue diseñado para reflejar la población infanto-juvenil de las edades comprendidas entre los 6 y 16 años del Área Metropolitana de Buenos Aires (AMBA). En el diseño del plan de muestreo se consideraron tres variables para la estratificación, el sexo, la edad, y la zona.

Se consultaron los datos del Censo (Instituto Nacional de Estadísticas y Censos [INDEC], 2001) acerca de la población escolar de 6 a 16 años en cada región y, sobre esta base, se calculó para cada uno de los 11 grupos de edad la cantidad de casos necesarios según sexo, edad y región que debían recolectarse. Para cada grupo de edad, se consideran tres subgrupos clasificados de a tres meses, por ejemplo, 6:0 a 6:3/ 6:4 a 6:7/ 6:8 a 6:11, con cantidad semejante de casos en cada subgrupo.

Se establecieron como criterios de exclusión a los siguientes: (a) haber sido evaluado mediante cualquier instrumento de medición de inteligencia en los últimos 6 meses; (b) dificultades visuales y auditivas sin corregir o limitantes; (c) falta de fluidez verbal; (c) fallos severos en la expresión verbal o de comunicación; (d) problemas motores en extremidades superiores; (e) estar en tratamiento psiquiátrico, psicológico o psicofarmacológico y (f) diagnóstico previo de alguna enfermedad que pueda disminuir el desempeño en el test.

Entre diciembre de 2009 y noviembre de 2010 se recolectaron los 1485 casos estipulados pero 48 debieron ser desechados por datos incompletos o fallos en la administración o puntuación. Por tal razón la muestra final fue de 1437 casos. De estos, el 50.4% fueron niñas y el 49.6%, niños. La cantidad de casos por franja de edad osciló entre 120 y 142 y, en cuanto a la zona de residencia, el 34.9% fue de partidos de la zona sur, el 34.7% de la zona norte y el 30.2 de la zona oeste del Gran Buenos Aires.

Aunque no fue un criterio para la estratificación, es interesante notar los datos acerca del nivel educativo del adulto a cargo del niño—que, en lo usual, es la madre o el padre— ya que es un buen indicio del nivel socioeconómico del hogar. En nuestro estudio, consideramos aquellos casos en los que el registro de la educación se comparaba con las categorías globales de cantidad de años de estudios aprobados del INDEC (2004) ($n = 1053$) y se eliminaron aquellos cuyo registro fue confuso o incompleto ($n = 391$). Así, se relevaron 14 casos con menos de 4 años de estudios (1.4%), 84 casos con 4 a 6 (8.3%); 288 con entre 7 y 9 años de educación (27.3%); 344 con 10 a 12 años (32.6%) y 293 con más de 13 años de estudios (27.8%). El análisis de los datos sugiere una concordancia importante entre los datos promedio suministrados por el INDEC (4.5%, 6.1%, 34.3%, 31.4% y 25.4%, respectivamente) y los obtenidos en la muestra de tipificación, ya indicados. De esta manera puede considerarse que se ha logrado una buena representación de la población infanto-juvenil del AMBA, en especial en cuanto al sexo, la edad, el nivel educativo de los padres y la región de residencia.

Técnicas e instrumentos de recolección de datos

- *Escala de Inteligencia de Wechsler para Niños - Cuarta Edición (WISC-IV) en su adaptación para Buenos Aires (Taborda, Barbenza y Brenlla, 2011).* Se trata de un instrumento de administración individual para evaluar la inteligencia de niños o adolescentes de entre 6 años 0 meses y 16 años 11 meses (6:0 - 16:11). El WISC-IV consta de 10 subtests principales y 5 subtests optativos y otorga puntuaciones compuestas que proveen información sobre el funcionamiento intelectual en cuatro áreas cognitivas específicas (Puntuaciones Índice) y un Cociente Intelectual Total (CIT), tal como fue descripto en Figura 1.

Su fiabilidad y validez ha sido estudiada tanto en EE.UU (Wechsler, 2003) y España (Wechsler, 2005) como en Argentina, en los cuales se encontraron evidencias satisfactorias de las propiedades psicométricas del WISC-IV. En términos resumidos, para nuestro medio, las evidencias de consistencia interna—estimadas a partir del método de división por mitades— muestran coeficientes entre 0.79 y 0.90 y los de estabilidad test-retest, entre 0.77

y 0.86, así como adecuada discriminación entre grupos clínicos y evidencias satisfactorias de validez interna. Al igual que en el estudio de E.E. U.U., se observó que de los cinco modelos propuestos para el análisis de las puntuaciones del test, el más consistente resultó el de cuatro factores, debido a que el modelo de 5 factores no mejoró significativamente al de 4 y estuvo formado por un sola prueba (*Aritmética*). Para una descripción detallada de los procedimientos, consúltase la edición publicada por Paidós (Wechsler, 2011). Dichos procedimientos y resultados fueron evaluados y aprobados por el comité científico de la editora norteamericana.

Procedimiento

Entre diciembre de 2009 y noviembre de 2010 se recolectaron 1437 casos de sujetos de 6 a 16 años. Cada uno fue administrado individualmente por un evaluador entrenado a tal fin. Se llevaron a cabo cursos de capacitación para la administración y puntuación del test. Se solicitó el consentimiento del padre, madre o tutor del niño a fin de cumplir con los requerimientos éticos de investigaciones en psicología. Se registró el nivel educativo alcanzado por los padres o tutores así como los datos del niño incluidos en la Hoja de Registro.

La identificación de los niños que reunieran los requisitos de inclusión de la muestra de tipificación se realizó a través de reclutadores entrenados (70%) y de examinadores independientes (30 %), a cargo de tres grupos de distintas unidades académicas.

Los datos de la muestra de tipificación sirvieron para construir bases de datos en planillas Excel y bajo la plataforma del programa *Statistical Package for Social Sciences 18* (SPSS 18). En cuanto a las técnicas de análisis de datos, se utilizaron modelos de Análisis Factorial Confirmatorio (AFC), que permiten contrastar la validez ajustando un modelo teórico previamente estipulado y diagnosticando su bondad de ajuste a través de distintos índices. En este modelo, cada ítem satura únicamente sobre el factor-dimensión del que se supone que constituye un indicador válido. Para su cálculo se utilizó el programa AMOS 18 (Arbuckle, 2009).

Resultados

Los objetivos principales de este estudio fueron (a) valorar si el WISC-IV evalúa los mis-

mos constructos en distintos grupos de edad, tal como ocurre en la versión norteamericana, (b) analizar si la estructura de puntuaciones compuestas de cuatro dominios específicos y un factor general postulada por los autores del WISC-IV se replica en nuestro medio y (c) poner a prueba el modelo alternativo basado en la CHC propuesto por Keith et al. (2006) y por Chen et al. (2009). Para realizar una exposición ordenada, cada objetivo se brinda en forma de pregunta y luego se describen los resultados obtenidos.

¿La adaptación Buenos Aires del WISC-IV evalúa los mismos constructos en distintos grupos de edad, tal como ocurre en el estudio original?

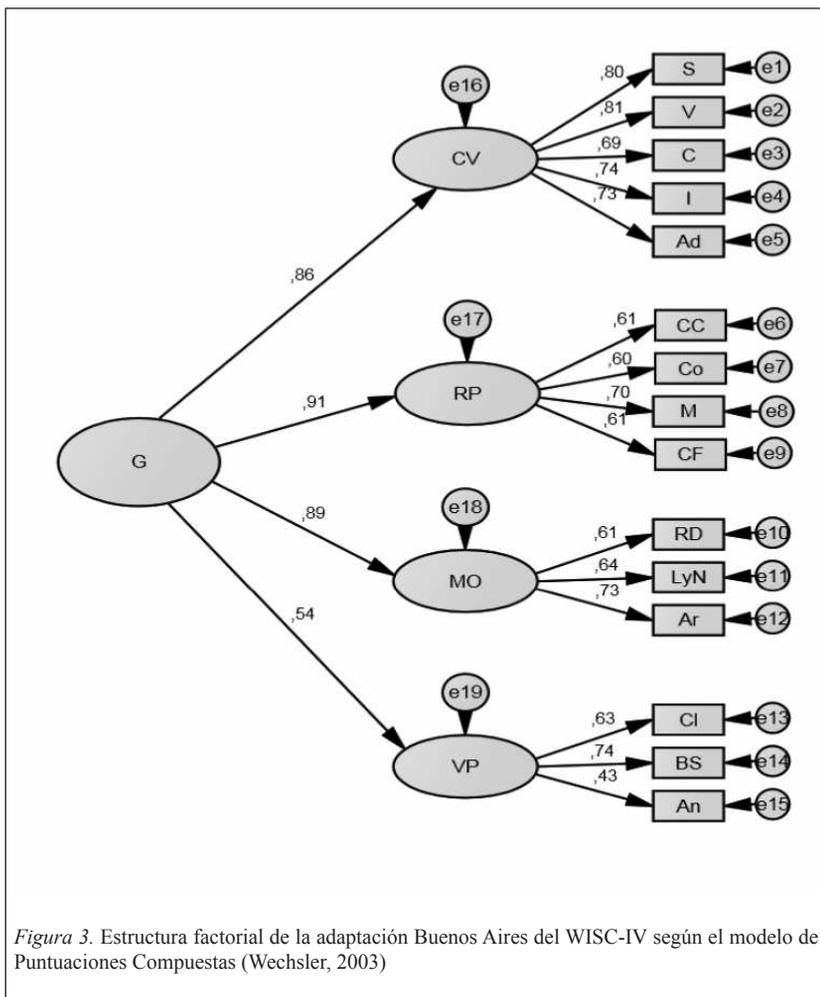
La primera cuestión que se analizó fue si los subtests del WISC-IV evalúan los mismos constructos en los distintos grupos de edad para lo cual se aplicaron métodos de AFC multigrupo. Al igual que en los estudios de Keith et al. (2006) y Chen et al. (2009), el primer paso consistió en comparar la equivalencia de las matrices de covarianza de los subtests según los rangos de edad considerados (6-7 años; 8 a 10 años; 11 a 13 años y 14 a 16 años). En un segundo paso se analizó la invariancia de un modelo estructural específico, para lo cual las matrices de varianza y de covarianza se calcularon a fin que sean similares en los distintos grupos de edad. El modelo resultante se acopló bien a los datos, ya que presentó buenos índices de ajuste (SRMR = .05; RMSEA = .05; TLI = .94; NFI = .92). Esto significa que, respecto de los subtests del WISC-IV, las correlaciones observadas en la muestra y las predichas por el modelo difieren en promedio solo en un .05 —que es la interpretación usual de SRMR—. Ello sugiere que, independientemente del nivel de edad, las puntuaciones que obtuvieron los niños en el WISC-IV expresan claros e invariantes patrones de correlaciones entre los subtests. En este sentido los hallazgos en cuanto a la estructura interna del WISC-IV son aplicables a todas las edades y refrendan la estructura de puntuaciones compuestas propugnada por los autores originales del test (Wechsler, 2003) y otros alternativos que presenten evidencias sólidas de validez.

¿La adaptación Buenos Aires del WISC-IV evalúa cuatro dominios diferenciados y un factor general de inteligencia, tal como se propone en el estudio original?

Tal como se indicó en la introducción, el WISC-IV propone que la interpretación del test se realice sobre la base de las puntuaciones compuestas de cociente intelectual e índices de dominios específicos. Para comparar si esta estructura de puntuaciones es válida en nuestro medio, se utilizaron técnicas de AFC para analizar el comportamiento de los subtests principales y optativos en la muestra de Buenos Aires. A diferencia de lo realizado en oportunidad de los estudios de validez anteriores (Taborda et al., 2011), en los que se trabajó con un análisis de factores relacionados, en este caso, se analizó un modelo jerárquico constituido por un factor general, cuatro factores correspondientes a CV, RP, MO y VP y, como variables observadas, las puntuaciones a los 15 subtests del WISC-IV.

En la Figura 3 se muestra la solución hallada para el modelo de 4 factores mediante la aplicación del método de análisis factorial confirmatorio (AFC).

Puede notarse que los coeficientes de regresión estandarizados confirman los factores que se conjeturaron (CV, RP, MO y VP) y que los que tienen mayor carga en *g* son los de Comprensión Verbal ($\beta = .86$), Razonamiento perceptivo ($\beta = .91$) y Memoria operativa ($\beta = .89$) en tanto que, Velocidad de procesamiento ($\beta = .54$), es el de menor carga en *g*. Es de resaltar que todos los coeficientes son muy altos y que las cargas factoriales de los subtests se ajustan con nitidez a la estructura de puntuaciones compuestas ideada en el WISC-IV. Por ejemplo, los coeficientes de los subtests que componen el índice de Comprensión verbal tienen cargas factoriales que oscilan entre .69 a .81; los de Razonamiento perceptivo, entre .60 a .70, los de Memoria operativa, de .61 a .73 y los de Velocidad de procesamiento, entre .43 a .74. Además, todas las asociaciones especificadas por el modelo mostraron un peso significativo ($p < .001$) sobre el factor en cuestión.



Para la evaluación de la bondad de ajuste, se analizaron estadísticos de ajuste absoluto, relativo y parsimonioso. Entre los primeros, las medidas más utilizadas son el estadístico χ^2 o índice de razón de verosimilitud y el índice *goodness-of-fit* (GFI). El primero solo indica la significación estadística del modelo completo y permite contrastar la hipótesis nula de que todos los errores de éste son nulos, vale decir, que hay un ajuste total entre el modelo teórico y la muestra utilizada. Por tanto es deseable no rechazar la hipótesis nula. El modelo produjo una χ^2 significativa (384.250, $p < .001$). Como es conocido, el índice χ^2 raramente es utilizado como prueba única o concluyente de bondad del ajuste de un modelo ya que se basa en la hipótesis excesivamente restrictiva de que S es igual a Σ . Además, en muchos casos, la distribución de los datos no se ajusta a una distribución χ^2 o su valor está influenciado por el tamaño de la muestra (Arias, 2008; Tanaka, 1993; The Structural Equation Modeling Discussion Network [SEMNET], 2010). Por ello, muchos autores (Arbuckle, 2009; Tanaka, 1993) recomiendan examinar no solo la significación sino también la magnitud del χ^2 a través del índice χ^2/df , donde los grados de libertad se utilizan como un estándar para juzgar si ese valor de χ^2 es grande o pequeño. Se consideran excelentes los valores entre 1 y 2 (Stahl, 2010), buenos los menores a 3 (Tabachnik & Fidell, 2001), aceptables a los comprendidos entre 3 y 5, sobre todo si el análisis se realiza sobre muestras grandes como es este caso y, rechazables, si son mayores que 5 (Bentler & Bonnet, 1980). De acuerdo a estos criterios, este índice reflejó valores aceptables (4.46), en la muestra total ($n = 1437$) y, cuando se replicó el mismo análisis según los grupos de edad – que incluían menos casos –, los valores obtenidos fueron muy buenos ya que oscilaron entre 1.5 y 2.5.

Por su lado, el GFI mide la cantidad de varianza y covarianza en la matriz observada que está predicha por la matriz reproducida. Se recomienda que su valor sea igual o mayor a .95 (Arbuckle, 2009; Ruiz, Pardo y San Martín, 2010), por lo que el valor obtenido (.96) puede ser considerado muy bueno. El otro índice de ajuste absoluto que se evalúa es el AGFI (*adjusted goodness-of-fit*). El AGFI es el índice GFI ajustado por los grados de libertad y permite comparar el modelo propuesto con un modelo nulo que se toma de referencia. Este modelo

nulo estipula la falta absoluta de asociación entre las variables y es calculado por el programa AMOS por defecto. El AGFI indica la proporción de la varianza explicada por el modelo y su valor de referencia es de $\beta.95$ (Arbuckle, 2009; Tabachnik & Fidell, 2001). Como puede notarse el valor obtenido (.95) indica un muy buen ajuste según los criterios establecidos.

El RMSR (*Root mean square residual*), es la raíz cuadrada de la media de los residuos entre los elementos correspondientes a la matriz observada y a la reproducida. Si los residuos se aproximan a 0, el valor de RMSR será cero que se interpreta como un ajuste perfecto. Por lo general, los valores de $\beta.05$ se consideran “premonitorios” de un buen ajuste aunque se consideran aceptables –sobre todo con muestras grandes– los valores ≤ 08 . En este caso, el valor fue de .08, que indica un ajuste aceptable entre el modelo propuesto y la matriz de observaciones. El RMSEA es un índice del error de aproximación medio cuadrático a valores de la población. De acuerdo con Arbuckle (2009) los valores de ≤ 08 indican que el modelo ajusta apropiadamente. Así, los índices obtenidos (.047; .044 – .054 $p .670$) indican que existe un 67% de probabilidad de que el valor real sea menor a .05.

El NFI es un índice de ajuste comparativo y se interpreta como un porcentaje de incremento en la bondad de ajuste sobre el modelo nulo. Por ejemplo, un NFI de .90 sugiere que el modelo propuesto ajusta 90% mejor que el nulo. Usualmente se toma como valor de referencia un NFI ≥ 95 , por lo que el índice obtenido (.95), refleja un buen ajuste. De igual manera, el *Comparative Fit Index* (CIF) – análogo al anterior – mostró valores considerados de buen ajuste (.96). En conjunto, los resultados obtenidos a partir del AFC indican que la estructura de puntuaciones compuestas de la adaptación argentina del WISC-IV se ajusta a la idea de que el test evalúa un factor general de la inteligencia y cuatro dominios específicos, tal como fue propuesto en la edición original (Wechsler, 2003).

¿La adaptación Buenos Aires del WISC-IV evalúa las habilidades de acuerdo a la teoría CHC?

Para contestar a esta pregunta se realizaron AFC de las puntuaciones de la muestra de tipificación argentina reagrupando los ítems según los dominios supuestos por la CHC.

Se tuvo en cuenta tanto el modelo de Keith et al. (2006) como el de Chen et al. (2009) que se ilustraron en la Figura 2.

El primer modelo no produjo una solución satisfactoria. Se registraron varianzas negativas y la inspección de los coeficientes beta estandarizados indicó que el subtest Matrices presentó cargas negativas y exiguas en razonamiento fluido (-.03) y cargas positivas y significativas en procesamiento visual (.75), estimación contraria al modelo propuesto por Keith et al.(2006). Estos rasgos hacen que la solución hallada no sea admisible.

En cambio, el modelo propuesto por Chen et al. (2009) presentó un buen ajuste. Todos los coeficientes estandarizados van en la dirección esperada, sus valores son, por lo general, medios o altos y estadísticamente significativos. En la figura 4 se muestra la solución hallada. Se observa que los coeficientes de regresión estandarizados confirman los factores que

se conjeturaron (procesamiento visual; razonamiento fluido; habilidad cristalizada; memoria y velocidad de procesamiento) y que los que tienen mayor carga en *g* son Procesamiento visual ($\beta = .96$), Razonamiento fluido ($\beta = .95$), Habilidad Cristalizada ($\beta = .80$) y Memoria ($\beta = .79$) mientras que el de menor carga es Velocidad de Procesamiento ($\beta = .46$).

Los índices de bondad de ajuste calculados son similares a los obtenidos para el modelo de puntuaciones compuestas del WISC-IV ($\chi^2/gf = 4.15$, $GFI = .96$, $AGFI = .95$, $RMSEA = .04$ [$LO90 = .042$ / $HI90 = .052$, $p = .833$], $RMSR = .07$, $NFI = .95$ y $CFI = .97$). En la Tabla 1 se brindan los datos de ambos modelos: puntuaciones WISC-IV y Modelo CHC de Chen et al. (2009).

A pesar de esta similitud en cuanto al ajuste empírico, hay que señalar que en el modelo CHC alternativo las cargas factoriales son menos nítidas que las observadas en el del WISC-IV. Por ejemplo, el modelo CHC supone que las

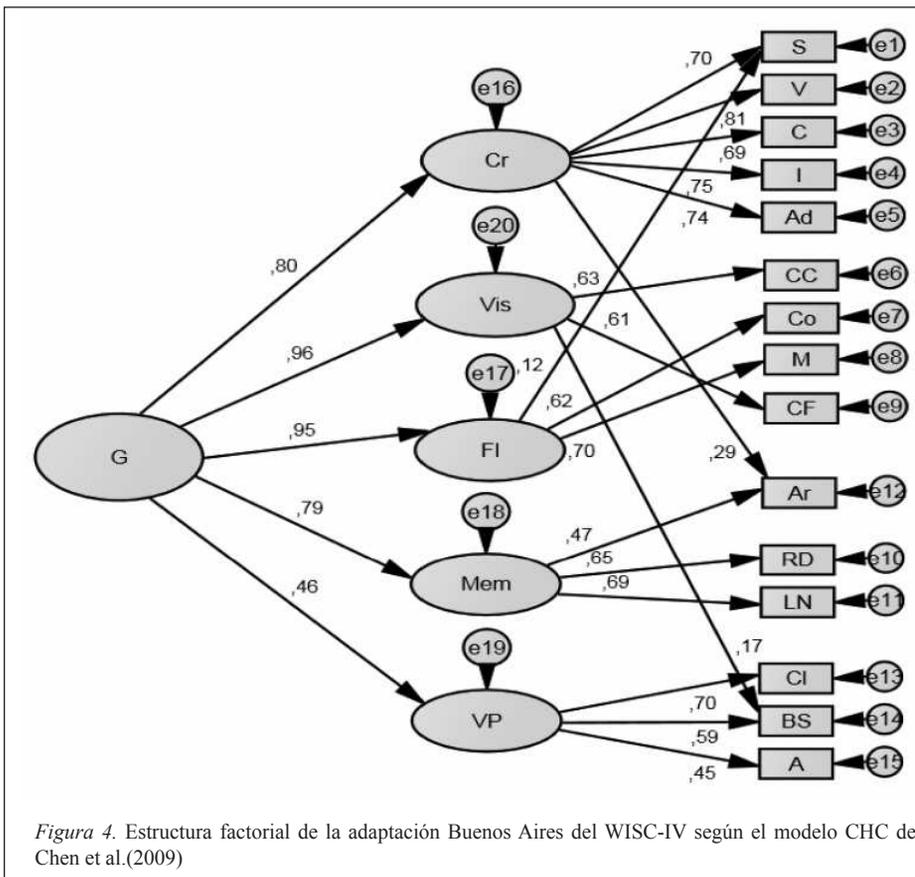


Figura 4. Estructura factorial de la adaptación Buenos Aires del WISC-IV según el modelo CHC de Chen et al.(2009)

Tabla 1
Estadísticos de bondad de ajuste modelos WISC-IV y CHC alternativo

Modelo	χ^2	gl	GFI	AGFI	NFI	CFI	RMSEA	SRMR
WISC-IV Puntuaciones	384.250	86	.96	.95	.95	.96	.04	.08
CHC Modelo Chen et al.	341.008	82	.96	.95	.95	.97	.04	.07

tareas de Búsqueda de símbolos tendrán una carga compartida entre el procesamiento visual y la velocidad de procesamiento. Si bien los resultados así lo indican, los coeficientes son bajos para el primero ($\beta = .17$) pero sustantivos para el segundo ($\beta = .59$). De manera análoga, Aritmética tiene una carga baja en inteligencia cristalizada ($\beta = .29$) y más alta para Memoria ($\beta = .47$) y, Semejanzas, cargas muchos más importantes en Habilidad cristalizada que en Razonamiento fluido ($\beta = .70$ y $\beta = .12$, respectivamente). Además, hay que notar que Matrices –que en el modelo de Keith se agrupaba con Aritmética y Conceptos y que había presentado varianza negativa con Razonamiento fluido– registró coeficientes altos y significativos en este dominio ($\beta = .70$).

¿Qué evalúan los subtests de la adaptación Buenos Aires del WISC-IV?

Como se indicó, tanto el modelo de puntuaciones compuestas del WISC-IV como el modelo CHC de Chen et al. (2009) reflejaron buenos o muy buenos índices de ajuste a los supuestos teóricos establecidos *a priori*. En consecuencia ambos proveen inferencias teóricas válidas acerca de la estructura interna del WISC-IV. Dado que ello tiene importancia a la hora de interpretar el test, a continuación se describen cuáles son los procesos involucrados en la resolución de las tareas de los subtests del WISC-IV a fin que el profesional disponga de una guía basada en el estudio de los dos modelos alternativos de inteligencia en la muestra de tipificación de Buenos Aires.

Semejanzas: En virtud del análisis de los modelos WISC-IV y CHC, puede considerarse que este subtest mide, primariamente, la comprensión verbal y la inteligencia cristalizada y, secundariamente, el razonamiento fluido.

Vocabulario, Comprensión, Información y Adivinanzas: comprensión verbal, inteligencia cristalizada.

Construcción con Cubos: razonamiento perceptivo, específicamente, procesamiento visual.

Completamiento de Figuras: razonamiento perceptivo y, secundariamente, procesamiento visual.

Conceptos y Matrices: razonamiento perceptivo, particularmente, el razonamiento fluido.

Retención de Dígitos, Letras y Números: memoria operativa y de corto plazo.

Aritmética: memoria operativa y de corto plazo y, secundariamente, la habilidad cristalizada.

Clave y Animales: velocidad de procesamiento.

Búsqueda de Símbolos: velocidad de procesamiento y, secundariamente, procesamiento visual.

Discusión y Conclusiones

Los resultados a los que se arribó en este estudio permiten decir que tanto la estructura de puntuaciones compuestas propuesta por los autores del WISC-IV como el modelo alternativo CHC resultan igualmente significativos para interpretar las puntuaciones del test y que las inferencias son aplicables a todas las edades consideradas –6 a 16 años–.

En este sentido, el estudio factorial confirmatorio realizado en Buenos Aires replica los hallazgos del estudio original en cuanto a la consistencia del modelo de cuatro factores en distintos grupos de edad (Wechsler, 2003), coincide con el llevado a cabo en Taiwan (Chen et al., 2009) en el que tanto el modelo propuesto en el WISC-IV y el basado en la teoría CHC presentan ajustes satisfactorios y difiere del estudio de Keith et al. (2006), que sostiene que otro modelo inspirado en la CHC es más sustantivo que el de cuatro factores para la interpretación del WISC-IV.

En particular, los resultados de nuestro estudio revelan que los subtests Aritmética,

Matrices, Completamiento de figuras y Búsqueda de símbolos presentan cargas factoriales compartidas en ambos modelos. Aritmética puede considerarse primariamente como una medida de memoria operativa tal como postula el modelo WISC-IV pero también de habilidad cristalizada. Matrices, refleja primariamente el razonamiento perceptivo y el razonamiento fluido mientras que Búsqueda de símbolos y Completamiento de figuras expresan la habilidad en tareas de procesamiento visual. De este modo, los resultados de este estudio avalan la idea de una visión integrativa entre los modelos WISC-IV y CHC para la interpretación de las puntuaciones del test.

Además los resultados sugieren que el razonamiento fluido es, al igual que los cuatro dominios propuestos en el WISC-IV, una excelente medida de *g* en tanto que el factor de Velocidad de Procesamiento es el que menos contribuye a su explicación. Entre los subtests, Semejanzas, Vocabulario, Construcción con cubos, Conceptos, Aritmética, Matrices, Letras y Números y Retención de dígitos representan una porción importante de la varianza explicada y, en ambos modelos, Animales es el de menor peso explicativo.

Así como se han encontrado estas coincidencias sustanciales con las investigaciones de la versión original del WISC-IV (Wechsler, 2003) y con el estudio de Taiwan (Chen et al., 2009), también se hallaron discrepancias con otros. En especial, algunos resultados son claramente diferentes de los informados por Keith et al. (2006) para la muestra norteamericana. El desmembramiento de la medida de Razonamiento perceptivo en los factores de Procesamiento visual y de Razonamiento fluido propugnado por Keith en su modelo ajusta satisfactoriamente en la población infantil de EE.UU pero no en la muestra de Buenos Aires ni en la de Taiwan. Si bien se observó en los dos estudios que Construcción con cubos y Completamiento de figuras cargan sobre Procesamiento visual y Matrices y Conceptos sobre Razonamiento fluido no obstante, las puntuaciones obtenidas por los niños revelan una correlación más importante y significativa entre ambos que la observada en el país de origen. Tomando un argumento de Chen et al. (2009) podría conjeturarse que cuando a los niños se les presenta un estímulo visual, podría activarse casi automáticamente la pizarra viso-espacial del sistema de memoria operativa (Baddeley, 2003) y poner en marcha

así los procesos del razonamiento fluido. De esta manera, tanto la interpretación basada en el modelo WISC-IV –razonamiento perceptivo– como el modelo alternativo CHC de Buenos Aires y Taiwan –que propone la consideración de dos factores relacionados– son igualmente sustantivos para el análisis de las puntuaciones.

Otra discrepancia de interés es la encontrada respecto de Completamiento de figuras (CF). Keith et al. (2006) reportaron cargas cruzadas de este subtest en Procesamiento visual (.42) y en Habilidad cristalizada (.31) mientras que, en nuestro estudio, las cargas fueron importantes para el primero (.61) y no significativas para la segunda. Como se indicó, en EE.UU. los valores son cercanos y moderados en cambio, en las muestras de Buenos Aires y de Taiwan, la carga principal recae sobre el procesamiento visual pero no sobre la inteligencia cristalizada. Como recordará el lector, en CF el entrevistado tiene que decir que parte le falta a un dibujo que se presenta en láminas de colores por lo que la tarea propuesta requiere de concentración, memoria, análisis y síntesis visuales y conocimiento visual adquirido. En este sentido, la tarea se comparece con la definición de McGrew (2009) acerca del procesamiento visual como la “habilidad para generar, almacenar, recuperar y transformar sensaciones e imágenes visuales” más que con una asociación con la inteligencia cristalizada como indicaba el estudio de Keith et al. (2006). Además, el hecho que la carga fuera nítida coadyuva a considerar que CF es una tarea definida de procesamiento visual no-verbal. Por lo tanto, los resultados de Buenos Aires y Taiwan en cuanto a la pregnancia del procesamiento visual en CF son teóricamente consistentes y pueden contribuir a la comprensión de los procesos psicológicos implicados en este subtest.

De acuerdo con el citado estudio de Keith, en el subtest de Aritmética (Ar), los niños norteamericanos puntúan de manera tal que puede considerarse como una medida del razonamiento fluido. Si bien ello coincide en parte con el análisis realizado en nuestro medio, también es verdad que los resultados muestran ciertas diferencias cuando se analiza en forma conjunta la influencia de otros factores. Tanto en la muestra de Buenos Aires como en la de Taiwan, Aritmética mostró cargas compartidas con la Inteligencia cristalizada (.29 y .30, respectivamente) y con la Memoria de corto plazo (.47 y .51, en el mismo orden), por lo que estos

factores parecen ser más importantes que otros en su explicación. Estas discrepancias respecto de qué evalúa Ar ya se habían identificado en estudios previos con el WISC-III, donde se mostró que este subtest evaluaba razonamiento fluido pero también inteligencia cristalizada y memoria de corto plazo (Keith & Witta, 1997). Además, como señalan Keith et al. (2006) en las limitaciones de su estudio, el hecho de no haber contado con medidas específicas de razonamiento numérico (RN) pudo haber distorsionado los resultados ya que la evidencia previa ha mostrado que cuando se incluyen medidas de RN –por ejemplo, rendimiento en matemáticas-Aritmética carga en razonamiento numérico pero no en fluido.

Keith et al. (2006) indican que diversas condiciones pueden afectar el rendimiento en este subtest, tales como los programas escolares de matemáticas o el nivel de sofisticación para la resolución de problemas. Hay que recordar que la inteligencia cristalizada se relaciona significativamente con el conocimiento adquirido en el sistema escolar en tanto que el razonamiento fluido implica manipular abstracciones, reglas, generalizaciones y relaciones lógicas. Según estos datos, los mecanismos y procesos que subyacen a este subtest podrían ser diferentes para los niños de EE.UU. y de las muestras de Buenos Aires y de Taiwan. Así, las puntuaciones en Aritmética en esta investigación parecen depender más de los conocimientos adquiridos y el entrenamiento previo que de la aplicación de leyes lógicas.

¿Qué representan estos datos? ¿qué la resolución de estos problemas se relaciona con conocer los procedimientos específicos que se aprenden en la escuela, con la comprensión verbal y con el procesamiento auditivo de los enunciados más que con procesos de deducción o inducción lógicas? Sea cual fuere la respuesta, lo interesante es que en culturas tan distintas como la argentina y taiwanesa se encontraron similitudes en la estructura factorial del WISC-IV que difieren en algunos puntos importantes con la hallada en el estudio norteamericano.

Si bien los análisis factoriales son de importancia para clarificar la naturaleza de los constructos que se pretenden evaluar no hay que perder de vista, tal como indican Prifitera, Weiss, Saklofske y Rohlfus (2005), que el análisis factorial es solo una herramienta para conocer cuál es la mejor manera de interpretar las relaciones entre los subtests y que, cuando se seleccio-

nan o eligen ciertos factores, ello debería estar orientado por su potencial utilidad clínica y no por un modo mecanizado de interpretar el test. Por ejemplo, cuando el rendimiento de un niño en las pruebas de razonamiento perceptivo sea errático o inconsistente, el profesional debería revisar las puntuaciones considerando su consistencia con los dominios de razonamiento fluido o de procesamiento visual.

En términos globales, los resultados reseñados en este artículo proveen información sólida para comprender la estructura interna del WISC-IV y su relación con los procesos cognitivos en distintas culturas. Si bien en este trabajo se analizó una muestra amplia y representativa de niños residentes en el Área Metropolitana de Buenos Aires –cuya población representa el 32% de la población argentina–, no obstante se considera como una limitación el que no abarque las otras regiones del país. Asimismo, otra limitación es que este estudio se focalizó en las habilidades del estrato general y amplio indicadas por la CHC pero no exploró las habilidades limitadas o próximas.

Como conclusión puede señalarse que tanto la estructura de cuatro factores propuesta por los autores del WISC-IV (Wechsler, 2003) como el modelo alternativo basado en la teoría CHC son enfoques sustantivos para realizar la interpretación del WISC-IV en nuestro medio. Además, los resultados de este estudio coadyuvan a la comprensión de cuáles son los constructos principales que evalúa el WISC-IV y qué particularidades y diferencias se encontraron respecto de los estudios de origen. En consonancia con ello, al realizar la interpretación de las puntuaciones del WISC-IV, es importante notar no solo las similitudes sino también las discrepancias respecto de las habilidades cognitivas subyacentes en distintas culturas.

Referencias

- Arbuckle, J. L. (2009). *Amos 18 User's Guide*. Chicago, IL.: SPSS Inc.
- Arias, B. (2008). *Desarrollo de un ejemplo de análisis factorial confirmatorio. Seminario de actualización en investigación sobre discapacidad SAID 2008*. Recuperado de <http://www.benitoarias.com/articulos/afc.pdf>
- Baddeley, A. (2003). Working Memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829-839.
- Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.

- Blaaha, J., & Wallbrown, F. H. (1996). Hierarchical factor structure of the Wechsler intelligence scale for children-III. *Psychological Assessment*, 8, 214-218.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, England: Cambridge University Press.
- Carroll, J. B. (1997). The three-stratum theory of cognitive abilities. En D. P. Flanagan, J. L. Genshaft & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 122-130). Nueva York: Guilford Press.
- Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin*, 38, 592.
- Cattell, R.B. (1957). *Culture fair intelligence test, a measure of "g": Scale 3, forms A and B (high school pupils and adults of superior intelligence)*. Savoy, IL: Institute for Personality and Ability Testing
- Chen, H.Y., Keith, T.Z., Chen, Y.H., & Chang, B.S. (2009). What does the WISC-IV measure? Validation of the Scoring and CHC-based Interpretative Approaches. *Journal of research in Education Sciences*, 54 (3), 85-108.
- Cohen, R. J., y Swerdlik, M.E. (2000): *Pruebas y Evaluación Psicológicas*. México: McGraw Hill
- Gustafsson J. E., & Undheim J. O. (1996). Individual differences in cognitive functions. En D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 186-242), Nueva York: MacMillan.
- 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: Wiley.
- Horn, J. L. (1988). Thinking about human intelligence. En J. R. Nesselroad & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (pp. 645-685). Nueva York: Plenum Press.
- Horn, J. L. (1991). Measurement of intellectual capabilities: A review of theory. En K. S. McGrew, J. K. Werder & R. W. Woodcock (Eds.), *Woodcock-Johnson technical manual* (pp. 197-232). Itasca, IL: Riverside.
- Horn, J. L. & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. En D. P. Flanagan, J. L. Genshaft & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53-91), Nueva York: Guilford Press.
- Instituto Nacional de Estadísticas y Censos. (2001). *Censo nacional de población, hogares y vivienda 2001* Buenos Aires: Autor.
- Instituto Nacional de Estadísticas y Censos. (2004). Situación educacional de la población de Argentina según el Censo 2001. *Revista informativa del censo 2001*. Número 1.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger Publishers.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8 User's Reference Guide*. Chicago: Scientific Software International.
- 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., Reynolds, M.R., & Kranzler, J.H. (2006). Higher Order, Multisample, Confirmatory Factor Analysis of the Weschler Intelligence Scale for Children-Fourth Edition: What does It Measure? *School Psychology Review*, 35 (1), 108-127.
- Keith, T. Z., & Witta, E. L. (1997). Hierarchical and cross-age confirmatory factor analysis of the WISC-III: What does it measure? *School Psychology Quarterly*, 12, 89-107.
- Marín, G. (1986). Metodología de la Investigación Psicológica. *Acta Psiquiátrica y Psicológica de América Latina*, 32, 3-5.
- Martínez Arias, R. (1996) *Psicometría: Teoría de los Tests Psicológicos y Educativos*. Madrid: Síntesis
- McGrew, K., & Flanagan, D. (1998). *The intelligence test desk reference (ITDR): The Gf-Gc cross-battery assessment*. Boston: Allyn & Bacon.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, a Multidisciplinary Journal*, 37, 1-10.
- Prifitera, A., Weiss, L.G., Saklofske, D.H., & Rolfhus, E. (2005). The WISC-IV in the clinical assessment text. En A. Prifitera, D.H. Saklofske & L.G. Weiss (Eds.), *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 3-32). San Diego, CA: Academic Press.
- Roid, G. H. (2003). *Stanford-Binet Intelligence Scales, Fifth Edition (SB:V)*. Itasca, IL: Riverside Publishing.
- Ruiz, M. A., Pardo, A. y San Martín, R. (2010). Modelos de ecuaciones estructurales. *Papeles del psicólogo*, 31(1), 34-45.
- Sattler, J. M. (2001). *Assessment of children: Cognitive applications (4ª ed.)*. San Diego, CA: Author.
- Schneider, W. J., & McGrew, K. (2012) The Cattell-Horn-Carroll model of intelligence. En D. Flanagan & P. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues (3ª ed.)*. (pp. 99-144) Nueva York: Guilford.
- Spearman, C. (1904). The Proof and Measurement of Association between two things. *American Journal of Psychology*, 11, 72-73.
- Spearman, C. (1927). *The Nature of Intelligence and the Principles of Cognition*. Nueva York: MacMillan.
- Stahl, D. (2010). *Introduction to Path Analysis and Structural Equation Modelling with AMOS Biostatistics and Computing*. Recuperado de <http://www.kcl.ac.uk/iop/depts/biostatistics/teaching/previouslecturenotes/amos/amosweek2.pdf>
- Tabachnick, B.G., & Fidell, L.S. (2001). *Using Multivariate Statistics, Fourth Edition*. Needham Heights, MA: Allyn & Bacon.
- Taborda, A.R., Brenlla, M.E., & Barbenza, C. (2011). Adaptación argentina de la Escala de Inteligencia de Wechsler para Niños IV (WISC-IV) En D. Wechsler (Ed.), *Escala de Inteligencia de Wechsler para Niños IV (WISC-IV)* (pp.37-55). Buenos Aires: Paidós.
- Tanaka, J.S. (1993). Multifaceted conceptions of fit in structural equation models. En K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models*. (pp. 10-39) Newbury Park, CA: Sage.
- The Structural Equation Modeling Discussion Network (2010). The Structural Equation Modeling Discussion Network Archives. Disponible en <http://www2.gsu.edu/~mkteer/semnet.html>.
- Triandis, H. C., Lambert, W. W., Berry, J. W., Brislin, R. W., Draguns, J., Lonner, W., & Heron, A. (1980). *Handbook of cross-cultural psychology (6 vol.)*. Boston, MA: Allyn & Bacon.

- Wechsler, D. (1939). *The Measurement of Adult Intelligence*. Baltimore, MD: Williams & Wilkins.
- Wechsler, D. (1991). *Wechsler intelligence scale for children-third edition*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (2003). *Wechsler intelligence scale for children-fourth edition*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (2004). *Wechsler Scale of Intelligence, Fourth ed*. London: Pearson Assessment
- Wechsler, D. (2005). *Escala de Inteligencia Wechsler para Niños IV. (WISC-IV)*. Madrid: TEA.
- Wechsler, D. (2011). *Escala de Inteligencia de Wechsler para Niños IV (WISC-V)*. Buenos Aires: Paidós.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Test*. Riverside Publishing Company. Itasca, IL. Terry L. Blackwell.

Para citar este artículo:

Brenlla, M.E. (2013). Interpretación del WISC-IV. Puntuaciones compuestas y Modelos CHC. *Ciencias Psicológicas VII* (2): 183 - 197.

Recibido: 12/2012
Revisado: 06/2013
Aceptado: 08/2013