

# Estandarización de la Escala Wechsler de Inteligencia Para Adultos-Cuarta Edición en Chile

## Standardization of the Wechsler Intelligence Scale for Adults-Fourth Edition in Chile

Ricardo Rosas, Marcela Tenorio, Marcelo Pizarro, Patricio Cumsille y Angélica Bosch  
Pontificia Universidad Católica de Chile

Susana Arancibia  
Universidad Católica del Norte

Marcos Carmona-Halty  
Universidad de Tarapacá

Claudia P. Pérez-Salas  
Universidad de Concepción

Eduardo Pino  
Universidad de Magallanes

Beatriz Vizcarra  
Universidad de la Frontera

Pamela Zapata-Sepúlveda  
Universidad de Tarapacá

La Escala Wechsler de Inteligencia para Adultos en su cuarta edición (WAIS-IV) es el instrumento por excelencia que permite la evaluación cognitiva de adultos entre los 16:00 y 90:11 años. Dada la necesidad de adaptar este tipo de instrumentos en función de particularidades culturales, se realizó la estandarización del instrumento para la población chilena. Se utilizó una muestra intencionada de 887 sujetos, estratificados por edad, nivel educativo y género, que se agruparon en 7 conglomerados según región geográfica. A partir de los parámetros entregados en los Estándares para Evaluación Psicológica y Educativa (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1999), se realizaron análisis de consistencia interna y validez. La consistencia interna de la prueba es 0,941, evaluada con alfa de Cronbach, registrándose niveles adecuados de consistencia interna de cada subprueba. Hay evidencia de validez de contenido, de constructo y convergente. Se encontró sesgo en las pruebas verbales a favor de los niveles socioeconómicos más aventajados. La principal limitación del estudio es la ausencia de evidencia en grupos clínicos.

*Palabras clave:* evaluación, inteligencia, cognición, estandarización, WAIS-IV

The Wechsler Intelligence Scale for Adults, fourth edition (WAIS-IV) is the gold standard for assessing cognitive abilities in adults from 16:00 to 90:11 years of age. Given the need to adjust this instrument to specific cultural characteristics, its standardization for Chilean population was undertaken. A purposive sample of 887 subjects was used, stratified by age, educational level and gender, and grouped into 7 clusters by geographical region. Upon the basis of the parameters set out in the Standards for Psychological and Educational Assessment (American Educational Research Association et al., 1999), analyses of internal consistency and validity were performed. The internal consistency of the test is 0.941, analyzed with Cronbach's alpha. Each subtest shows good levels of internal consistency. There is evidence of content, construct, and convergent validity. Bias was found in the verbal tests in favor of people with better socio-economic status. The main limitation of the study is the absence of evidence in clinical groups.

*Keywords:* assessment, intelligence, cognition, standardization, WAIS-IV

---

Ricardo Rosas y Marcela Tenorio, Escuela de Psicología y Centro de Desarrollo de Tecnologías de Inclusión (CEDETi UC), Pontificia Universidad Católica de Chile, Santiago, Chile; Marcelo Pizarro, Centro de Desarrollo de Tecnologías de Inclusión (CEDETi UC), Pontificia Universidad Católica de Chile, Santiago, Chile; Patricio Cumsille, Escuela de Psicología, Pontificia Universidad Católica de Chile, Santiago, Chile; Angélica Bosch, MIDE UC, Pontificia Universidad Católica de Chile, Santiago, Chile; Susana Arancibia, Escuela de Psicología, Facultad de Humanidades, Universidad Católica del Norte, Antofagasta, Chile; Marcos Carmona, Departamento de Ciencias Sociales, Universidad de Tarapacá, Iquique, Chile; Claudia P. Pérez-Salas, Departamento de Psicología, Universidad de Concepción, Chile; Eduardo Pino, Escuela de Psicología, Universidad de Magallanes, Punta Arenas, Chile; Beatriz Vizcarra, Escuela de Psicología, Universidad de la Frontera, Temuco, Chile; Pamela Zapata-Sepúlveda, Unidad Interdisciplinaria de Investigaciones Psicosociales y Jurídicas, Departamento de Filosofía y Psicología, Universidad de Tarapacá, Arica, Chile.

El presente trabajo se enmarca en el proyecto FONDEF D09I1238, fondo que financió la investigación.

La correspondencia relativa a este artículo debe ser dirigida a Ricardo Rosas D., Centro de Desarrollo de Tecnologías de Inclusión (CEDETi UC), Pontificia Universidad Católica de Chile, Edificio MIDE UC 2° piso, Macul, Santiago, Chile. E-mail: rrosas@uc.cl

## **Evolución de las Escalas Wechsler de Inteligencia**

Wechsler creó su primera escala de inteligencia a partir de pruebas desarrolladas por Binet y un grupo de psicólogos que trabajaron durante la Primera Guerra Mundial. El interés fue ofrecer un instrumento que permitiera aproximarse dinámicamente a la inteligencia, con alto impacto clínico (Lichtenberger & Kaufman, 2009).

La primera escala diseñada por este autor fue la *Wechsler-Bellevue Intelligence Scale* (Wechsler, 1939) que permitía la evaluación de jóvenes y adultos. En esta escala se recuperaron pruebas que habían sido diseñadas sin un propósito clínico para modificarlas y unir las hasta alcanzar una evaluación dinámica de la cognición. Este diseño incorporó pruebas verbales y no verbales con la idea de que esta división fuera una mínima condición para una evaluación factorial de inteligencia.

La primera versión de la Escala de Inteligencia para Adultos de Wechsler (WAIS; Wechsler, 1955) incluyó una escala verbal y una manipulativa, que en conjunto entregaban un Coeficiente Intelectual Total (CIT). La revisión de esta escala (WAIS-R; Wechsler, 1981), contempló una actualización de las normas e incorporación de nuevos ítems que buscaban disminuir sesgos sociales y raciales (Kaufman, 1994), pero conservó su estructura factorial. WAIS-III (Wechsler, 1997) tuvo actualizaciones de las subpruebas y normas; también incorporó nuevas herramientas de análisis factorial que permitieron un salto en la comprensión de la inteligencia, al identificar dos escalas generales y cuatro índices subyacentes. En la elaboración de WAIS-IV (Wechsler, 2008) se incorporaron investigaciones sobre inteligencia, desarrollo cognitivo y neurociencias (Coalson & Raiford, 2008). Los modelos matemáticos actuales permitieron avanzar en dos sentidos: (a) se identificó una estructura de cuatro factores que abandona la antigua división verbal-no verbal y (b) se incorporó un algoritmo de toma de decisiones que permite la evaluación de procesos cognitivos (Coalson & Raiford, 2008; Wechsler, 2008).

## **Adaptaciones en Chile de las Escalas Wechsler de Inteligencia**

WAIS (Wechsler, 1955) fue la primera escala de Wechsler adaptada a la población chilena (Berdicewski & Herreros, 1960), pero esta versión no incluyó licencia para comercialización. En los años 80 se realizó un nuevo estudio con una muestra de 48 sujetos clasificados por género y nivel socioeconómico (NSE) pero, nuevamente, sin licencia. El resultado de este estudio fue un manual de WAIS-R, en el que se agregó información sobre cada prueba, criterios de suspensión y procedimientos de administración (Hermosilla, 1982).

Desde esa investigación y hasta hoy no ha habido nuevas investigaciones en el área. Los profesionales en Chile informan un uso frecuente del material de WAIS (Wechsler, 1955) con el manual de Hermosilla (1982) o uso de versiones extranjeras con adaptaciones libres (Ematris Consultores, 2012).

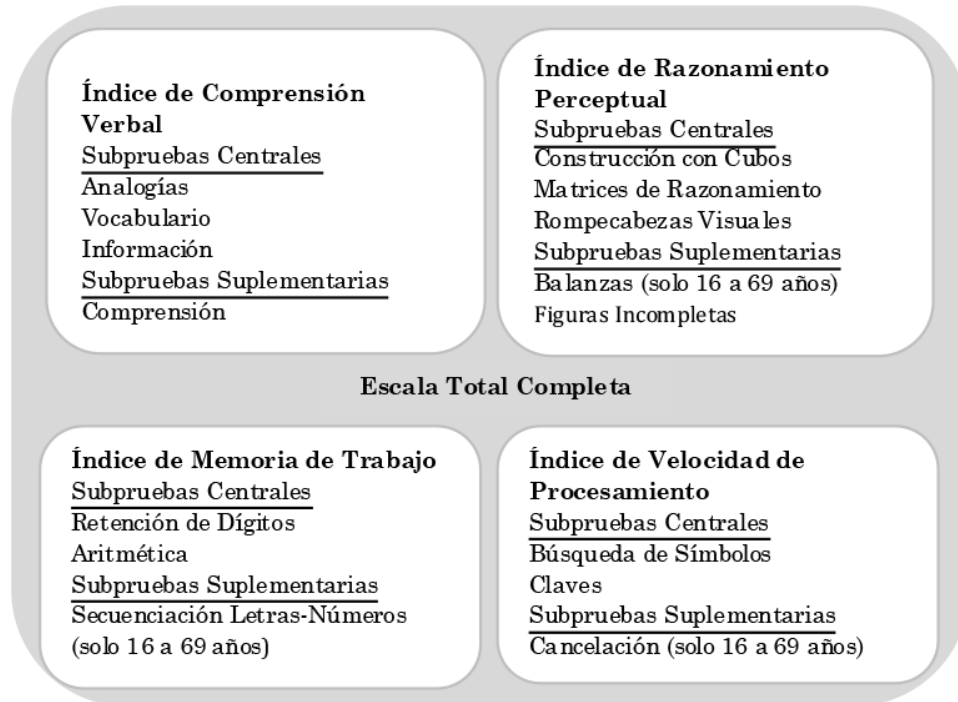
En esta investigación realizamos la adaptación y estandarización de WAIS-IV para la población chilena, con el objetivo de ofrecer un instrumento actualizado, de ítems ajustados a los tiempos modernos y normas actualizadas. Este esfuerzo investigativo contribuye a la compensación del *efecto Flynn*. Este es un fenómeno descrito por Flynn (1984), quien ha demostrado, a partir de datos empíricos disponibles en varios países, que hay un incremento sustancial y sostenido en el tiempo en valores de medición de inteligencia representados en coeficiente intelectual. Aunque el efecto Flynn no ha sido explicado, sí ha sido posible establecer que cada 10 años se deben modificar las pruebas estandarizadas de medición de inteligencia con el fin de evitar sobreestimaciones.

## **La Escala Wechsler de Inteligencia para Adultos-Cuarta Edición**

La Escala Wechsler de Inteligencia Para Adultos-Cuarta Edición es una prueba de administración individual que permite la evaluación comprensiva de la inteligencia desde los 16:00 hasta los 90:11 años. Su resultado es un conjunto de cuatro puntajes compuestos que corresponden a los índices factoriales. Estos son el Índice de Comprensión Verbal (ICV), el Índice de Razonamiento Perceptual (IRP), el Índice de Memoria de Trabajo (IMT) y el Índice de Velocidad de Procesamiento (IVP). Entrega, además, un Índice de Habilidad General (IHG), un CIT y una serie de variables que permiten analizar el proceso general realizado por el sujeto.

## Subpruebas y Principios Generales Para Análisis

WAIS-IV está compuesto por 15 subpruebas, 10 son centrales y cinco son suplementarias. Las centrales son necesarias para la estimación de los puntajes compuestos, CIT e IHG; las suplementarias entregan mayor información sobre el funcionamiento cognitivo del sujeto y permiten acciones de sustitución, en caso de ser necesario (ver Figura 1)



*Figura 1.* Distribución de las subpruebas centrales y suplementarias en WAIS-IV. La estructura factorial es compartida en la estandarización norteamericana y en la chilena.

Con la aplicación de WAIS-IV el profesional obtiene datos cuantitativos y cualitativos que requiere integrar en la interpretación. Excede el objetivo de esta publicación presentar en detalle el proceso de interpretación, pero ofrecemos al examinador interesado nuestra propuesta de organización de las subpruebas en función del Modelo de Procesamiento de Información (Shannon, 1948). Este modelo sugiere que la actividad de un sujeto para resolver un problema se realiza en tres niveles: un análisis del tipo de información que recibe, un conjunto de procesos ejecutados por el sistema cognitivo y, por último, la conducta observada (ver Tabla 1).

Al momento de reportar los hallazgos, hay dos cuidados básicos que el evaluador deberá considerar. Primero, hay un cambio en los descriptores de rendimiento respecto de aquellos sugeridos en el manual de Hermsilla (1982). Es importante estar alineados para facilitar la comunicación profesional, así es que ofrecemos una comparación entre los descriptores de WAIS-R y WAIS-IV (ver Tabla 2). Segundo, el método de análisis estadístico que aplicamos en WAIS-IV nos permite profundizar en aspectos específicos de la cognición del evaluado, es decir, analizar diferencias entre índices y dar cuenta de la armonía del desarrollo, dispersión intra-escalar, con lo cual vemos si hay unidad de constructo. Además, vemos fortalezas y debilidades del sujeto, que nos permitirán establecer mejores vías para la intervención de un sujeto, aun si su rendimiento se encuentra en uno de los extremos de la curva. Toda esta riqueza de información debe estar interpretada e integrada en el reporte de resultados, tomando en consideración que en presencia de indicadores de disarmonía significativa, los niveles de CIT e índices pueden perder validez.

Tabla 1  
*Características de las Subpruebas Centrales y Suplementarias de WAIS-IV*

Tipo	Subprueba	Entrada	Procesamiento cognitivo	Conducta esperada	Factores de influencia
Índice de Comprensión Verbal					
Centrales	Analogías	Estímulos verbales complejos	Conocimiento lexical, abstracción y categorización verbal	Expresión verbal simple	Flexibilidad cognitiva, interés, negativismo (“no se parecen en nada”) e interés en la lectura
	Vocabulario	Estímulos verbales simples	Conocimiento lexical, conceptualización verbal y evocación de información desde memoria semántica	Expresión verbal compleja	Oportunidades culturales, experiencia con otras lenguas, curiosidad intelectual, interés en la lectura y riqueza del ambiente
	Información	Estímulos verbales complejos	Evocación de información desde almacenes de memoria de largo plazo, recuperación de conocimiento adquirido	Expresión verbal simple	Alerta hacia el ambiente, oportunidades culturales, contacto con otras culturas, curiosidad intelectual e interés en la lectura
Suplementaria	Comprensión	Estímulos verbales complejos	Abstracción y comprensión de situaciones sociales, juicio moral	Expresión verbal compleja	Oportunidades culturales, desarrollo de la conciencia y la norma social, flexibilidad cognitiva y negativismo
Índice de Razonamiento Perceptual					
Centrales	Construcción con Cubos	Estímulos visuales y auditivos complejos	Procesamiento perceptual – vía visual, integración de información viso-espacial y generación de estrategias para manipular información no-verbal. Comprensión de instrucciones verbales	Coordinación mano-ojo, integración viso-espacial	Flexibilidad cognitiva, interés, negativismo (“no alcanzan los cubos”), interés en la lectura y experiencias con juegos similares
	Matrices de Razonamiento	Estímulos visuales y auditivos complejos	Resolución de problemas no-verbales, sin influencia de la velocidad, razonamiento fluido. Comprensión de instrucciones verbales	Coordinación mano-ojo	Tipo de estilo cognitivo (dependiente o independiente del campo), ceguera al color, flexibilidad cognitiva, nivel de motivación, negativismo y persistencia conductual

(continúa)

Tipo	Subprueba	Entrada	Procesamiento cognitivo	Conducta esperada	Factores de influencia
Suplementarias	Rompecabezas Visuales	Percepción visual de estímulos complejos y percepción auditiva de estímulos complejos	Resolución de problemas no-verbales, sin influencia de la velocidad, procesamiento visual	Coordinación mano-ojo	Flexibilidad cognitiva, nivel de motivación, negativismo (“esto es imposible”), persistencia conductual, ansiedad, preocupación por presión de tiempo, impulsividad, habilidad para auto-monitoreo
	Balanzas	Estímulos visuales y auditivos complejos	Resolución de problemas no-verbales, sin influencia de la velocidad, razonamiento fluido. Comprensión de instrucciones verbales	Coordinación mano-ojo	Flexibilidad cognitiva, nivel de motivación, negativismo (“ninguna va ahí”), persistencia, habilidad para transformar el set, problemas en funciones ejecutivas, ansiedad o preocupación por la presión de tiempo, impulsividad y habilidad para auto-monitoreo
	Figuras Incompletas	Estímulos visuales y auditivos complejos	Resolución de problemas no-verbales. Comprensión de instrucciones verbales y acceso a léxico	Expresión verbal simple, coordinación mano-ojo	Nivel de alerta frente al ambiente, estilo cognitivo (dependiente o independiente del campo), atención sostenida y capacidad de trabajo bajo presión
Índice de memoria de trabajo					
Centrales	Retención de Dígitos	Estímulos verbales simples	Memoria de Trabajo Verbal, codificación memoria de corto plazo	Respuesta verbal simple	Habilidad para recibir estímulos pasivamente, span atencional, ansiedad, distractibilidad, flexibilidad cognitiva, habilidades de aprendizaje, presencia de TDAH y negativismo
	Aritmética	Estímulos verbales complejos e información matemática	Memoria de trabajo verbal, solución de problemas presentadas por canal auditivo-verbal	Respuesta verbal simple	Span atencional, ansiedad, atención sostenida, tendencia a la distractibilidad, presencia de dificultades en el aprendizaje y capacidad de trabajo bajo presión de tiempo
Tipo	Subprueba	Entrada	Procesamiento cognitivo	Conducta esperada	Factores de influencia
Suplementaria	Secuenciación de Letras y Números	Estímulos verbales simples	Memoria de trabajo verbal y codificación de información el memoria de corto plazo	Respuesta verbal simple	Span atencional, ansiedad, atención sostenida, flexibilidad cognitiva, dislexia u otras dificultades del aprendizaje, negativismos y persistencia

(continúa)

Índice de Velocidad de Procesamiento					
	Búsqueda de Símbolos	Estímulos abstractos presentados vía visual	Velocidad de procesamiento frente a material no-verbal, procesos de focalización y sostenimiento atencional	Respuesta óculo-manual coordinada	Ansiedad, atención sostenida, dificultades del aprendizaje/TDAH, nivel motivacional, tendencia a la atención obsesiva a los detalles, problemas visuales y habilidad para trabajar bajo presión de tiempo
Centrales	Claves	Estímulos visuales	Atención sostenida	Respuesta óculo-manual coordinada	Ansiedad, atención sostenida, dificultades de aprendizaje, atención obsesiva a los detalles, persistencia conductual, problemas viso-perceptuales y habilidad para trabajar bajo presión
Suplementaria	Cancelación	Estímulos visuales abstractos	Atención sostenida, barrido visual, rapidez de procesamiento de información	Respuesta óculo-manual coordinada	Ansiedad, atención sostenida, agudeza visual, impulsividad, habilidades de planeación y capacidad de trabajo bajo presión

Tabla 2  
*Comparación de Descriptores de Rendimientos WAIS-R Versus WAIS-IV*

Rango de puntaje compuesto	Descriptor WAIS-R	Descriptor WAIS-IV
> 130	Muy superior	Muy superior
120-129	Superior	Superior
110-119	Normal superior	Sobre el promedio
90-109	Normal promedio	Promedio
80-89	Normal lento	Bajo el promedio
70-79	Limítrofe	Limítrofe
< 70	Retardo	Muy bajo

### Método

Realizamos el proceso de estandarización en tres etapas: (a) traducción y adaptación de los ítems, (b) fase piloto y (c) fase experimental. Describimos a continuación los instrumentos, participantes, procedimientos y análisis de datos realizados en cada etapa. Todos los procedimientos que aplicamos fueron aprobados por el Comité de Ética de la Escuela de Psicología de la Pontificia Universidad Católica de Chile y por la comisión del fondo que financió el proyecto.

### Traducción y Adaptación de Ítems

El instrumento que traducimos y adaptamos fue *Wechsler Intelligence Scale for Adults, fourth edition* (WAIS-IV; Wechsler, 2008). Incluye estímulos para presentación al evaluado consignados en dos cuadernos de estímulos, un set de nueve cubos, un protocolo de administración, Cuaderno de Estímulos 1 y 2, plantillas de corrección, un Manual de Administración y Corrección y un Manual Técnico Interpretativo (Coalson & Raiford, 2008).



**Traducción de manuales y protocolos.** Tradujimos el Manual de Administración y Corrección y el Protocolo de Registro que utilizamos durante el pilotaje. En esta etapa hubo entrega de avances parciales a los directores de proyecto para incorporar correcciones y sugerencias.

**Traducción y adaptación de subpruebas verbales.** Realizamos adaptaciones a las subpruebas de Analogías, Vocabulario, Información y Comprensión con el fin de lograr adecuación a las características culturales de Chile. Presentamos información específica sobre las adaptaciones realizadas, sin incluir detalles con el fin de evitar usos inapropiados de WAIS-IV.

**Analogías.** Tradujimos todos los ítems y agregamos tres adicionales con el objetivo de contar con una reserva para eliminar aquellos que no cumplieran con los criterios psicométricos necesarios luego del análisis de resultados. Los tres ítems adicionales fueron: celular-computador, Brasilia-Montevideo y Amazonía-Antártida.

**Vocabulario.** Para la selección de las palabras de esta subprueba fue necesario analizar la versión original para identificar el modo en que esta fue construida. Nos guiamos por el *Corpus of Contemporary American English* (Brigham Young University, 2012) para obtener la frecuencia de aparición y registro de uso (oral, medios masivos de comunicación escrita y textos académicos) de cada palabra en inglés. Posteriormente, comparamos las palabras con el *Corpus Lifcach* (Sadowsky & Martínez, 2012) para seleccionar aquellas que tuvieran la misma clase gramatical, registro de uso y frecuencia que su contraparte norteamericana. A la lista de ítems originales agregamos: cómodo, ominoso, patrimonial, subrepticio e ignorante.

**Información.** De los 26 ítems originales de esta subprueba, tradujimos 19 directamente desde el inglés. Los siete restantes requirieron adaptación, pues los originales no resultaban adecuados para el contexto latinoamericano. En este proceso procuramos que los ítems creados aludieran a contenidos similares que los de la versión norteamericana. En este sentido, la pregunta “Who was Martin Luther King Jr.?” la adaptamos a “¿Quién fue J. F. Kennedy”, debido a que J. F. Kennedy es una figura política conocida a nivel latinoamericano. La pregunta “Who wrote Hamlet?”, cuyo contenido es el reconocimiento de un autor de una obra clásica de la lengua inglesa, la adaptamos a “¿Quién escribió Don Quijote de la Mancha?”. La pregunta “On what continent is Brazil?”, cuyo contenido apunta a la ubicación de un país que se encuentra en otro continente, la adaptamos a “¿En qué continente está India?”. La pregunta “Who was the president of the United States during the civil war?” la adaptamos a “¿Quién fue el líder de la revolución cubana?”, en la medida en que ambas hacen alusión a un líder durante un periodo de turbulencia a nivel regional. La pregunta “Who was Sacagawea?” la adaptamos a “¿Quién fue Ícaro?”, debido a que ambas apuntan al conocimiento de una figura popular. La pregunta “Who was Catherine the Great?” la adaptamos a “¿Quién fue Francisco Franco?”, para mantener el contenido acerca de gobernantes de países extranjeros. La pregunta “Who created Sherlock Holmes?” la adaptamos a “¿Quién creó a Condorito?”, para mantener el contenido sobre los autores de personajes literarios populares.

Adicional a las adaptaciones anteriores agregamos dos ítems para pilotaje: “¿Quién fue Rigoberta Menchú?” y “¿Por qué flota el hielo?”

**Comprensión.** En esta subprueba tradujimos todos los ítems, excepto los refranes, para los cuales buscamos símiles en español que tuvieran el mismo sentido que su contraparte norteamericana. De este modo, el refrán “Fall seven times, stand up eight” lo adaptamos a su equivalente “Si te caes 7 veces levántate 8”, el refrán “If you can’t bite, don’t show your teeth”, a “Perro que ladra no muerde”, el refrán “One frost does not make a winter”, a “Una golondrina no hace verano” y el refrán “Shallow brooks are noisy”, a “Mucho ruido y pocas nueces”.

Adicionalmente, agregamos dos ítems, para probarlos durante la fase piloto y experimental con el objetivo de reemplazar el ítem traducido “¿Qué debería hacer usted si encuentra un sobre, con una estampilla nueva y con la dirección puesta?”, pensando que el uso de correo electrónico ha reemplazado al correo tradicional para el uso de mensajería personal. Estos ítems fueron: “¿Por qué algunos países aprueban leyes que discriminan positivamente a las minorías?” y “¿Por qué no se deben hacer copias ilegales de productos inscritos como propiedad intelectual?”.

## Fase Piloto

La fase piloto tuvo varios objetivos: identificar los ítems con mejor comportamiento psicométrico, verificar los puntos de inicio y suspensión sugeridos para la versión norteamericana, estimar las bonificaciones por tiempo y establecer el criterio de dificultad de cada ítem.

**Participantes.** Participaron voluntariamente 27 sujetos (14 hombres y 13 mujeres) entre 16:00 y 90:11 años ( $M = 39,2$ ,  $DS = 20,1$ ), distribuidos en tres niveles educativos (básico, medio y alto), todos residentes en la Región Metropolitana. Los criterios de inclusión fueron cumplir con la edad y nivel educativo requeridos, no informar por autoreporte dificultades neurológicas o psiquiátricas que hubieran requerido tratamiento médico, informar que no había ingesta de medicamentos cuyo efecto podría influir sobre el rendimiento cognitivo y tener un funcionamiento motor sin alteraciones. A todos los sujetos de 69 años o más les administramos pruebas de tamizaje cognitivo con el objetivo de garantizar una línea base apropiada.

Para el contacto de los participantes utilizamos métodos de muestreo intencionado, solicitudes boca a boca y llamados por redes sociales.

**Procedimiento.** El procedimiento estándar lo iniciamos con el establecimiento del contacto con los sujetos, verificación de antecedentes personales y firma de consentimiento informado. Aplicamos todas las subpruebas, excepto Retención de Dígitos y Secuenciación Letras-Números, cuyos criterios de suspensión habían sido probados en versiones anteriores de Escalas Wechsler de Inteligencia (Wechsler, 1997). En la fase piloto aplicamos todos los ítems de las subpruebas, sin considerar criterios de ítem de inicio, secuencia inversa ni suspensión. El tiempo promedio de evaluación por sujeto fue de tres horas.

**Análisis de datos.** Realizamos análisis descriptivos con el fin de determinar si los puntos de corte y suspensión presentes en la estandarización norteamericana eran adecuados para la muestra chilena. Utilizamos información descriptiva para analizar el nivel de dificultad de los ítems, realizando las reorganizaciones pertinentes en función de dicho nivel.

## Fase Experimental

Con la revisión de la evidencia acumulada en la fase piloto, construimos el instrumento que estandarizamos. En esta fase obtuvimos información de una muestra nacional, con la cual analizamos la validez y consistencia interna de la escala.

**Participantes.** Participaron voluntariamente 887 sujetos. La muestra fue intencionada, diseñada por conglomerados (regiones) e incluyó estratificación por edad, NSE y género (ver Tabla 3). Consideramos siete conglomerados: Región Metropolitana, Biobío, Arica y Parinacota, Tarapacá, Antofagasta, Araucanía y Magallanes. Para el NSE usamos como variable de aproximación los años de educación formal, decisión justificada en reportes previos que han demostrado la presencia de correlaciones altas y directas entre estas dos variables (Caro, 2002; Fundación para la Superación de la Pobreza, 2006). Incluimos tres niveles: *bajo*, para personas que completaron ocho años de educación formal; *medio*, para personas que avanzaron más allá del periodo de educación básica y llegaron como máximo a tener educación superior técnica completa o incompleta (entre 9 y 15 años de educación formal aproximadamente) y *alto*, para personas que completaron la formación universitaria o siguieron adelante con formación de post-grado (más de 17 años de educación formal aproximadamente). Para los sujetos entre 16:00 y 24:11 años tomamos el nivel educativo de los padres como proxy del NSE del joven. En caso de haber en el hogar un único padre o apoderado, ingresamos solo este dato; en caso de haber dos, ingresamos un promedio de ambos. En la variable género procuramos incluir igual cantidad de hombres y mujeres por celdilla de edad y nivel educativo, principio que tuvimos que violar en los grupos de edad mayores de 75:00 años, con más mujeres que hombres, situación que refleja la situación demográfica del país.

**Procedimiento.** Tomamos cuidados específicos para garantizar la calidad de los datos, entre ellos: una experta en evaluación cognitiva y neuropsicológica realizó la capacitación de todos los evaluadores del país, coordinamos el trabajo con los líderes de cada región y los coordinadores de campo, tuvimos entregas y revisiones parciales de los protocolos y garantizamos un espacio de comunicación para atender dudas. Todos los protocolos de la muestra nacional fueron revisados por un equipo de expertos.



Solicitamos la firma de consentimiento informado a cada sujeto y realizamos una verificación general del estado neurológico y psiquiátrico de los participantes. Para personas hasta los 59:11 años empleamos un cuestionario para recabar antecedentes personales relevantes a la evaluación e instruimos a los evaluadores para descartar casos en los que habían componentes conductuales que llamaran su atención (e.g., exceso de inquietud motora, divagaciones, sintomatología positiva activa, entre otros). A todos los adultos mayores de 60 años les aplicamos el Examen Mínimo Mental (Folstein, Folstein & McHugh, 1975) y Frontal Assessment Battery (Dubois, Slachevsky, Litvan & Pillon, 2000). Incluimos solo a los sujetos que obtuvieron rendimiento promedio en estas pruebas, según normas nacionales (Quiroga, Albala & Klaasen, 2004).

**Análisis de datos.** En busca de evidencia de confiabilidad realizamos análisis de consistencia interna con el coeficiente alfa de Cronbach. Dada la estructura de Búsqueda de Símbolos, Claves y Cancelación, se requería aplicar el método test-retest; sin embargo, no realizamos este procedimiento durante la estandarización en Chile. Como estas subpruebas se incluyeron sin cambios respecto de la versión norteamericana, hemos supuesto los mismos índices de confiabilidad test-retest reportados para la versión original (Coalson & Raiford, 2008).

Para aportar evidencia de validez realizamos los siguientes procedimientos: (a) expertos nacionales verificaron la correspondencia de los ítems a los originales después de la traducción, (b) analizamos la matriz de correlaciones producto-momento de Pearson para verificar la estructura interna de la prueba y (c) realizamos un análisis factorial confirmatorio con modelamiento de ecuaciones estructurales (Jöreskog & Sörbom, 1993).

Como indicadores de ajuste del modelo factorial tomamos en consideración: (a)  $\chi^2$ ; (b)  $\chi^2/gl$ , para corregir la posible sensibilidad por el tamaño muestral (Byrne, 2001); (c) el Índice corregido de Bondad de Ajuste (AGFI), que evalúa el grado en que las varianzas y covarianzas del modelo reproducen correctamente la matriz de varianzas y covarianzas observada, una vez que se ajusta por los grados de libertad del modelo; (d) el Índice de Bondad de Ajuste Comparativo (CFI), que evalúa en qué medida el modelo ajustado es mejor que un modelo de referencia, el modelo de independencia, en el cual la covarianza entre los indicadores es fijada en cero y (e) el Error Medio Cuadrático de Aproximación (RMSEA), un indicador de cuán bien se ajusta la matriz de varianza-covarianza estimada a la matriz de covarianza poblacional, una vez que se corrige por los grados de libertad usados en el modelo estimado. Para  $\chi^2/gl$  consideramos valores  $< 2$  como buen ajuste y  $< 3$  como ajuste adecuado. Thompson (2000) señala que los valores para AGFI debieran estar sobre 0,90 o 0,95. Hu y Bentler (1999) recomiendan un CFI 0,95 o superior. Para RMSEA consideramos valores  $< 0,05$  para buen ajuste y  $< 0,07$  para adecuado ajuste (Browne & Cudeck, 1993).

Con los datos obtenidos en la muestra nacional realizamos un análisis de sesgo con regresión logística y Mantel-Haenszel, considerando género, edad y NSE para las subpruebas verbales. Para la construcción de normas hemos seguido el procedimiento de normalización inferencial (Wilkins, Rolfhus, Weiss & Zhu, 2005, Abril).

Realizamos los análisis con los paquetes estadísticos SPSS v.17 y SAS. Fijamos el punto de significación en 0,05 para todos los casos y en el análisis del tamaño del efecto consideramos las recomendaciones de Cohen (1992), quien indica que 0,2 es un tamaño de efecto pequeño, 0,5 es medio y 0,8 es amplio. En el caso de los coeficientes de consistencia interna seguimos las recomendaciones de Cicchetti (1994), fijando el punto deseable en 0,7.

Tabla 3

*Muestra Empleada Para la Estandarización de WAIS-IV en Chile, Según Conglomerado, Nivel Educativo y Grupo de Edad*

Conglomerado	Nivel educativo	Grupo de edad												Total	
		16-17	18-19	20-24	25-29	30-34	35-44	45-54	55-64	65-69	70-74	75-79	80-84		85-90
Santiago	Básico	11	11	11	10	18	13	10	12	12	11	7	9	8	143
	Medio	11	10	11	9	12	10	10	15	11	9	9	6	7	130
	Alto	12	11	10	11	10	9	11	11	9	9	4	5	7	119
	Total	34	32	32	30	40	32	31	38	32	29	20	20	22	392
Concepción	Básico	4	4	4	3	2	2	4	4	3	4	3	3	4	44
	Medio	5	4	4	4	4	4	4	4	4	4	3	4	4	52
	Alto	4	4	4	4	4	4	4	4	4	4	4	5	4	53
	Total	13	12	12	11	10	10	12	12	11	12	10	12	12	149
Arica	Básico	1	2	2	3		2	2	1	2	2	1	2	1	21
	Medio	3	2	2	2	4	2	2	4	1	2	2	2		28
	Alto	2	2	2	2	2	3	1	3	2	2	1	2		24
	Total	6	6	6	7	6	7	5	8	5	6	4	6	1	73
Iquique	Básico	2	2	2	2	2	2	2	2	1	1	2	4	1	25
	Medio	3	2	3	3	1	3	3	2		4	1	1	1	27
	Alto		3	2	1	3	2	2	1	1			2		17
	Total	5	7	7	6	6	7	7	5	2	5	3	7	2	69
Antofagasta	Básico	2	2	2	1		1	2	2	2	2	1	1	1	19
	Medio	3	2	3	2	3	3	1	2	2	2	1			24
	Alto	1	2	2	2	2	1	2	2	2	2				18
	Total	6	6	7	5	5	5	5	6	6	6	2	1	1	61
Temuco	Básico	2	2	3	1		1	1	2		1	1	1	1	16
	Medio	2	2	2	3	3	3	3	2	2	2	2	1		27
	Alto	2	2	2	2	2	2	2	2	2	2		1		21
	Total	6	6	7	6	5	6	6	6	4	5	3	3	1	64
Punta Arenas	Básico	2	2	2	2	2	2	3	2	2	2	2	2	1	26
	Medio	2	2	2	2	2	2	2	2	2	2	2	2	2	26
	Alto	2	2	3	2	2	3	2	2	2	2	3	2		27
	Total	6	6	7	6	6	7	7	6	6	6	7	6	3	79
Total	Básico	24	25	26	22	24	23	24	25	22	23	17	22	17	294
	Medio	29	24	27	25	29	27	25	31	22	25	20	16	14	314
	Alto	23	26	25	24	25	24	24	25	22	21	12	17	11	279
	Total	76	75	78	71	78	74	73	81	66	69	49	55	42	887

### Resultados

Los datos descriptivos generales muestran que no hay diferencias en CIT entre regiones,  $F(6, 880) = 1,56$   $p = 0,153$ , con la media más baja en Iquique ( $M = 97,5$ ,  $DS = 12,2$ ) y la más alta en Punta Arenas ( $M = 103,8$ ,  $DS = 11,0$ ).

Aparecen diferencias en CIT por género,  $F(1, 885) = 8,75$ ,  $p = 0,030$ ,  $\eta^2 = 0,010$ , 95% IC [0,99, 4,92], con mejor rendimiento en los hombres ( $M = 101,88$ ,  $DS = 15,39$ ) sobre las mujeres ( $M = 98,9$ ,  $DS = 14,4$ ). Registramos también diferencias entre los tres NSE,  $F(2, 884) = 387,05$ ,  $p < 0,001$ ,  $\eta^2 = 0,46$ . Comparaciones post-hoc de Bonferroni mostraron diferencias significativas ( $p < 0,001$ ) entre todos los NSE: 95% ICs, [8,13, 12,44], [12,8, 17,05], [23,02, 27,40] para bajo, medio y alto, respectivamente. El nivel bajo ( $M = 87,1$ ,  $DS = 11,5$ ) es superado por el nivel medio ( $M = 102,1$ ,  $DS = 11,4$ ) y este, por el nivel alto ( $M = 112,3$ ,  $DS = 9,7$ ).

El análisis de sesgo mostró un sesgo por edad en un único ítem en Comprensión que eliminamos. No encontramos evidencia de sesgo por género, pero sí por NSE en las subpruebas verbales. Detalles sobre esta condición los discutiremos en otra publicación.

### **Evidencia de Consistencia Interna**

La prueba total arroja un coeficiente alfa de Cronbach de 0,941. Los coeficientes de consistencia interna a nivel de subpruebas van de 0,628 (45-55 años) en Cancelación hasta 0,937 (30-34 años) en Información (Tabla 4). En todas las variables de la batería hay adecuados coeficientes de consistencia interna.

### **Evidencia de Validez**

Respecto de la validez de contenido de la prueba, WAIS-IV evalúa múltiples dimensiones de la cognición, como procesamiento auditivo, manipulación de información visual, organización de información, velocidad del procesamiento de información, entre otras. Esto ha sido verificado y discutido por el panel de expertos que generó el instrumento y análisis posteriores (Coalson & Raiford, 2008; Lichtenberger & Kaufman, 2009).

A nivel de validez de constructo de la prueba, obtuvimos resultados que se organizan en dos niveles: estructura de correlaciones de la prueba y resultado del análisis factorial confirmatorio.

La matriz de correlaciones se organiza en función de lo esperado teóricamente (Tabla 5). Las relaciones más fuertes aparecen entre subpruebas que pertenecen a un mismo índice. Son especialmente llamativas a nivel de subpruebas las correlaciones entre Comprensión y Vocabulario ( $r = 0,716$ ), Construcción con Cubos y Rompecabezas Visuales ( $r = 0,589$ ), Aritmética y Retención de Dígitos ( $r = 0,626$ ) y Búsqueda de Símbolos y Claves ( $r = 0,651$ ).

Respecto de las correlaciones entre las subpruebas y los índices factoriales, las más fuertes aparecen entre Vocabulario e ICV ( $r = 0,904$ ), Construcción con Cubos e IRP ( $r = 0,848$ ), Retención de Dígitos e IMT ( $r = 0,908$ ) y Búsqueda de Símbolos e IVP ( $r = 0,908$ ).

Respecto del análisis factorial confirmatorio, investigaciones previas y resultados en fases tempranas de investigación mostraron que la estructura de WAIS-IV responde a un modelo de segundo orden, con cuatro factores de primer orden (Comprensión Verbal, Razonamiento Perceptual, Memoria de Trabajo y Velocidad de Procesamiento) que contribuyen al factor g (Escala Total).

En la estandarización de WAIS-IV para Chile pusimos a prueba el modelo que mejor ajustó en la versión norteamericana. Este modelo explica la organización que fundamenta los procesos de corrección y análisis.

En tanto WAIS-IV tiene tres subpruebas suplementarias que solo deben ser aplicadas a personas hasta los 69:11 años, estimamos diferencialmente modelos factoriales confirmatorios para la población total (sin estas subpruebas), un modelo para la población entre 16:00 y 69:11 años y otro modelo para la población entre 70:00 y 90:11 años.

Tabla 4  
*Coefficientes de Consistencia Interna Para las Subpruebas e Índices de WAIS-IV Estandarizado en Chile*

Subprueba	Grupo de edad													Promedio
	16-17	18-19	20-24	25-29	30-34	35-44	45-54	55-64	65-69	70-74	75-79	80-84	85-90	
CC	0,761	0,791	0,774	0,750	0,788	0,788	0,785	0,767	0,787	0,724	0,785	0,758	0,807	0,775
An	0,723	0,775	0,791	0,795	0,813	0,822	0,866	0,859	0,814	0,801	0,791	0,810	0,815	0,808
RD	0,880	0,877	0,885	0,889	0,891	0,907	0,879	0,843	0,904	0,863	0,872	0,811	0,895	0,879
MR	0,870	0,878	0,876	0,881	0,906	0,912	0,897	0,913	0,925	0,896	0,863	0,849	0,824	0,887
Voc	0,885	0,888	0,893	0,937	0,924	0,902	0,924	0,926	0,927	0,907	0,878	0,910	0,896	0,909
Ari	0,830	0,872	0,868	0,831	0,866	0,838	0,828	0,829	0,822	0,758	0,793	0,766	0,740	0,823
BS	0,779	0,771	0,775	0,736	0,709	0,664	0,652	0,756	0,797	0,780	0,808	0,795	0,799	0,759
RV	0,858	0,883	0,881	0,878	0,910	0,889	0,884	0,852	0,798	0,854	0,864	0,764	0,717	0,856
Inf	0,910	0,905	0,899	0,914	0,937	0,927	0,932	0,935	0,936	0,919	0,932	0,930	0,933	0,925
Cla	0,813	0,812	0,819	0,785	0,813	0,794	0,779	0,850	0,863	0,780	0,808	0,795	0,799	0,810
SLN	0,814	0,813	0,836	0,813	0,854	0,844	0,811	0,870	0,888	-	-	-	-	0,840
Bal	0,907	0,898	0,888	0,904	0,910	0,906	0,892	0,873	0,881	-	-	-	-	0,896
Com	0,810	0,82	0,812	0,881	0,858	0,815	0,853	0,856	0,889	0,868	0,832	0,889	0,877	0,853
Can	0,779	0,771	0,775	0,736	0,690	0,641	0,628	0,744	0,788	-	-	-	-	0,733
FI	0,759	0,781	0,765	0,828	0,875	0,828	0,815	0,856	0,845	0,789	0,855	0,799	0,788	0,817
CCSB	0,721	0,787	0,759	0,683	0,779	0,775	0,771	0,755	0,781	0,728	0,785	0,765	0,808	0,763
RDD	0,770	0,771	0,762	0,698	0,757	0,798	0,752	0,746	0,816	0,771	0,734	0,708	0,767	0,759
RDI	0,780	0,761	0,754	0,742	0,777	0,819	0,736	0,699	0,793	0,651	0,726	0,649	0,715	0,742
RDS	0,730	0,734	0,798	0,779	0,775	0,754	0,764	0,725	0,786	0,782	0,795	0,693	0,779	0,763
ICV	0,933	0,939	0,939	0,958	0,958	0,951	0,960	0,961	0,958	0,946	0,944	0,954	0,952	0,951
IRP	0,863	0,881	0,881	0,867	0,902	0,897	0,882	0,884	0,891	0,837	0,864	0,844	0,857	0,874
IMT	0,909	0,920	0,923	0,913	0,922	0,923	0,920	0,897	0,932	0,897	0,908	0,851	0,913	0,912
IVP	0,838	0,862	0,874	0,846	0,841	0,830	0,814	0,874	0,880	0,840	0,870	0,863	0,860	0,854
CIT	0,915	0,942	0,962	0,944	0,945	0,960	0,942	0,956	0,938	0,925	0,933	0,931	0,921	0,941

CC = Construcción con Cubos, An = Analogías, RD = Retención de Dígitos, MR = Matrices de Razonamiento, Voc = Vocabulario, Ari = Aritmética, BS = Búsqueda de Símbolos, RV = Rompecabezas Visuales, Inf = Información, Cla = Claves, SLN = Secuenciación de Letras y Números, Bal = Balanzas, Com = Comprensión, Can = Cancelación, FI = Figuras Incompletas, CCSB = Construcción con Cubos sin Bonificación de Tiempo, RDD = Retención de Dígitos Orden Directo, RDI = Retención de Dígitos Orden Inverso, RDS = Retención de Dígitos Secuenciados, ICV = Índice de Comprensión Verbal, IRP = Índice de Razonamiento Perceptual, IMT = Índice de Memoria de Trabajo, IVP = Índice de Velocidad de Procesamiento y CIT = Coeficiente Intelectual Total.

Tabla 5  
Matriz de Correlaciones de Subpruebas, Puntajes de Procedimiento e Índices de WAIS-IV

Subprueba	ICV					IRP				IMT			IVP			Variables de proceso				Índices compuestos				
	An	Voc	Inf	Com	CC	MR	RV	Bal	FI	RD	Ari	SLN	BS	Cla	Can	CCSB	RDD	RDI	RDS	ICV	IRP	IMT	IVP	
Voc	0,70	1																						
Inf	0,66	0,71	1																					
Com	0,67	0,72	0,65	1																				
CC	0,45	0,47	0,52	0,45	1																			
MR	0,54	0,55	0,60	0,54	0,55	1																		
RV	0,39	0,43	0,46	0,40	0,59	0,52	1																	
Bal	0,53	0,56	0,59	0,53	0,57	0,58	0,56	1																
FI	0,44	0,46	0,49	0,42	0,51	0,46	0,47	0,48	1															
RD	0,51	0,55	0,57	0,48	0,47	0,52	0,42	0,55	0,42	1														
Ari	0,52	0,52	0,55	0,48	0,53	0,55	0,51	0,63	0,41	0,63	1													
SLN	0,52	0,53	0,56	0,49	0,46	0,55	0,44	0,56	0,43	0,74	0,57	1												
BS	0,39	0,43	0,43	0,40	0,43	0,40	0,47	0,43	0,39	0,42	0,41	0,44	1											
Cla	0,49	0,52	0,52	0,50	0,48	0,46	0,39	0,48	0,44	0,50	0,44	0,47	0,65	1										
Can	0,34	0,30	0,32	0,29	0,31	0,30	0,33	0,34	0,35	0,35	0,36	0,36	0,50	0,43	1									
CCSB	0,45	0,46	0,52	0,45	0,97	0,55	0,56	0,55	0,50	0,46	0,51	0,45	0,42	0,46	0,29	1								
RDD	0,40	0,45	0,44	0,38	0,35	0,40	0,30	0,44	0,31	0,80	0,51	0,58	0,31	0,36	0,27	0,34	1							
RDI	0,43	0,47	0,45	0,39	0,39	0,48	0,40	0,47	0,36	0,85	0,54	0,64	0,36	0,42	0,32	0,39	0,58	1						
RDS	0,45	0,45	0,51	0,41	0,42	0,44	0,38	0,48	0,38	0,81	0,52	0,64	0,37	0,45	0,30	0,41	0,47	0,56	1					
ICV	0,88	0,90	0,89	0,76	0,54	0,63	0,48	0,63	0,52	0,61	0,60	0,60	0,47	0,57	0,36	0,54	0,49	0,51	0,52	1				
IRP	0,55	0,58	0,63	0,55	0,85	0,83	0,83	0,68	0,57	0,56	0,63	0,57	0,48	0,53	0,37	0,83	0,42	0,51	0,49	0,65	1			
IMT	0,57	0,59	0,62	0,53	0,55	0,59	0,52	0,65	0,46	0,91	0,89	0,73	0,59	0,52	0,40	0,54	0,74	0,78	0,75	0,67	0,66	1		
IVP	0,48	0,52	0,52	0,49	0,50	0,48	0,42	0,50	0,46	0,50	0,47	0,50	0,91	0,91	0,51	0,49	0,37	0,43	0,45	0,57	0,55	0,54	1	
CIT	0,76	0,79	0,81	0,71	0,74	0,77	0,68	0,73	0,60	0,75	0,76	0,71	0,66	0,73	0,47	0,72	0,58	0,64	0,65	0,88	0,87	0,84	0,77	

Notas. Todas las correlaciones son significativas a nivel de  $p < 0,001$ .

$n = 887$  para todas las correlaciones, excepto las que consideran las subpruebas de Cancelación, Balanzas y Secuencias de Letras-Número ( $n = 672$ ).

An = Analogías, Voc = Vocabulario, Inf = Información, Com = Comprensión, CC = Construcción con Cubos, MR = Matrices de Razonamiento, RV = Rompecabezas Visuales, Bal = Balanzas, FI = Figuras Incompletas, RD = Retención de Dígitos, Ari = Aritmética, SLN = Secuenciación de Letras y Números, BS = Búsqueda de Símbolos, Cla = Claves, Can = Cancelación, CCSB = Construcción con Cubos sin Bonificación de Tiempo, RDD = Retención de Dígitos Orden Directo, RDI = Retención de Dígitos Orden Inverso, RDS = Retención de Dígitos Secuenciados, ICV = Índice de Comprensión Verbal, IRP = Índice de Razonamiento Perceptual, IMT = Índice de Memoria de Trabajo, IVP = Índice de Velocidad de Procesamiento y CIT = Coeficiente Intelectual Total.

En la Figura 2 presentamos la organización de las 10 subpruebas centrales considerando la muestra total. A pesar de que los resultados del test de  $\chi^2$  y  $\chi^2/gl$  no indican un buen ajuste del modelo,  $\chi^2(31, N = 887) = 106,58, p < 0,001, \chi^2/gl = 3,44$ , en los restantes índices encontramos niveles buenos o aceptables, AGFI = 0,995, CFI = 0,984, RMSEA = 0,053, lo que nos permitió mantenerlo.

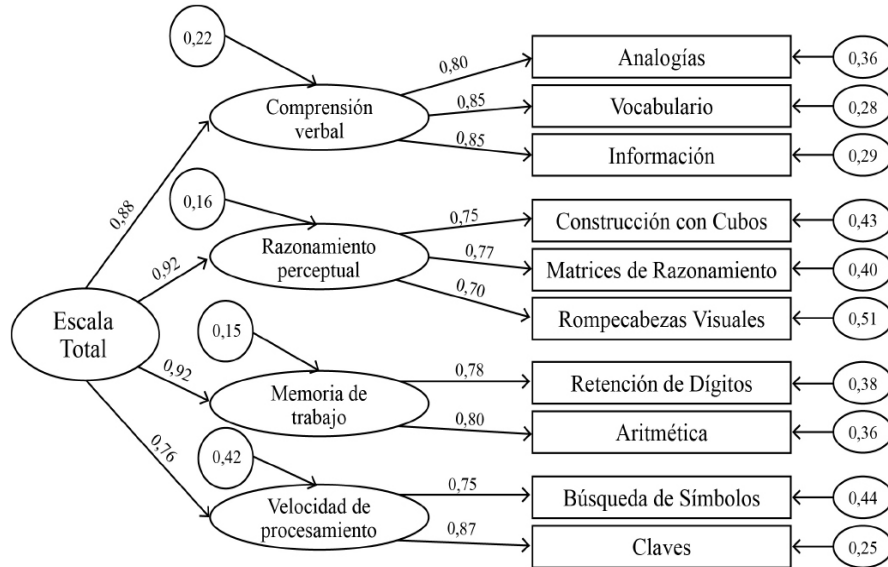


Figura 2. Modelo jerárquico de cuatro factores para las subpruebas centrales en la muestra completa de estandarización para Chile.

En la Figura 3 presentamos el análisis factorial confirmatorio con las 672 personas que están entre los 16:00 y los 69:11 años. Aunque los resultados de  $\chi^2$  y  $\chi^2/gl$  son similares a los observados en la muestra completa,  $\chi^2(86, N = 672) = 296,50, p < 0,001, \chi^2/gl = 3,45$ , también en este caso encontramos niveles buenos o aceptables en los demás índices de ajuste, AGFI = 0,922, CFI = 0,965, RMSEA = 0,060. La distribución justifica la organización factorial que proponemos y confirma la correspondencia de las pruebas suplementarias en los índices indicados.

Para el grupo de 215 adultos mayores (Figura 4), a pesar de que no se rechaza la hipótesis nula en el test  $\chi^2, \chi^2(50, N = 215) = 89,75, p < 0,001$ , el indicador de  $\chi^2/gl$  indica un buen ajuste del modelo,  $\chi^2/gl = 1,795$ . Esto se observa también para los otros indicadores de ajuste: AGFI = 0,922, CFI = 0,965, RMSEA = 0,061.

Respecto de la evidencia de la validez convergente de la prueba, ofrecemos un análisis de su relación con la Prueba de Evaluación de Inteligencia Fluida (FIX; Rosas et al., 2012). En teoría, el FIX es una prueba que permite una evaluación rápida, con buen nivel de confiabilidad y alto nivel de discriminación para el CIT de adultos entre 16:00 y 90:11 años. El FIX fue aplicado en paralelo a WAIS-IV durante la fase de estandarización a una muestra de 375 personas de la Región Metropolitana. Los resultados indican que el FIX en su forma A alcanza una correlación con el CIT de WAIS-IV de 0,79 ( $p < 0,001$ ) y en su forma B, de 0,82 ( $p < 0,001$ ) (Lange et al., 2013).



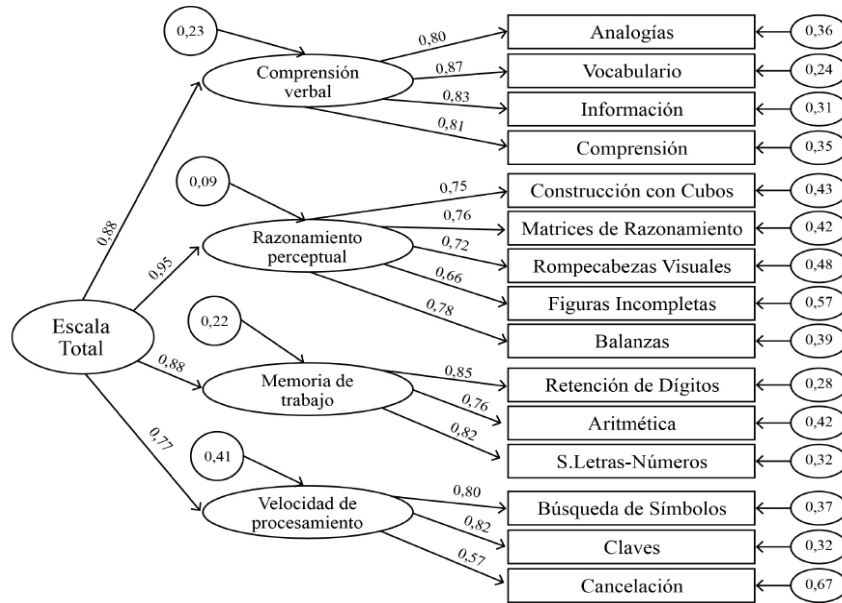


Figura 3. Modelo jerárquico de cuatro factores para las subpruebas centrales y suplementarias en la muestra de 16:00 a 69:11 años de estandarización para Chile.

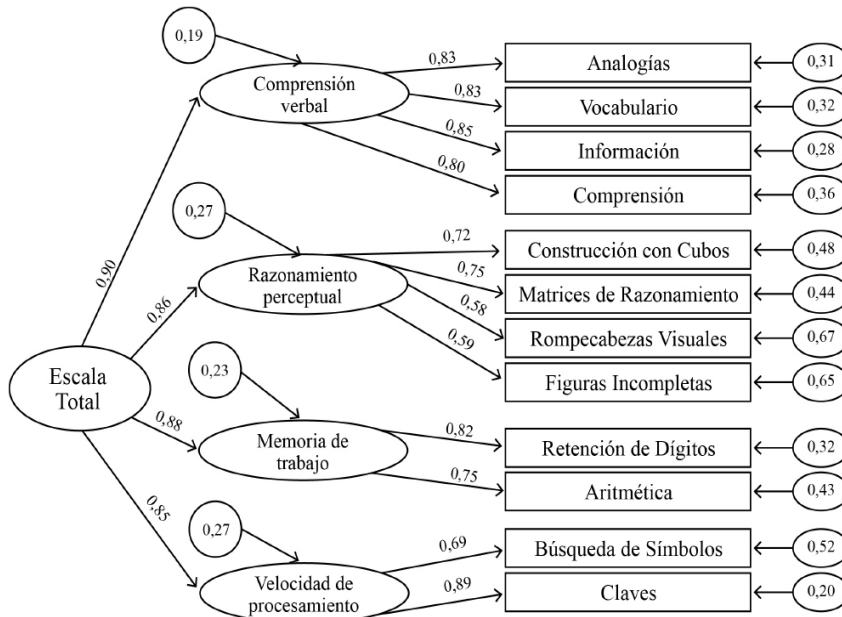


Figura 4. Modelo jerárquico de cuatro factores para las subpruebas centrales y suplementarias en la muestra de 70:00 a 90:11 años de estandarización para Chile.

Analizamos también el comportamiento de indicadores de inteligencia fluida y cristalizada. Utilizamos datos del ICV como indicador cristalizado y de IRP como indicador fluido. Para hacer posible la comparación inter-edades los puntajes brutos de las subpruebas que componen cada índice los transformamos a puntaje  $Z$ , sumándolos y, sobre los valores obtenidos, normalizamos nuevamente con media 0. Considerando este último valor, comparamos las medias por edad. Los resultados corresponden a lo reportado previamente (Baltes, Cornelius, Spiro, Neseelroade & Willis, 1980; Beauducel & Kersting, 2002). Para el componente cristalizado observamos un incremento en los jóvenes que aún están expuestos a la educación formal, produciéndose después un mínimo decremento con la edad (Figura 5a). En contraste, en el componente fluido hay un rendimiento alto y sostenido en los años de educación formal, produciéndose después un declive pronunciado con la edad, con un punto de quiebre a los 25 años (Figura 5b).

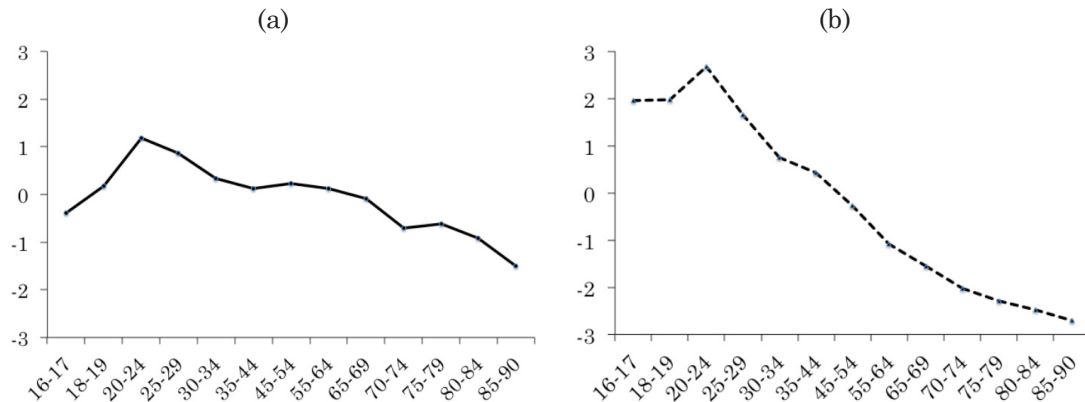


Figura 5. Rendimiento de la muestra de estandarización para Chile diferenciado en (a) componentes cristalizados y (b) componentes fluidos.

Respecto de la evidencia de validez obtenida con grupos especiales, está disponible la investigación realizada por Fuica et al. (2013) en la que se estudió una muestra de jóvenes entre 16 y 19 años de edad, residentes de zonas urbanas y rurales de la Región del Biobío. Los resultados pusieron en evidencia una interacción entre zona de residencia y nivel educativo de los padres. Se observó que cuando los jóvenes pertenecen a familias formadas por padres de bajo nivel educativo, no hay diferencias en el rendimiento en el CIT en función de la zona de residencia, siendo este el grupo de más bajo resultado. En contraste, cuando los padres tienen un alto nivel educativo, hay un rendimiento significativamente mayor en jóvenes que residen en áreas urbanas.

## Discusión

Hemos presentado el proceso y los principales resultados de la estandarización para Chile de la Escala Wechsler de Inteligencia para Adultos, en su cuarta edición. Los datos evidencian apropiada confiabilidad y validez. Es destacable que esta versión de WAIS-IV alcance una consistencia interna en rango muy bueno. La estructura factorial muestra una organización en cuatro factores que se corresponde teóricamente con las nuevas tendencias de evaluación de inteligencia.

La interacción que hemos reportado entre la zona de residencia (urbana versus rural) y el nivel educativo de los padres es una evidencia de la posición de la inteligencia como variable dependiente. Al cambiar la posición de la inteligencia, desplazándola desde el lugar de variable independiente, donde la ubican las teorías tradicionales, y reubicarla en posición de variable dependiente, se hace posible defender que el rendimiento de un sujeto en una evaluación estandarizada de inteligencia está en estrecha relación con la estimulación y las oportunidades educativas. Este resultado pone de manifiesto la importancia del acceso a la educación como medio que permite la compensación de deprivaciones ambientales.

Tal como se ha reportado en otros estudios (Kieffer, 2008), las subpruebas verbales resultan ser las más afectadas en situación de vulnerabilidad, ejerciendo con ello una presión significativa en el rendimiento total en la prueba. Son también las habilidades más sensibles a un ambiente familiar estimulante, logrando jalonar positivamente el desarrollo cuando se crea un ambiente de apoyo para este.

Aunque el estudio de la muestra nacional pone en evidencia la presencia de diferencias en el rendimiento en función del NSE —siempre a favor de los más aventajados—, hemos tomado la decisión de ofrecer baremos nacionales únicos. Estos ayudarán a ubicar al sujeto en función de una situación educativa ideal y a trabajar para lograr mejor acceso y mayor calidad en las oportunidades de enseñanza-aprendizaje. Es necesario llamar la atención de quienes harán uso de este instrumento en calidad de examinadores, pues deberán incorporar la presencia de estas diferencias al momento de realizar las interpretaciones.

En este sentido, hay dos aspectos que permitirán la mejor interpretación de los resultados. En primer lugar, tanto el resultado expresado en el CIT como el rendimiento expresado en índices y subpruebas deben realizarse en función del contexto social del evaluado. Así, personas de niveles sociales más bajos tienden a obtener resultados ubicados hasta una desviación estándar bajo el promedio, lo que no significa que haya problemas cognitivos. En el otro extremo, personas de niveles sociales más altos tienden a obtener mejores resultados, lo que tampoco constituye necesariamente la presencia de habilidades extraordinarias. Lo que sí debe llamar nuestra atención son situaciones poco esperables, como el hallazgo de alto rendimiento en sectores vulnerables y resultados promedio o inferiores en sectores aventajados.

En segundo lugar, a consecuencia del efecto Flynn (1984, 1999), la incorporación de esta nueva versión de WAIS generará cambios significativos respecto de los resultados obtenidos en evaluaciones previas. La tendencia reportada es hacia la sobre-estimación del rendimiento al usar pruebas obsoletas (Flynn, 1998). A consecuencia de ello, es esperable que en una reevaluación con WAIS-IV la persona obtenga un puntaje inferior al registrado previamente. Esta es una característica que no debe ni sorprender ni preocupar al evaluador. En función de la evidencia presentada, puede estar seguro de que el resultado obtenido con WAIS-IV es tanto confiable como válido y su tarea será entonces re-pensar la situación del evaluado.

Consideramos relevante recordar que la evaluación cognitiva es un paso del proceso evaluativo que debemos cumplir cuando nos interesa integrar los resultados en una conclusión con alto impacto social, como la adecuación a un empleo, un diagnóstico clínico o una necesidad educativa especial. En este sentido, WAIS-IV es una prueba que debe ser empleada junto a otros instrumentos de medición, seleccionados a partir de las necesidades de evaluación y las hipótesis a verificar.

Nuestro trabajo tiene varias limitaciones, siendo la más importante la falta de estudios con poblaciones clínicas que aporten a la validez del instrumento. Es igualmente notoria la ausencia de una comparación entre WAIS-IV y la versión chilena de WISC-III (Ramírez & Rosas, 2007) para el grupo entre 16:00 y 16:11 años. Este artículo falla también en ofrecer información detallada sobre el sesgo observado en pruebas verbales en función del NSE, información que presentaremos en otro espacio. A futuro esperamos ofrecer nuevos estudios con poblaciones especiales y ampliar el análisis de relación con otras variables a partir de la comparación de WAIS-IV con pruebas gold estándar de dominios específicos, como sistemas de memoria o funciones ejecutivas.

## Referencias

- American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999). *The standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Baltes, P. B., Cornelius, S. W., Spiro, A., Neseelroade, J. R. & Willis, S. L. (1980). Integration versus differentiation of fluid/crystallized intelligence in old age. *Developmental Psychology*, *16*, 625-635. doi:10.1037/0012-1649.16.6.625
- Beauducel, A. & Kersting, M. (2002). Fluid and crystallized intelligence and the Berlin Model of Intelligence Structure (BIS). *European Journal of Psychological Assessment*, *18*, 97-112. doi:10.1027//1015-5759.18.2.97
- Berdicewski, O. & Herreros, R. (1960). *Normas de adaptación WAIS a Chile* (Tesis de Grado no publicada), Universidad de Chile, Santiago, Chile.
- Brigham Young University (2012). *Corpus of Contemporary American English* [Base de Datos]. Provo, UT: Autor. Extraído de <http://corpus.byu.edu/cocal>
- Browne, W. M. & Cudeck, R. (1993). Alternative ways of assessing model fit. En K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp.132-162). Newbury Park, CA: Sage.
- Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum.

- Caro, D. (2002). *Estimación del nivel socioeconómico de las familias: propuesta metodológica para la evaluación nacional de rendimiento del 2001*. Lima, Perú: Ministerio de Educación. Extraído de [http://www2.minedu.gob.pe/umc/admin/images/menanexos/menanexos\\_44.pdf](http://www2.minedu.gob.pe/umc/admin/images/menanexos/menanexos_44.pdf)
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6, 284-290. doi:10.1037/1040-3590.6.4.284
- Coalson, D. L. & Raiford, S. E. (2008). *WAIS-IV Technical and Interpretative Manual*. San Antonio, TX: Pearson.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159. doi:10.1037/0033-2909.112.1.155
- Dubois, B., Slachevsky, A., Litvan, I. & Pillon, B. (2000). The FAB: A frontal assessment battery at bedside. *Neurology*, 55, 1621-1626. doi:10.1212/WNL.55.11.1621
- Ematris Consultores (2012). *Estudio de mercado* (Concurso I+D, FONDEF Construcción y estandarización de instrumentos de evaluación psicométrica para el desarrollo de una cultura de evaluación ética y rigurosa). Santiago, Chile: Pontificia Universidad Católica de Chile, CEDETi.
- Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95, 29-51. doi:10.1037/0033-2909.95.1.29
- Flynn, J. R. (1998). WAIS-III and WISC-III IQ gains in the United States from 1972 to 1995: How to compensate for obsolete norms. *Perceptual & Motor Skills*, 86, 1231-1239. doi:10.2466/pms.1998.86.3c.1231
- Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist*, 54, 5-20. doi:10.1037/0003-066X.54.1.5
- Folstein, M. F., Folstein, S. E. & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.
- Fuica, P., Lira, J., Alvarado, K., Aranceda, C., Lillo, G., Miranda, R. ... Pérez Salas, C. P. (2013). *Habilidades cognitivas, contexto rural y urbano: comparación de perfiles WAIS-IV en jóvenes de la región del Biobío*. Manuscrito sometido para publicación.
- Fundación para la Superación de la Pobreza (2006). *Umbrales sociales 2006: propuesta para una futura política social*. Santiago, Chile: Autor.
- Hermosilla, M. (1982). *La Escala de Inteligencia de Wechsler para adultos*. Manuscrito no publicado, Escuela de Psicología, Pontificia Universidad Católica de Chile, Santiago, Chile.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55. doi:10.1080/10705519909540118
- Jöreskog, K. G. & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Hillsdale, NJ: Lawrence Erlbaum.
- Kaufman, A. S. (1994). *Nuevas alternativas para la interpretación del WISC-III*. México DF, México: Manual Moderno.
- Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology*, 100, 851-868. doi:10.1037/0022-0663.100.4.851
- Lange, M., Riveros, R., Figueroa, P., Benavente, C., Aparicio, A. & Rosas, R. (2013). *FIX: propuesta de nueva herramienta de evaluación de inteligencia fluida en adultos*. Manuscrito sometido para publicación.
- Lichtenberger, E. O. & Kaufman, A. S. (2009). *Essentials of WAIS-IV assessment*. Hoboken, NJ: John Wiley & Sons.
- Quiroga, P., Albala, C. & Klaasen, G. (2004). Validación de un test de tamizaje para el diagnóstico de demencia asociada a edad, en Chile. *Revista Médica de Chile*, 132, 467-478. doi:10.4067/S0034-98872004000400009
- Ramírez, V. & Rosas, R. (2007). Estandarización del WISC-III en Chile: descripción del test, estructura factorial y consistencia interna de las escalas. *Psykhe*, 16(1), 91-109. doi:10.4067/S0718-22282007000100008
- Rosas, R., Riveros, R., Aparicio, A., Benavente, C., Figueroa, P. & Lange, M. (2012). *Examen de inteligencia fluida*. Santiago, Chile: Pontificia Universidad Católica de Chile, Centro de Desarrollo de Tecnologías de Inclusión.
- Sadowsky, S. & Martínez, R. (2012). *Corpus Lifcach* [Base de Datos]. Temuco, Chile: Universidad de la Frontera. Extraído de <http://sadowsky.cl/lifcach.html>
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379-423, 623-656.
- Thompson, B. (2000). Ten commandments of structural equation modeling. En L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding MORE multivariate statistics* (pp. 261- 283). Washington, DC: American Psychological Association.
- Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore, MD: William and Wilkins.
- Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale (WAIS)*. New York, NY: Psychological Corporation.
- Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised (WAIS-R)*. New York, NY: Psychological Corporation.
- Wechsler, D. (1997). *Manual for the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III)*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2008). *WAIS-IV Administration and Scoring Manual*. San Antonio, TX: Psychological Corporation.
- Wilkins, C., Rolffus, E., Weiss, L. & Zhu, J. J. (2005, Abril). *A new method for calibrating translating test with small simple sizes*. Ponencia presentada en la 2005 Annual Meeting of the American Educational Research Association, Montreal, Canadá.

Fecha de recepción: Octubre de 2012.

Fecha de aceptación: Mayo de 2013.