

BAT-7, BATERÍA DE APTITUDES DE TEA: DESCRIPCIÓN Y DATOS PSICOMÉTRICOS
BAT-7, TEA ABILITIES BATTERY: DESCRIPTION AND PSYCHOMETRIC PROPERTIES

Fernando Sánchez-Sánchez. David Arribas Águila
Dpto. I+D+i de TEA Ediciones

<http://dx.doi.org/10.17060/ijodaep.2014.n1.v2.450>

Fecha de recepción: 15 de Diciembre de 2013

Fecha de admisión: 30 de Marzo de 2014

ABSTRACT

BAT-7 is a new psychometric tool for assessing three intellectual domains, general factor (g), fluid intelligence (Gf) and crystallized intelligence (Gc), and eight cognitive abilities: Verbal (V), Spatial (E), Attention (A) and Concentration (CON), Reasoning (R), Numerical (N), Mechanical (M) and Orthography (O). It consists of three forms of increasing difficulty (E, M and S) and it is focused on the assessment of school and college students, as well as adults with different educational background.

The present study aims to provide a description of the BAT-7 development and the main features of it, and a summary of its psychometric properties, mainly its reliability and validity.

BAT-7 was administered to a Spanish standardization sample of 4,263 students (age between 12 and 18) and 1,507 adults. Reliability was studied with ordinal alpha, Cronbach's alpha and tests information functions. Validity was explored by multigroup CFA.

3P IRT model showed a statistical fit for all items. Ordinal alpha ranged from .79 to .91 for test scores and from .91 to .97 for composite scores. The model based on CHC theory showed a good fit to the observed data (RMSEA=.034; SRMR=.026; CFI=.981).

The results, in terms of TRI model fitting, reliability and validity, suggest that the BAT-7 is a useful, flexible and highly reliable test for assessing the cognitive abilities, and provide evidences supporting the use of the battery for the assessment of intellectual abilities in children and adults. An estimate of general ability (g), fluid intelligence (Gf) and crystallized intelligence (Gc) also can be obtained by applying the BAT-7.

Keywords: Intelligence, cognitive abilities, assessment, *g* factor.

BAT-7, BATERÍA DE APTITUDES DE TEA: DESCRIPCIÓN Y DATOS PSICOMÉTRICOS

RESUMEN

El BAT-7 es una nueva batería que permite estimar la inteligencia (factor general, *g*; inteligencia fluida, *Gf*; e inteligencia cristalizada, *Gc*) y evaluar 8 aptitudes cognitivas: razonamiento verbal (*V*), de razonamiento espacial (*E*), atención (*A*) y Concentración (*Con*), razonamiento abstracto (*R*), razonamiento numérico (*N*), razonamiento mecánico (*M*) y ortografía (*O*). Se compone de tres formas o niveles de dificultad creciente (Elemental, Medio y Superior) y está diseñada para la evaluación intelectual colectiva e individual de los estudiantes de E. Secundaria y universitarios, así como adultos con diferentes niveles educativos.

El presente trabajo tiene como objetivo ofrecer una descripción del desarrollo de la prueba y de las principales características de la misma así como un resumen de sus propiedades psicométricas, fundamentalmente de su fiabilidad y validez.

La muestra de tipificación española estuvo formada por 4.263 alumnos (de 1.º de ESO hasta 2.º de Bachillerato) y 1.507 adultos de diferente nivel educativo y edad. La fiabilidad fue estudiada por medio de las funciones de información y de los coeficientes alfa de Cronbach y alfa ordinal. La validez se exploró mediante análisis factoriales confirmatorios multigrupo (ACF).

Los ítems que componen cada prueba muestran un buen ajuste al modelo TRI de tres parámetros utilizado. Los valores del alfa ordinal variaron de 0,79 hasta 0,91 para las puntuaciones de la prueba y 0,91-0,97 para las puntuaciones compuestas o índices. El modelo factorial basado en la teoría CHC mostró un buen ajuste a los datos observados (RMSEA = 0,034 ; SRMR = 0,026 ; CFI = 0,981).

Los resultados obtenidos en términos de ajuste al modelo de TRI, fiabilidad y validez indican que el BAT-7 es una medida útil, flexible y altamente fiable para la evaluación de las aptitudes cognitivas y aportan evidencias que apoyan el uso de la batería para la evaluación de las aptitudes intelectuales en escolares y adultos. Igualmente, mediante su aplicación puede obtenerse una estimación de la capacidad general (*g*), de la inteligencia fluida (*Gf*) y de la inteligencia cristalizada (*Gc*).

Palabras clave: Inteligencia, aptitudes intelectuales, evaluación, factor *g*.

INTRODUCCIÓN

El estudio de las facultades de orden superior del ser humano es sin duda uno de los fenómenos que ha gozado de mayor atención y dedicación en Psicología, existiendo multitud de publicaciones que tratan de caracterizar, definir y estructurar esas facultades.

En la actualidad, existe una contundente y abundante evidencia empírica, acumulada durante más de un siglo de investigación, sobre cómo se relaciona la inteligencia con infinidad de criterios externos de tipo social, demográfico, biológico y psicológico (p. ej., Burks *et al.*, 2009; Colom, 2002; Deary, 2012; Deary *et al.*, 2007; Gottfredson, 1997, 2000; Jensen, 1998; Lubinski, 2000; Ree y Caretta, 2002; Schmidt, 2002; Strenze, 2007). Entre los criterios que se tienen en cuenta en las investigaciones, la relación con los tipos de rendimiento académico y profesional son dos de los aspectos más estudiados y sobre los que se ha acumulado una mayor cantidad de evidencia empírica (Kuncel y Hezlett, 2007; Kuncel, Hezlett y Ones, 2001; Sackett, Boerneman y Connelly, 2008). Según la APA (American Psychological Association; Neisser *et al.*, 1996), los tests estandarizados de inteligencia correlacionan 0,50 con el rendimiento escolar, 0,55 con el número de años de escolarización y 0,54 con el rendimiento laboral. Independientemente de la fuente que se analice, parece claro que las puntuaciones en los tests cognitivos son una fuente de información relevante a tener en cuenta a la hora de tomar decisiones sobre una persona.

En el campo educativo, las correlaciones entre las medidas de inteligencia y el rendimiento escolar dependen de la etapa de escolarización, estando entre 0,50 y 0,70 o más en la educación básica y entre 0,30 y 0,50 o más en la educación secundaria y bachillerato (p. ej., Brody, 1997; Jensen, 1981; Haywood, 2004; Santamaría *et al.*, 2005; Sattler, 2001; Süß, 2001). Estos resultados indican que la ejecución en los tests cognitivos es un factor relevante que convendría tener en cuenta para la predicción del rendimiento de un alumno (Detterman y Thompson, 1997). Sin duda existen otros factores relevantes, como la motivación (Duckworth *et al.*, 2011), los intereses (Burks *et al.*, 2009) o la

disciplina (Duckworth y Seligman, 2005), pero los tests de inteligencia permiten predecir, sin duda también, el rendimiento educativo de un alumno y el aprovechamiento que realiza de sus estudios. Parece claro que la evaluación de las aptitudes intelectuales en los alumnos de distintas etapas educativas ayuda al profesional a predecir qué dificultades puede presentar cada uno de ellos y a diseñar las intervenciones psicológicas y educativas más adecuadas a cada caso. Además, en el ámbito de la orientación académica y profesional de los alumnos, la evaluación de las aptitudes es de gran utilidad para realizar orientaciones basadas tanto en los intereses y motivaciones del alumno como en su perfil de capacidades, aptitudes y competencias.

Las medidas aptitudinales de aplicación individual permiten realizar una evaluación exhaustiva y profunda del funcionamiento cognitivo, pero resultan poco prácticas cuando se desea recabar información de grupos numerosos, debido a la cantidad de tiempo que requieren. Por este motivo se decidió desarrollar el BAT-7 (Arribas *et al.*, 2013), una batería de nueva generación que permitiera la evaluación colectiva de diferentes aptitudes intelectuales y a su vez, obtener una estimación de la inteligencia basada en una de las teorías con más apoyo empírico, la teoría de Cattell-Horn-Carroll (CHC; Carroll, 1993). En el presente trabajo se describen las principales características del BAT-7 y se describen brevemente los resultados de los estudios de fiabilidad y validez de la batería.

MÉTODO

Participantes

Tras un extenso estudio piloto en el que participaron 895 alumnos de E. Secundaria y Bachillerato, se realizó el estudio de tipificación de la batería para obtener el conjunto de datos necesario para construir unos baremos amplios y representativos de la población escolar española. La muestra de tipificación fue recogida por más de 70 los profesionales, aportando casos procedentes de 31 provincias españolas.

El número total de alumnos evaluados fue de 4.263, distribuidos por nivel, curso y sexo de acuerdo a los datos que se presentan en la tabla 1. En cuanto a la *región geográfica* de procedencia, el 31,5% de los casos pertenecían a la zona Centro, el 17,0% a la zona Este, el 35,1% a la zona Norte y el 16,4% de los casos a la zona Sur¹. En cuanto al *tipo de centro*, el 17,6% eran alumnos que estudiaban en centros públicos, el 77,6% en centros concertados y el 4,8% en centros privados. Además, se recogieron datos sobre el país de procedencia y la lengua principal en el hogar; los datos indicaron que el 90,1% de los evaluados procedían de España, frente al 9,9% procedente de otro país, y que la lengua principal del 93,8% era el castellano, frente al 3,9% cuya lengua principal era el catalán, el gallego o el euskera y el 2,3% correspondiente a cualquier otro idioma.

Tabla 1. Distribución de la muestra de tipificación por curso y sexo

| | Varones | % | Mujeres | % | TOTAL ² |
|------------------------|--------------|-------------|--------------|-------------|--------------------|
| NIVEL Elemental | | | | | |
| 1.º ESO | 427 | 51,1 | 408 | 48,9 | 841 |
| 2.º ESO | 322 | 45,8 | 381 | 54,2 | 704 |
| NIVEL Medio | | | | | |
| 2.º ESO | 42 | 46,7 | 48 | 53,3 | 91 |
| 3.º ESO | 374 | 48,4 | 398 | 51,6 | 777 |
| 4.º ESO | 283 | 49,6 | 288 | 50,4 | 575 |
| CFGM | 43 | 32,8 | 86 | 67,2 | 131 |
| NIVEL Superior | | | | | |
| 4.º ESO | 74 | 50,3 | 73 | 49,7 | 151 |
| 1.º BACHILLERATO | 251 | 50,4 | 247 | 49,6 | 505 |
| 2.º BACHILLERATO | 68 | 41,5 | 79 | 58,5 | 135 |
| CFGS | 65 | 59,1 | 45 | 40,9 | 110 |
| UNIVERSITARIOS | 57 | 23,6 | 185 | 76,4 | 243 |
| TOTAL | 1.994 | 47,1 | 2.240 | 52,9 | 4.263 |

² Los valores de esta columna no se corresponden con la suma de valores parciales, aleatoriamente debido a errores de redondeo.

BAT-7, BATERÍA DE APTITUDES DE TEA: DESCRIPCIÓN Y DATOS PSICOMÉTRICOS

Instrumento de evaluación: Descripción del BAT-7, Batería de Aptitudes de TEA

El BAT-7 es una batería que evalúa las aptitudes cognitivas de los escolares y de los adultos. Está formada por 3 cuadernillos o niveles de dificultad creciente (E, M y S; Elemental, Medio y Superior), cada uno de los cuales incluye 7 tests: **V** (aptitud verbal), **E** (aptitud espacial), **A** (atención), **R** (razonamiento abstracto), **N** (aptitud numérica), **M** (aptitud mecánica) y **O** (ortografía). La mayor cantidad de información y precisión de los resultados se obtiene aplicando todos ellos, aunque también se pueden utilizar de forma aislada o combinada en función de las necesidades del profesional. Los tipos de estímulos y tareas que incorpora cada uno de los test se describe brevemente en la tabla 2.

Tabla 2. Descripción de los test de la batería

| |
|--|
| Test V (aptitud verbal; 12 minutos): Analogías verbales compuestas por cuatro términos a completar eligiendo una entre cuatro opciones de respuesta posibles. Para ello, se debe descubrir cómo se relacionan dos de los términos verbales de la analogía para aplicar después la misma lógica a uno o más de los términos restantes, obteniendo como resultado una frase verdadera y con pleno sentido. |
| Test E (aptitud espacial; 15 minutos): Desarrollo de superficies o ensamblaje de figuras en tres dimensiones. El evaluado debe visualizar y rotar mentalmente de una a cuatro figuras tridimensionales a partir de sus correspondientes desarrollos bidimensionales y comparar los resultados con un modelo dado en términos de igual-diferente. |
| Test A (atención-concentración; 8 minutos): Tarea de velocidad de cancelación o búsqueda visual. En cada ítem se debe identificar el número de veces que aparece un patrón estímulo entre otros de características muy similares. Test R (razonamiento abstracto; 20 minutos): Series de figuras o dibujos ordenados de acuerdo a una ley o principio. El examinando debe descubrir la lógica subyacente y aplicarla al último elemento de la serie para seleccionar, entre cuatro opciones de respuesta, la que continúa la secuencia. |
| Test N (aptitud numérica; 20 minutos): Incluye 3 tipos de tareas (igualdades numéricas, series y tablas incompletas) con mayor carga de razonamiento numérico que de cálculo. Se trata de ítems de elección múltiple en los que el examinando debe seleccionar entre 5 opciones de respuesta el valor numérico que cumple una igualdad, que continúa una serie o que completa una tabla de valores. |
| Test M (aptitud mecánica; 10 minutos): Se representan diferentes situaciones prácticas para que el evaluado prevea el efecto de aplicar principios físicos ligados al movimiento de los cuerpos en el espacio. |
| Test O (ortografía; 10 minutos): Tarea o caso de identificación de una palabra incorrectamente escrita entre cuatro opciones posibles. |

La batería permite obtener **8 puntuaciones de aptitud** (una por cada test y una evaluación de la concentración derivada del test A), **una puntuación total** (Total BAT), el **estilo atencional** del examinando y **3 índices de inteligencia**: *g* (capacidad general), *Gf* (inteligencia fluida) y *Gc* (inteligencia cristalizada). Las puntuaciones de aptitud del BAT-7 se obtienen transformando las puntuaciones directas a percentiles mediante un sistema informático de corrección o consultando las tablas de baremos. Existen diferentes tablas de baremos por curso o edad con las que comparar el rendimiento de los escolares. La tabla 3 presenta una breve definición operativa de lo que representa cada puntuación de aptitud.

Tabla 3. Definición operativa de las puntuaciones de aptitud

| Puntuación | Aptitud | Definición operativa |
|------------|----------------|---|
| V | Verbal | Destreza para formular y comprobar hipótesis acerca de conceptos e ideas expresados verbalmente. Implica cierto grado de conocimiento léxico y la comprensión semántica de nombres, verbos y adjetivos. |
| E | Espacial | Capacidad para visualizar, recordar y transformar mentalmente imágenes visuales en dos y tres dimensiones. |
| A | Atención | Habilidad para identificar rápida y selectivamente los aspectos relevantes de un estímulo y para ignorar los irrelevantes. Puede interpretarse como una medida de la velocidad de procesamiento. |
| CON | Concentración | Precisión en el procesamiento de la información visual independiente de la velocidad. Puede interpretarse como una medida de la calidad del procesamiento. |
| R | Raz. abstracto | Capacidad para resolver problemas novedosos aplicando leyes lógicas de tipo deductivo y estableciendo correlatos entre figuras abstractas. |
| N | Númerica | Capacidad para razonar de modo inductivo o deductivo con conceptos matemáticos en términos de relaciones y propiedades. |
| M | Mecánica | Grado de comprensión de los principios mecánicos relacionados con el equilibrio y el movimiento de los cuerpos sometidos a cualquier fuerza. |
| O | Ortografía | Habilidad en la aplicación del conocimiento almacenado de las reglas gramaticales. |

Los tres índices para la estimación de la inteligencia que incorpora el BAT-7 son *g* (capacidad general), *Gf* (inteligencia fluida) y *Gc* (inteligencia cristalizada). Estos índices se basan en la aplicación de la técnica del análisis factorial confirmatorio a las puntuaciones TRI, por lo que su cálculo requiere el uso de procedimientos informáticos dada su laboriosidad.

El índice *g* de capacidad general se basa en el rendimiento del examinando en los 7 tests, por lo que es necesario aplicar la batería al completo para obtener su valor. No obstante, pueden obtenerse dos aproximaciones menos precisas de la capacidad general a partir de la aplicación de algunos tests concretos.

El índice *Gf* de inteligencia fluida se obtiene aplicando los tests R y N, por lo que siempre que se apliquen ambos se podrá calcular su valor. Del mismo modo, el índice *Gc* de inteligencia cristalizada se deriva de los tests V y O, por lo que igualmente será necesario aplicar ambos tests para obtener su valor. La tabla 4 presenta una visión general de los rasgos que evalúa cada uno de los índices.

Tabla 4. Definición operativa de los índices intelectuales

| Puntuación | Índice | Definición operativa |
|------------|---------------------------|--|
| <i>g</i> | Capacidad general | Capacidad mental que permite comprender y relacionar ideas complejas, trabajar eficientemente con contenidos abstractos y verbales, extraer conclusiones lógicas, resolver problemas novedosos y aprender con rapidez. |
| <i>Gf</i> | Inteligencia fluida | Capacidad para razonar inductivo y deductivamente con los problemas abstractos y formales que requieran el uso del lenguaje interno para generar y poner a prueba hipótesis e identificar la solución correcta. |
| <i>Gc</i> | Inteligencia cristalizada | Capacidad para aplicar los conocimientos y habilidades adquiridos culturalmente a la resolución de problemas relacionados con materiales y procesos verbales. |

BAT-7, BATERÍA DE APTITUDES DE TEA: DESCRIPCIÓN Y DATOS PSICOMÉTRICOS

Diseño y procedimiento del estudio

Durante el desarrollo del BAT-7 se estableció como objetivo general la construcción de una batería para la evaluación de las aptitudes cognitivas en los ámbitos escolar y profesional, cuya resolución demandara del examinando la puesta en marcha de algunos de los factores de la teoría CHC que se consideran importantes para el desarrollo académico, la adaptación al entorno y el desempeño profesional. Bajo esta premisa, se delimitaron varios objetivos concretos, como fueron:

Representar adecuadamente varias aptitudes intelectuales de la teoría CHC relacionadas con el rendimiento y el potencial académico y profesional.

Ajustar la dificultad de las tareas a diferentes niveles aptitudinales para aumentar la capacidad de discriminación de la batería.

Crear un instrumento psicométrico con la mayor fiabilidad y validez posibles de acuerdo a los enfoques actuales en teoría de tests.

Obtener una prueba de fácil aplicación y corrección, con el máximo ahorro de tiempo.

Facilitar una gran cantidad de información que permita conocer los puntos fuertes y las debilidades aptitudinales de varios tipos de poblaciones.

Para alcanzar estos objetivos el proceso de desarrollo se estructuró en siete grandes fases:

Diseño de la batería. Durante esta fase se siguieron las recomendaciones de Navas (2001) y Hambleton y Swaminathan (1985) respecto a los objetivos y especificaciones de la evaluación.

Construcción de los ítems. Se construyeron un total de 700 ítems que fueron revisados por un panel formado por 5 expertos de acuerdo a unos criterios de calidad preestablecidos. Tras la revisión, se seleccionaron para el estudio piloto un total de 642 ítems (V: 100 ítems; E: 100 ítems; A: 80 ítems; R: 85 ítems; N: 115 ítems; M: 70 ítems; O: 92 ítems) revisados y equilibrados en cuanto al nivel de dificultad apriorístico.

Estudio piloto. La muestra con la que se llevó a cabo el estudio piloto para la calibración inicial de los ítems estuvo formada por 895 escolares que cursaban estudios de 1.º de ESO a 2.º de Bachillerato en 3 institutos, uno de ellos de carácter público (46,9% de los casos) y los otros de tipo concertado (53,1%). La edad media de los evaluados fue de 14,31 años ($Dt=2,05$); el 51,3% fueron varones y el 48,7% mujeres.

Análisis y calibración de los ítems. Para el análisis de los ítems de las formas piloto se calcularon los estadísticos de dificultad y discriminación basados en la TCT. Antes de la calibración de los ítems de acuerdo a la TRI se examinaron los supuestos de independencia local y unidimensionalidad, obteniéndose resultados satisfactorios en ambos casos. A continuación se procedió a la estimación de los parámetros TRI de los ítems, utilizando un modelo logístico de 3 parámetros (Birnbaum, 1968; Lord y Novick, 1968) tras examinar la bondad de ajuste de cada ítem al modelo. Tras estos análisis un total de 9 ítems fueron eliminados por su mal funcionamiento o ajuste.

Ensamblaje de los niveles para la tipificación. Finalmente, tras los análisis anteriores, se seleccionaron los ítems que se debían incluir en cada uno de los 3 niveles del BAT-7 y se confeccionaron los cuadernillos que se utilizarían en la fase de tipificación.

Estudio de tipificación. La versión definitiva de la batería fue aplicada a una muestra representativa de escolares compuesta por 4263 escolares.

Análisis de datos y preparación de la versión final del BAT-7. Se analizaron los datos procedentes del estudio de tipificación para estimar la fiabilidad y la validez de la prueba y posteriormente se confeccionaron los baremos de la batería.

RESULTADOS

Fiabilidad: consistencia interna

La fiabilidad de las puntuaciones del BAT-7 fue estudiada por medio de las funciones de información y de los coeficientes alfa de Cronbach y alfa ordinal.

Los valores del coeficientes alfa de Cronbach (Cronbach, 1951) de los tres niveles para los principales grupos normativos y puntuaciones se muestran en la tabla 5. En el caso de la puntuación Total BAT-7, de los indi-

ces de inteligencia (*g*, *Gf*, *Gc*) y de las estimaciones de *g* (*g_{VERNO}* y *g_{VR}*), se utilizó la fórmula de la fiabilidad de un test compuesto (Martínez Arias, 1995).

Tabla 5. Alfa de Cronbach de las puntuaciones y los índices

| | V | E | A | CON | R | N | M | O | α-INTL BAT-7 | B | Gf | Gc | g _{VERNO} | g _{VR} |
|---------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----------------|-----|-----|-----|--------------------|-----------------|
| NIVEL E | | | | | | | | | | | | | | |
| SCALA INTELIGENCIA | .81 | .84 | .81 | .81 | .80 | .80 | .77 | .83 | .85 | .84 | .81 | .80 | .84 | .83 |
| 2x ESD ¹ | .85 | .81 | .81 | .81 | .80 | .80 | .78 | .87 | .85 | .84 | .80 | .80 | .84 | .84 |
| 1x ESD ² | .84 | .81 | .80 | .80 | .80 | .80 | .77 | .86 | .85 | .84 | .80 | .80 | .84 | .84 |
| ADULTOS (NIVEL E) | .83 | .82 | .80 | .80 | .80 | .80 | .78 | .87 | .85 | .84 | .80 | .80 | .84 | .84 |
| TOTAL NIVEL E: | .81 | .84 | .83 | .81 | .80 | .80 | .78 | .86 | .85 | .84 | .80 | .80 | .84 | .84 |
| NIVEL M | | | | | | | | | | | | | | |
| ESCOLARES (NIVEL M) | .77 | .82 | .82 | .81 | .80 | .80 | .71 | .85 | .85 | .83 | .83 | .87 | .81 | .80 |
| 2x ESD ¹ | .80 | .81 | .82 | .82 | .80 | .80 | .70 | .87 | .87 | .81 | .80 | .78 | .81 | .79 |
| 1x ESD ² | .80 | .81 | .82 | .82 | .80 | .80 | .70 | .87 | .87 | .81 | .80 | .78 | .81 | .79 |
| ADULTOS (NIVEL M) | .80 | .81 | .82 | .82 | .80 | .80 | .70 | .87 | .87 | .81 | .80 | .78 | .81 | .79 |
| TOTAL NIVEL M: | .79 | .82 | .82 | .81 | .80 | .80 | .75 | .85 | .85 | .83 | .83 | .87 | .81 | .80 |
| NIVEL S | | | | | | | | | | | | | | |
| SCALA INTELIGENCIA | .80 | .84 | .81 | .81 | .81 | .81 | .80 | .80 | .85 | .81 | .81 | .81 | .84 | .84 |
| 2x ESD ¹ | .81 | .81 | .80 | .80 | .80 | .80 | .78 | .87 | .86 | .82 | .80 | .80 | .83 | .84 |
| 1x ESD ² | .81 | .81 | .80 | .80 | .80 | .80 | .78 | .87 | .86 | .82 | .80 | .80 | .83 | .84 |
| ADULTOS (NIVEL S) | .81 | .81 | .80 | .80 | .80 | .80 | .78 | .87 | .86 | .82 | .80 | .80 | .83 | .84 |
| TOTAL NIVEL S: | .80 | .84 | .81 | .81 | .81 | .81 | .80 | .80 | .85 | .81 | .81 | .81 | .84 | .84 |
| OTROS | | | | | | | | | | | | | | |
| UNIVERSITARIOS | .85 | .87 | .81 | .82 | .81 | .80 | .78 | .87 | .86 | .84 | .80 | .80 | .83 | .80 |
| ADULTOS (NIVEL S) | .87 | .87 | .81 | .82 | .81 | .80 | .78 | .87 | .86 | .84 | .80 | .80 | .83 | .80 |
| TOTAL NIVEL S: | .86 | .87 | .81 | .82 | .81 | .80 | .78 | .87 | .86 | .84 | .80 | .80 | .83 | .80 |

*. Este curso no se tuvo en cuenta para el cálculo de los valores totales del grupo de escolares.

Además del cálculo del alfa de Cronbach, se calcularon dos índices adicionales: el coeficiente alfa ordinal (basado en McDonald, 1999) y el coeficiente omega ordinal (basado en Heise y Bohrnstedt, 1970). Estos coeficientes son dos métodos alternativos para el cálculo de la fiabilidad del test que se han considerado mejores estimadores de la consistencia interna por ajustarse mejor a la naturaleza de las respuestas a los ítems del BAT-7 (de tipo dicotómica, acierto/error). Los valores de los coeficientes ordinales, tanto de los tests como de la puntuación total y de los índices, son muy elevados y en buena parte similares, aunque algo superiores como cabía esperar, a los obtenidos mediante el procedimiento continuo (alfa de Cronbach) debido al grado de unidimensionalidad de los tests (McDonald, 1999; Abad *et al.*, 2011).

Validez: Evidencias de validez interna

Durante el desarrollo del BAT-7 se han recogido diversas evidencias de validez para apoyar la interpretación de las puntuaciones. En primer lugar se analizó la unidimensionalidad de las pruebas de la batería. El procedimiento que se utilizó fue el AFC sobre la matriz de correlaciones tetracóricas. Los datos obtenidos permiten afirmar que los tests son aceptablemente unidimensionales y presentan un elevado grado de homogeneidad. Estos resultados suponen una fuente de evidencia que apoya la solvencia psicométrica de las unidades más básicas que forman la batería: sus ítems.

Para examinar la estructura subyacente de las aptitudes que evalúa el BAT-7 y de las puntuaciones e índices que incorpora, el AFC ha sido la técnica utilizada. Partiendo de los postulados básicos de la teoría CHC sobre la inteligencia y las aptitudes cognitivas (Carroll, 1993, 2005; Cattell, 1971, 1985; Cattell y Horn, 1985; Horn, 1978; Horn, 1988, 1991; McGrew y Wendling, 2010) y del patrón de relaciones entre las variables, se determinó un modelo jerárquico en el que un factor general (*g*; Estrato III) en la cúspide del sistema daba cuenta de las relaciones existentes entre las aptitudes generales (*Gf*, *Gc*, *Gv*, *Gs*, *Gkr*; Estrato II), las cuales incluyen como indicadores una o dos aptitudes específicas (*V* y *O*, *R* y *N*, *M*, *E* y *A*; Estrato I).

BAT-7, BATERÍA DE APTITUDES DE TEA: DESCRIPCIÓN Y DATOS PSICOMÉTRICOS

Se pusieron a prueba cuatro modelos:

Modelo nulo: Modelo de referencia con el que se comparan los demás en el que las aptitudes no se relacionan entre sí.

Modelo 1: Todos los tests dependen de un único factor latente (*g*).

Modelo 2: Un factor superior (*g*) explica el rendimiento en tres variables observables (E, M y A) y en dos factores latentes (*Gf* y *Gc*), los cuales a su vez subyacen a dos variables observables cada uno (R y N, factor *Gf*; V y O, factor *Gc*).

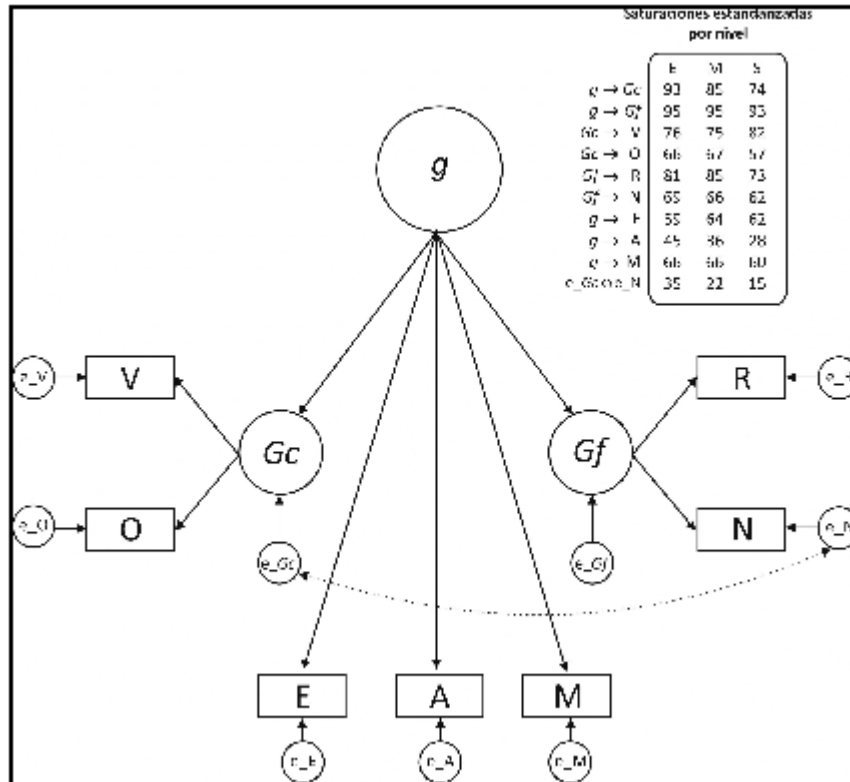
Modelo 3: Modificación parcial del modelo 2 en el que las variables N y *Gc* se relacionan entre sí en mayor medida de lo que cabría esperar en función del modelo.

Con estas premisas se aplicó un AFC multigrupo (en puntuaciones *theta*) para comprobar el ajuste de los cuatro modelos y la invarianza factorial de los tres niveles, es decir, si en cada uno de ellos existe una misma estructura de las aptitudes cognitivas. Para ello se utilizó el programa AMOS 16.0 (Arbuckle, 2007) y el método de estimación de máxima verosimilitud. Los resultados de los análisis (tabla 6) suponen una sólida evidencia sobre la naturaleza de la evaluación con el BAT-7. En términos absolutos, los tres modelos se ajustan a los datos, lo que supone que subyace una dimensión general (*g*) que explica el rendimiento en los tests (modelo 1) y también que dicha dimensión se puede desglosar en aptitudes generales, pero más independientes (modelos 2 y 3), formadas por los tests V y O (*Gc*) y R y N (*Gf*). En términos comparativos, el modelo que mejor ajuste presenta es el que incluye estas aptitudes generales y permite además que el test N tenga relación con el factor cristalizado *Gc* (figura 1), lo que posiblemente refleja la asociación que existe entre la aptitud numérica y la formación académica.

Tabla 6. Índices de ajuste del AFC

| | χ^2 | gl | RMSEA | SRMR | NFI | CFI | ECVI | $\Delta\chi^2$ | Sig. |
|-------------|----------|----|-------|-------|-------|-------|-------|----------------|-------|
| Modelo nulo | 10773,67 | 93 | 0,75 | 0,380 | 0,300 | 0,000 | 1,954 | - | - |
| Modelo 1 | 490,01 | 42 | 0,043 | 0,000 | 0,957 | 0,990 | 0,100 | 10305,655 | 0,000 |
| Modelo 2 | 288,72 | 39 | 0,036 | 0,000 | 0,973 | 0,976 | 0,060 | 179,266 | 0,000 |
| Modelo 3 | 238,06 | 39 | 0,034 | 0,000 | 0,978 | 0,981 | 0,001 | 50,030 | 0,000 |

Figura 1. Modelo de aptitudes cognitivas del BAT-7



DISCUSIÓN

El proyecto de creación del BAT-7, ha seguido un riguroso proceso de desarrollo que ha permitido obtener una completa y sólida batería para la evaluación de las aptitudes intelectuales en el ámbito educativo.

Los estudios sobre la consistencia interna del BAT-7 explorada desde dos perspectivas diferentes indican la consecución de uno de los objetivos que se planteó a la hora de su construcción: la creación de un instrumento psicométrico con un grado de fiabilidad que permita tomar decisiones grupales e individuales con plena garantía.

En cuanto a las evidencias de validez, los datos obtenidos permiten afirmar que los tests son unidimensionales y presentan un elevado grado de homogeneidad. Estos resultados suponen una fuente de evidencia que apoya la solvencia psicométrica de las unidades más básicas que forman la batería: sus ítems.

Además, los análisis factoriales confirmatorios evidencian que las puntuaciones del BAT-7 representan adecuadamente la estructura de las capacidades de orden superior. Tanto la estructura del modelo como la magnitud de las saturaciones estandarizadas, están en consonancia con las teorías actuales sobre la inteligencia y su estructura. Estos datos suponen un apoyo empírico a la capacidad del BAT-7 para estimar el factor *g* de inteligencia y las principales aptitudes generales que lo componen.

En resumen, el BAT-7, *Batería de Aptitudes de TEA*, es una prueba útil, válida y fiable para la evaluación de las aptitudes cognitivas e intelectuales de los escolares. La posibilidad de aplicarse colectivamente supone una gran ventaja para su uso en el ámbito educativo, frente a otras pruebas de evaluación de la inteligencia de aplicación individual.

BAT-7, BATERÍA DE APTITUDES DE TEA: DESCRIPCIÓN Y DATOS PSICOMÉTRICOS

REFERENCIAS

- Abad, F.J., Olea, J., Ponsoda, V. y García, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid: Síntesis.
- Arbuckle, J.L. (2007). *Amos (Versión 16.0) [programa informático]*. Chicago: SPSS.
- Arribas, D., Santamaría, P., Sánchez-Sánchez, F. y Fernández-Pinto, I. (2013). *BAT-7, Batería de aptitudes de TEA, Manual*. Madrid: TEA Ediciones.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. En F. M. Lord y M. R. Novick (Eds.), *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Brody, N. (1997). Intelligence, schooling and society. *American Psychologist*, 52, 1046-1050.
- Burks, S.V., Carpenter, J.P., Goette, L. y Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, 106 (19), 7745-7750.
- Byrne, M.B. (2004). Testing for Multigroup Invariance Using AMOS Graphics: A Road Less Traveled. *Structural Equation Modeling*, 11, 272-300.
- Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, Inglaterra: Cambridge University Press.
- Carroll, J.B. (2005). The three-stratum theory of cognitive abilities. En D.P. Flanagan y P.L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues (2.ª Ed.)*. Nueva York: Guilford Press.
- Cattell, R.B. (1971). *Abilities: Their structure, growth and action*. Boston: Houghton-Mifflin.
- Cattell, R.B. (1985). *Intelligence: Its structure, growth and action*. Amsterdam: North-Holland.
- Cattell, R.B. y Horn, J.L. (1985). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement*, 15, 139-164.
- Colom, R. (2002). En los límites de la inteligencia. ¿Es el ingrediente del éxito en la vida? Madrid: Pirámide.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Deary, I.J. (2012). Intelligence. *Annual Review of Psychology*, 63, 453-482.
- Deary, I.J., Strand, S., Smith, P. y Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, 35, 13-21.
- Detterman, D.K. y Thompson, L.A. (1997). IQ, Schooling, and Developmental Disabilities: What's so Special about Special Education? *American Psychologist*, 52, 1082-1091.
- Duckworth, A.L. y Seligman, M.E.P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, 16, 939-944.
- Duckworth, A.L., Quinn, P.D., Lynam, D.R., Loeber, R. y Stouthamer-Loeber, M. (2011). Role of test motivation in intelligence testing. *Proceedings of the National Academy of Sciences*, 108, 7716-7720.
- Gottfredson, L.S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24, 79-132.
- Gottfredson, L.S. (2000). g: highly general and highly practical. En R. Sternberg y E. Grigorenko (Ed.s), *The general intelligence factor: how general is it?* Nueva York: LEA.
- Hambleton, R. y Swaminathan, H. (1985). *Item Response Theory. Principles and Applications*. Boston: Nijhoff Publishing.
- Haywood, H. (2004). Thinking in, around, and about the curriculum: The role of cognitive education. *International Journal of Disability Development and Education*, 51(3), 231.
- Heise, D. R. y Bohrnstedt, G. W. (1970). Validity, invalidity and reliability. En E. F. Borgatta y G. W. Bohrnstedt (Eds.), *Sociological methodology*. San Francisco: Jossey-Bass.
- Horn, J.L. (1978). Psychometric studies of aging and intelligence. En S. Gershon y A. Raskin (Eds.),

- Genesis and Treatment of Psychological Disorders in the Elderly, vol. 2. Nueva York: Raven.
- Horn, J.L. (1988). Thinking about human abilities. En J. R. Nesselrode y R. B. Cattell (Eds.), *Handbook of multivariate psychology*. Nueva York: Academic Press.
- Horn, J.L. (1991). Measurement of intellectual capabilities: A review of theory. En K. S. McGrew, J. K. Werder y R. W. Woodcock (Eds.), *Woodcock-Johnson Technical Manual*. Chicago: Riverside.
- Jensen, A.R. (1981). *Straight talk about mental tests*. Londres: Methuen.
- Jensen, A.R. (1998). *The g Factor*. Nueva York: Praeger.
- Kuncel, N.R. y Hezlett, S.A. (2007). Standardized tests predict graduate students' success. *Science*, 315, 1080–1081.
- Kuncel, N.R., Hezlett, S.A. y Ones, D.S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, 127, 162–181.
- Lord, F.M. y Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. Reading, Massachusetts: Addison-Wesley Publishing Company.
- Lubinski, D. (2000). Scientific and social significance of assessing individual differences: "Sinking shafts at a few critical points." *Annual Review of Psychology*, 51, 405–444.
- Martínez Arias, R. (1995). *Teoría de los tests psicológicos y educativos*. Madrid: Síntesis.
- McDonald, R.P. (1999). *Test theory: a unified treatment*. Mahwah, NJ: LEA.
- McGrew, K.S. y Wendling, B.J. (2010). Cattell-Horn-Carroll cognitive-achievement relations: what we have learned from the past 20 years of research. *Psychology in the Schools*, 47, 651-675.
- Navas, M. J. (2001). *Métodos, diseños y técnicas de investigación psicológica*. Madrid: UNED.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J. y Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77-101.
- Ree, M. J. y Carretta, T. R. (2002). g2K. *Human Performance*, 15, 3–23.
- Sackett, P. R., Borneman, M. J. y Connelly, B. S. (2008). High-stakes testing in higher education and employment: Appraising the evidence for validity and fairness. *American Psychologist*, 63, 215–227.
- Santamaría, P., Arribas, D., Pereña, J. y Seisdedos, N. (2005). *EFAI. Evaluación Factorial de las Aptitudes Intelectuales*. Madrid: TEA Ediciones.
- Sattler, J.M. (2001). *Assessment of Children Cognitive Applications (4.ª edición)*. San Antonio: Sattler.
- Schmidt, F.L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, 15, 187–210.
- Strenze, T. (2007). Intelligence and socioeconomic success: A metaanalytic review of longitudinal research. *Intelligence*, 35, 401–426.
- Süß, H.M. (2001). Prädiktive Validität der Intelligenz im schulischen und außerschulischen Bereich. En E. Stern. y J. Guthke, (Eds.), *Perspektiven der Intelligenzforschung*. Lengerich: Pabst.

1 Las Comunidades autónomas que se agrupan en cada zona son: Centro: Extremadura, Comunidad de Madrid, Castilla La Mancha y Aragón; Este: Cataluña, Comunidad Valenciana, Murcia y Baleares; Norte: Galicia, Asturias, Cantabria, País Vasco, Navarra, La Rioja y Castilla-León; y Sur: Andalucía, Ceuta, Melilla y Canarias.

