

Diseño, confiabilidad y validez del Test Básico de Aprendizaje

Design, reliability and validity of the Basic Learning Test

Autor

Pablo San Martín Catalán. Psicopedagogo de la Universidad Tecnológica de Chile. Diplomado en Bioestadística del Instituto de Matemática, Física y Estadística de la Universidad de Las Américas, Chile. Investigador, Centro de Estudios Psicométricos y del Aprendizaje (CEPA). Antofagasta, Chile.

✉ psanmartin@cepaonline.cl

 <https://orcid.org/0000-0002-2230-1635>

Resumen

El presente estudio describe el diseño y las propiedades psicométricas del Test Básico de Aprendizaje (TBA). Esta batería de evaluación psicopedagógica multidimensional, informatizada y estandarizada, está diseñada para evaluar un amplio espectro de variables cognitivas y habilidades específicas relacionadas con los aprendizajes formales. Se trabajó con un enfoque teórico cuantitativo, utilizando una muestra de hasta 5321 personas (población chilena) entre 6 y 55 años, de distintos niveles académicos. En la integración cuantitativa, se llevaron a cabo Análisis Factoriales Exploratorios (AFE), Confirmatorios (AFC) y análisis de consistencia interna con alfa de Cronbach. Se obtuvo un diseño lógico y consistente para la medición de variables cognitivas y habilidades académicas, alcanzando una alta confiabilidad ($\alpha > 0.7$). Los análisis factoriales arrojaron resultados convergentes con las hipótesis teóricas. En conclusión, el TBA es un instrumento válido y confiable para la evaluación del perfil y potencial de aprendizaje de una persona.

Palabras claves: validez, confiabilidad, aprendizaje, batería psicopedagógica.

Abstract

The present study describes the design and psychometric properties of the Test Básico de Aprendizaje (TBA). This multidimensional, computerized and standardized psychopedagogical assessment battery is designed to assess a broad spectrum of cognitive variables and specific skills related to formal learning. We worked with a quantitative theoretical approach, using a sample of up to 5321 people (Chilean population) between 6 and 55 years old, from different academic levels. In the quantitative integration, Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA) and internal consistency analysis with Cronbach's alpha were carried out. A logical and consistent design was obtained for the measurement of cognitive variables and academic skills, achieving high reliability ($\alpha > 0.7$). Factor analyses yielded results convergent with the theoretical hypotheses. In conclusion, the TBA is a valid and reliable instrument for the evaluation of a person's learning profile and potential.

Keywords: validity, reliability, learning, psychopedagogical battery.

Sentido del Test Básico de Aprendizaje

El Test Básico de Aprendizaje (TBA) se fundamenta en la lógica del escalamiento cognitivo y está diseñado para evaluar a individuos en un rango etario de 6 a 55 años. Esta batería de pruebas mide capacidades y habilidades, en lugar de conocimientos o contenidos curriculares. Al referirnos al escalamiento cognitivo, las preguntas clave que

guían los algoritmos de análisis del TBA son: ¿En qué nivel de desarrollo cognitivo se encuentra la persona evaluada? ¿Cuál es su nivel de desarrollo en habilidades académicas?

Para responder a estas preguntas, se consideran factores como la edad y la experiencia escolar del evaluado. Se espera que un estudiante desarrolle habilidades de lectura, escritura y matemáticas en función de su experiencia educativa acumulada. De manera similar, en el ámbito cognitivo, se anticipa un desarrollo correspondiente de la memoria, la atención y el pensamiento lógico, de acuerdo con su etapa evolutiva.

Por ejemplo, en el contexto del escalamiento cognitivo, una prueba compuesta por varios ítems presentará un aumento progresivo en la dificultad. El evaluado avanzará hasta el punto en que logre comprender el material; cualquier ítem que supere su capacidad de entendimiento puede ser omitido sin repercusiones en la evaluación general.

En este sentido, el instrumento no se orienta hacia un diagnóstico específico, sino que ofrece un perfil de aprendizaje a partir del cual se pueden inferir ciertas condiciones, incluidas las dificultades específicas del aprendizaje. El diagnóstico debe basarse en criterios clínicos y en evidencia científica.

El TBA está diseñado para ser aplicado a toda persona que corresponda al rango de estandarización, independientemente de su condición. Por lo tanto, debe utilizarse únicamente con individuos cuyas características cognitivas, emocionales y físicas les permitan interactuar con un aparato tecnológico y responder a estímulos visuales y auditivos. Además, para la medición de las variables cognitivas, se considerará únicamente la edad, mientras que la evaluación de habilidades académicas se restringirá a sujetos lectores.

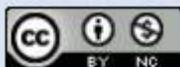
Estructura y funcionamiento del Test Básico de Aprendizaje

El origen de los instrumentos científicos de medición cognitiva se remonta al siglo XIX con las ideas de Galton sobre la psicología comparada, centradas en la inteligencia humana y su heredabilidad. Estas ideas, de alguna manera, sentaron las bases de la psicometría junto con Cattell y el concepto de “test mental” (Morales, 1990).

Posteriormente, en el siglo XX, en Francia, Binet y Simon desarrollaron la primera investigación científica para la medición de variables complejas. Este estudio fue solicitado por la Secretaría de Educación Pública del gobierno francés con el objetivo de identificar diferencias cognitivas y de aprendizaje en los estudiantes, permitiendo así brindar un apoyo educativo diferenciado (Morales, 1990).

Más adelante, Lewis Terman adaptó el Test Binet-Simon para su uso en Estados Unidos, en un trabajo realizado desde la Universidad de Stanford. Este instrumento, denominado Test Stanford-Binet, se hizo popular en la medición de la inteligencia general. En este contexto, se introdujeron los términos "edad mental" y "cociente intelectual (CI)", propuestos por Stern en el siglo XX. El CI se define como la relación entre la edad mental (EM) y la edad cronológica (EC), con un valor esperado de 1 o de 100 si se emplea la fórmula $CI = 100EM/EC$. Por ejemplo, una persona con una EM de 8 años y una EC de 10 años obtendría un CI de 80, lo que indicaría que su desempeño cognitivo promedio corresponde al de un niño de 8 años en lugar de 10, como se esperaría (Maureira, 2017).

Posteriormente, Spearman y Thurstone comenzaron a estudiar estructuras matemáticas complejas para la comprensión de la inteligencia humana y las habilidades asociadas. Este trabajo dio origen a los análisis factoriales, los cuales permiten identificar factores subyacentes que explican las varianzas de las variables medidas en los tests (Pérez y Medrano, 2010).



En cuanto al estudio de la confiabilidad de los tests, destacan los aportes de Cronbach, quien ajustó y utilizó un modelo matemático para evaluar el grado de consistencia y precisión de los instrumentos de medición. Según este modelo, un test es más preciso cuando la varianza de lo verdadero representa una mayor proporción de la varianza del resultado empírico u observado. En esencia, esta técnica se basa en el análisis de las correlaciones entre los ítems para determinar la coherencia interna de un instrumento (Celina y Campo, 2005).

Con respecto a las bases normativas de un instrumento psicométrico, existen directrices internacionales para el desarrollo de tests fiables, válidos, accesibles, con seguridad informática y basados en tecnología, incluyendo el uso de inteligencia artificial y aprendizaje adaptativo, de acuerdo con las Directrices para la Evaluación Basada en Tecnología (International Test Commission & Association of Test Publishers, 2022).

En este sentido, el TBA fue desarrollado en el Centro de Estudios Psicométricos y del Aprendizaje, siguiendo las directrices mencionadas anteriormente.

Definiciones de las variables integradas en el Test Cognitivo General (TCG)

La variable Bases del Pensamiento Lógico (BPL) se sustenta con el modelo teórico de la dinámica cognitiva del aprendizaje humano, desarrollado por San Martín y Carrasco (2022). Este modelo propone que el pensamiento, y en su actividad, razonamiento, es una función superior que hace posible el flujo mental entre la estructura previa del sujeto (esquemas de aprendizaje) y el objeto de aprendizaje, cuya dificultad es relativa a la inteligencia general (San Martín y Carrasco, 2022). Por esta razón, entre más difícil es el reactivo, más capacidad de deducción, inducción y espacialidad se requiere para su solución.

La variable Velocidad Atencional (A) se sustenta con el modelo de la dinámica cognitiva del aprendizaje humano, de acuerdo con el principio lógico b del sujeto que aprende (San Martín y Carrasco, 2022). En este sentido, A se entiende como un vector que dirige la capacidad cognitiva de las personas hacia ciertos estímulos, cuyo comportamiento cambia en función del tiempo y es ponderable con aciertos, errores y omisiones. Entonces, conforme con la ecuación propuesta por San Martín y Carrasco (2022), la velocidad atencional es la cantidad de atenciones que un sujeto genera por unidad de tiempo.

La variable Memoria Inmediata (MI), de acuerdo con San Martín y Carrasco (2022), es una función mental que permite codificar, almacenar, asociar y recuperar elementos en un tiempo determinado, cuya retentiva decrece en función del tiempo y su promedio mínimo y máximo de evocación tiene un límite matemático de siete más menos dos elementos (7 ± 2), dependiendo de la cantidad de objetos y no de la información requerida (bits) (Miller, 1956).

La variable Tasa de Aprendizaje (TA) se refiere a la proporción que determina la velocidad con la que un sujeto desarrolla habilidades. Esta variable permite predecir la curva de aprendizaje, proporcionando una medida de cómo un individuo adquiere y mejora sus competencias a lo largo del tiempo.

La variable Velocidad de Respuesta (VR) se integra en las Bases del Pensamiento Lógico y se define, en un contexto de problemas lógicos, como la cantidad de aciertos que un sujeto logra por unidad de tiempo. Esto se refiere al tiempo que un individuo tarda en encontrar una solución a un problema nuevo sin recurrir a conocimientos previos.

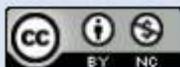
Tabla 1.

Diseño para la medición de variables a través del Test Cognitivo General TCG



Variable	Tiempo y/o lógica de prueba	Cantidad de reactivos	Fórmula de corrección	Diccionario
BPL	30 minutos	29	$\sum a$	Sumatoria de aciertos
A	2 minutos	1	$\frac{60 * (\sum a - (e + o))}{t (s)}$	a: acierto e: error o: omisión t (s): tiempo en segundos
MI	Tiempo variable. Ilimitado en general. 2 ítems con 30 segundos para la memorización de elementos, luego 1 minuto para la recuperación de elementos. 2 ítems de 5 segundos para la memorización de dictado de palabras, luego 1 minuto para la recuperación	4	$\frac{at1(v) + at2(v) + at3(a) + at4(a)}{4}$	at 1(v): aciertos en tarea 1 visual at 2(v): aciertos en tarea 2 visual at 3(a): aciertos en tarea 3 auditiva at 4(a): aciertos en tarea 4 auditiva
TA	Tiempo ilimitado. El reactivo se repite 8 veces y el sistema toma el tiempo de cada solución	1	Matriz: $\begin{matrix} t1 & t2 \\ t2 & t4 \\ t4 & t8 \end{matrix}$ $1 - \left(\frac{\frac{t2}{t1} + \frac{t4}{t2} + \frac{t8}{t4}}{3} \right)$	t1: tiempo primer ensayo t2: tiempo segundo ensayo t4: tiempo cuarto ensayo t8: tiempo octavo ensayo
VR	Es subvariable de BPL	29	$\frac{60 * \sum a}{t (s)}$	a: acierto t (s): tiempo en segundos

Fuente: elaboración propia para la validez de las puntuaciones empíricas en base a la modelación psicométrica de las ecuaciones del modelo de la dinámica cognitiva del aprendizaje humano de San Martín y Carrasco (2022).



Definiciones de las variables integradas en el Test de Aprendizajes Formales (TAF)

La variable Bases del Aprendizaje Matemático (BAM), se define como la capacidad de desarrollar conceptos numéricos y establecer relaciones tanto concretas como abstractas entre cantidades. Además, implica comprender conceptos probabilísticos y extraer información implícita de problemas de lógica matemática. Estas habilidades corresponden a las competencias matemáticas subyacentes propuestas por Chamorro (2003).

La variable Bases del Aprendizaje de la Lectoescritura (BALE) se define como la estructura mínima necesaria para el desarrollo de habilidades lectoescritoras. Esta variable incluye subvariables como la conciencia fonológica, la activación de esquemas fonológicos y las reglas gráficas, que son esenciales para la asociación de componentes alfabéticos (Cuetos, 2012).

La variable Comprensión Lectora (CL) se define como la capacidad de construir significados a través de la decodificación del lenguaje escrito, ya sea de manera literal o implícita (Cuetos, 1996).

La variable Velocidad Lectora (VL) se define como la cantidad de palabras leídas en voz alta por unidad de tiempo, siendo la medida universal utilizada palabras por minuto (p/min) (Fry, 1975).

La variable Ortografía Visual (OV) se define como la capacidad de formar imágenes mentales de palabras que son ortográficamente correctas (Cuetos, 2012).

La variable Ortografía Fonética (OF) se define como la capacidad de convertir unidades fonológicas en grafemas (Cuetos, 2012).

Tabla 2.

Diseño para la medición de variables a través del Test de Aprendizajes Formales TAF

Variable	Tiempo de prueba	de Cantidad reactivos	de	Fórmula de corrección	de Diccionario
BAM	30 minutos	26		$\sum a$	Sumatoria de aciertos
BALE	10 minutos	17		$\sum a$	Sumatoria de aciertos
CL	12 minutos	12		$\sum a$	Sumatoria de aciertos
VL	Ilimitado	1		$\frac{60 * P}{t (s)}$	p: cantidad de palabras que contiene el texto t (s): tiempo en segundos
OV	5 minutos	10		$\sum a$	Sumatoria de aciertos
OF	5 minutos	6		$\sum a$	Sumatoria de aciertos

Fuente: elaboración propia en base a la modelación psicométrica para la validez de las puntuaciones empíricas.



Definición y justificación de la Escala de Autopercepción del Aprendizaje (EAA)

La Escala de Autopercepción del Aprendizaje (EAA) es un instrumento diseñado para evaluar cómo los individuos perciben y valoran su propio proceso de aprendizaje en diferentes contextos educativos. La autoevaluación y la autopercepción desempeñan un papel crucial en el aprendizaje, permitiendo a las personas reflexionar sobre sus capacidades, identificar áreas de mejora y ajustar sus estrategias de estudio de manera autónoma.

La autopercepción se entiende, según Alexander (2006), como un tipo de creencias o imágenes motivacionales que representan "los modelos internos que las personas construyen acerca de sí mismas en relación con todas las áreas del desarrollo". Estas autopercepciones influyen en cómo se interpretan las situaciones de aprendizaje y orientan el desempeño académico. Además, son específicas por dominio o área de conocimiento, dinámicas y sensibles a las experiencias vividas en contextos académicos y sociales. Pueden manifestarse de formas más o menos conscientes y explícitas, y no siempre corresponden a la realidad (Garello, 2008, pp. 252-253).

En cuanto a la autopercepción del aprendizaje, es fundamental señalar la influencia del autoconcepto, que es un factor importante en la orientación de la conducta humana. El autoconcepto se refiere a la percepción que las personas tienen de sí mismas en un momento determinado (Bolívar y Rojas, 2015, p. 4). En el ámbito del aprendizaje, el autoconcepto se construye a partir de experiencias en diversos contextos, como el familiar, social, académico y laboral.

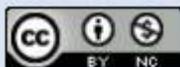
Los fundamentos para la construcción de la variable autopercepción del aprendizaje están dados por varias dimensiones e indicadores (Orozco y García, 2017). La autopercepción del aprendizaje se define como "la manera en que una persona se percibe a sí misma como aprendiz", abarcando las creencias, expectativas y actitudes sobre sus habilidades y capacidades de aprendizaje. Esta variable incluye las siguientes dimensiones e indicadores:

- Dimensión Cognitiva: Satisfacción y progreso en el aprendizaje.
- Dimensión Instruccional: Percepción del esfuerzo y la competencia.
- Dimensión Motivacional: Actitud hacia el aprendizaje.
- Dimensión Afectiva: Autoconfianza en el aprendizaje.

Sistema de análisis psicométrico

Para definir los niveles utilizando el TBA, se toma en cuenta la estimación del puntaje directo (PD), así como la media de la población correspondiente al grupo etario y a la experiencia escolar, junto con la desviación estándar de cada variable. Posteriormente, se calcula el valor Z (puntuación tipificada) para determinar la posición en la distribución normal. En este sentido, la batería trabaja con los niveles bajo, medio bajo, medio, medio alto y alto, asignándose conforme con la cantidad desviaciones típicas, donde por cada desviación por debajo o sobre la media, se genera un nivel de desarrollo distinto, considerándose el intervalo medio aquel comprendido entre -1 y 1 (Z) (>-1 y <1). Toda posición que alcance 2 o más desviaciones típicas, sobre o por debajo de la media, se considera alto o bajo, respectivamente (Morales, 1990).

Se consideran diversas medidas, tales como Memoria Inmediata (MI), Bases del Pensamiento Lógico (BPL), Velocidad de Respuesta (VR), Velocidad Atencional (A), Tasa de Aprendizaje (TA), Bases del Aprendizaje Matemático (BAM), Bases del Aprendizaje de



la Lectoescritura (BALE), Comprensión Lectora (CL), Velocidad Lectora (VL), Ortografía Visual (OV) y Ortografía Fonética (OF). A partir de estas variables, se generan indicadores globales, ponderando grupos de variables como: MI, BPL, A y TA (indicador cognitivo); BAM (indicador de matemáticas); BALE, CL y VL (indicador de lectura); y OV y OF (indicador de escritura (ortografía)).

Para la estimación del desarrollo cognitivo bajo un único factor, se consideran BPL, MI y A, calculando el Cociente Intelectual (CI) con una media de 100 y una desviación estándar de 15, así como el puntaje T con una media de 50 y una desviación estándar de 10; a este proceso se le denomina síntesis psicométrica (Morales, 1990).

Finalmente, para la estimación del cambio en las variables, se utiliza el cálculo basado en T, aplicando la ecuación $(T2-T1) / T1$ en relación con los indicadores globales de cognición, lectura, matemáticas y escritura.

Método

La investigación se desarrolló bajo un diseño cuantitativo transeccional no experimental (Sousa et al., 2007) para construir y validar la batería de evaluación psicopedagógica TBA en Chile. El proceso de estandarización del Test Básico de Aprendizaje (TBA) en Chile se inició en 2023 con el propósito de desarrollar un instrumento válido y confiable para evaluar el perfil de aprendizaje de las personas, independientemente del contexto en el que se encuentren. Este estudio involucró una muestra inicial compuesta por 1082 sujetos, de los cuales 895 fueron incluidos en el análisis específico presentado. Los participantes, tanto hombres como mujeres, abarcaron un amplio rango etario, desde la niñez hasta la adultez, y provenían mayoritariamente de la región de Antofagasta y la región Metropolitana.

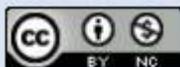
La muestra específica de 895 personas se distribuyó en cinco rangos de edad de la siguiente manera: 6 a 8 años con 194 personas (22%), 9 a 11 años con 241 personas (27%), 12 a 14 años con 186 personas (21%), 15 a 17 años con 112 personas (13%) y 18 a 55 años con 162 personas (18%). Esta submuestra constituyó una fracción representativa del total de 1082 personas, proporcionando una base estable para el análisis del perfil de aprendizaje.

El estudio principal utilizó un enfoque de muestreo combinado que integró tanto muestras incidentales como probabilísticas. Se aplicaron métodos de muestreo probabilístico simple y estratificado, lo que resultó fundamental para analizar patrones cognitivos y habilidades académicas en diferentes rangos de edad.

La aplicación del TBA se realizó mediante un sistema informático llamado Sistema TBA (software), que permite la interacción del sujeto con un dispositivo con conexión a internet, ya sea una tablet, un computador de escritorio o un notebook. En este proceso se utilizaron los tres tipos de dispositivos.

A razón de lo anterior, en 2024, se llevó a cabo una validación de los resultados obtenidos en el estudio inicial, utilizando una nueva muestra probabilística, integrando a personas de diversas regiones del país (Chile). Esta segunda fase tuvo como objetivo validar la consistencia de los hallazgos preliminares. Los resultados obtenidos confirmaron la fiabilidad del TBA como instrumento de evaluación del perfil de aprendizaje.

La validación de los resultados se realizó con una población de 2100 personas, de las cuales se seleccionaron al azar 1177 (56%). La distribución de la muestra en diferentes rangos de edad fue la siguiente: 6 a 7 años con 143 personas (12%), 8 a 9 años con 230 personas



(20%), 10 a 12 años con 267 personas (23%), 13 a 14 años con 156 personas (13%), 15 a 17 años con 151 personas (13%) y 18 a 55 años con 230 personas (20%).

El análisis del Test-Retest para el TBA se llevó a cabo aplicando 10 pruebas (TCG y TAF) a tres personas seleccionadas al azar, con un plazo de 2 semanas entre la primera y segunda aplicación.

El análisis de datos se realizó utilizando diversas funciones en R, distribuidas a través de varios paquetes. El proceso metodológico fue el siguiente:

En primer lugar, se utilizó la función *alpha* del paquete *psych* para calcular el coeficiente alfa de Cronbach, lo que permitió evaluar la consistencia interna de las escalas. Este coeficiente es una medida comúnmente empleada para determinar la fiabilidad de los instrumentos de medición.

Para evaluar la adecuación de la muestra para el análisis factorial, se aplicó la función *KMO*, también del paquete *psych*, que realiza la prueba de Kaiser-Meyer-Olkin. Esta prueba determina si los datos son adecuados para un análisis factorial, evaluando las correlaciones entre las variables.

La prueba de esfericidad de Bartlett se llevó a cabo mediante la función *bart_spher* del paquete *REdaS*, la cual verifica la hipótesis nula de que las variables no están correlacionadas, lo que es un requisito esencial para aplicar el análisis factorial.

El análisis factorial exploratorio se ejecutó utilizando la función *fa* del paquete *psych*, que permitió extraer un único factor utilizando el método de máxima verosimilitud (*ml*) y una rotación *Varimax*. Los resultados fueron visualizados mediante la función *fa.diagram*, que generó un diagrama que facilita la interpretación de las relaciones entre las variables y los factores extraídos.

Para la detección de valores atípicos, se utilizó la función *mahalanobis* para calcular las distancias, lo que permitió identificar puntos atípicos en el conjunto de datos. Posteriormente, se aplicó un umbral basado en la distribución chi-cuadrado para clasificar los valores atípicos y se filtraron los casos no atípicos utilizando índices lógicos.

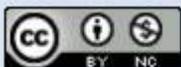
El análisis factorial paralelo se realizó utilizando la función *fa.parallel*, también del paquete *psych*, que ayudó a determinar el número adecuado de factores a extraer mediante un análisis iterativo.

Luego, para el análisis factorial confirmatorio (AFC), se utilizó el paquete *lavaan*. El modelo fue especificado mediante la función *cfa*, que permite ajustar el modelo teórico a los datos. Los resultados del ajuste del modelo fueron resumidos con la función *summary*, que mostró medidas de ajuste, coeficientes estandarizados y otras estadísticas relevantes.

La visualización del modelo de AFC se llevó a cabo con la función *lavaanPlot* del paquete *lavaanPlot*, que generó un diagrama que facilita la interpretación gráfica de las relaciones entre las variables latentes y sus indicadores, incluyendo los coeficientes estandarizados.

El Análisis Factorial de la batería tridimensional (TBA) se calculó sobre una muestra de 5321 personas.

Por último, para evaluar la consistencia test-retest, se utilizó el coeficiente de correlación de Pearson y un modelo de regresión lineal simple. El análisis se realizó con los siguientes códigos: primero se instaló y cargó el paquete *readxl* mediante *install.packages("readxl")* y *library(readxl)*. Posteriormente, se calcularon las correlaciones con la función *cor(Test_Retest, method = "pearson")* y se ajustó un modelo de regresión



lineal simple con $lm(Retest \sim Test, data = Test_Retest)$, permitiendo analizar tanto la relación lineal como la capacidad predictiva entre las puntuaciones de ambas mediciones.

Resultados

Estudio de fiabilidad del Test Básico de Aprendizaje (TBA)

Se realizó un análisis de consistencia interna mediante el cálculo del coeficiente alfa de Cronbach para evaluar la fiabilidad de diversos instrumentos de medición utilizados en el estudio. A través del entorno RStudio, se obtuvieron los valores correspondientes para cada instrumento, y se evaluaron los intervalos de confianza al 95% y el efecto de la eliminación de ítems sobre la consistencia interna.

Tabla 3.

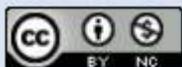
Instrumento	Alfa de Cronbach	Intervalo de Confianza (95%)	Alfa si se elimina el ítem mínimo	Alfa si se elimina el ítem máximo
Bases del Pensamiento Lógico	0.77	0.74 - 0.80	0.75	0.78
Memoria Inmediata	0.70	0.67 - 0.73	0.61	0.66
Bases del Aprendizaje Matemático	0.79	0.76 - 0.81	0.77	0.79
Bases del Aprendizaje de la Lectoescritura	0.87	0.85 - 0.88	0.85	0.89
Comprensión Lectora	0.85	0.83 - 0.87	0.83	0.85
Ortografía Fonética	0.66	0.62 - 0.69	0.58	0.65
Ortografía Visual	0.76	0.73 - 0.78	0.72	0.76
Autopercepción del Aprendizaje	0.90	0.87 - 0.93	0.88	0.90

Fuente: Elaboración propia.

La *tabla 3* presenta los resultados del análisis de consistencia interna para diversos instrumentos de medición, evaluados mediante el coeficiente alfa de Cronbach. El alfa de Cronbach para cada instrumento refleja la fiabilidad de este, indicando la consistencia interna de los ítems que lo componen. Los valores varían entre 0.66 (Ortografía Fonética) y 0.90 (Autopercepción del Aprendizaje), lo que sugiere una consistencia interna desde moderada hasta alta (Celina y Campo, 2005).

El intervalo de confianza al 95% indica el rango dentro del cual se espera que se ubique el valor real del alfa de Cronbach en la población, proporcionando una medida de precisión. En todos los casos, el intervalo es estrecho, lo que refuerza la estabilidad de las estimaciones. Asimismo, se presenta el alfa si se elimina el ítem mínimo o máximo, mostrando el efecto que tendría la eliminación de los ítems con menor o mayor influencia en la consistencia interna. En efecto, la eliminación de cualquier ítem no mejora significativamente el alfa, lo que indica que todos los ítems contribuyen adecuadamente a la fiabilidad de los instrumentos.

En lo que respecta al estudio de fiabilidad de las variables Atención (A), Tasa de Aprendizaje (TA) y Velocidad Lectora (VL), la naturaleza de estas variables no exige



cálculos de consistencia interna, ya que su lógica de medición se fundamenta en ecuaciones matemáticas de análisis directo. Por lo tanto, se consideran como ítems únicos.

Estudio de validez del Test Cognitivo General (TCG)

El análisis factorial (AF) fue implementado para validar la estructura subyacente del Test Cognitivo General (TCG), con el objetivo de confirmar la existencia de una dimensión latente que explique las relaciones entre sus ítems.

Previamente al análisis factorial, se evaluó la adecuación de los datos mediante el índice de Kaiser-Meyer-Olkin (KMO) y la prueba de esfericidad de Bartlett, siguiendo los criterios establecidos por Garrido Bermúdez et al. (2023). El KMO obtuvo un valor global de 0.65, indicando una relación moderada entre las variables, aunque por debajo del rango ideal de 0.7 a 1 para un análisis factorial óptimo.

La prueba de esfericidad de Bartlett arrojó un estadístico chi-cuadrado de 353.37 con 3 grados de libertad, alcanzando un nivel de significancia menor a 0.001. Esto permitió rechazar la hipótesis nula de ausencia de correlación entre las variables, confirmando que los datos son aptos para el análisis factorial.

En este caso, con un KMO cercano a 0.65 y una prueba de Bartlett adecuada para el análisis factorial exploratorio (AFE), se puede afirmar que las tres variables están fuertemente relacionadas. Además, al calcular el efecto de la ausencia de una de ellas (correlación parcial), se observa que el resto de las variables mantiene una relación fuerte entre sí. Esto puede justificarse mediante los modelos cognitivos que destacan cómo cada función cognitiva influye en las demás, relacionándose de forma conjunta o en subconjuntos dentro de las operaciones mentales. Este fenómeno refuerza la pertinencia del análisis factorial, al evidenciar que las variables del TCG están interconectadas y comparten un componente común que las vincula conceptualmente.

Con base en estos resultados, se concluye que el conjunto de datos satisface los requisitos mínimos para el análisis factorial, respaldando la viabilidad de identificar una dimensión latente que sustente la estructura del TCG.

Tabla 4.

Prueba	Valor	Interpretación
KMO (Global)	0.65	Marginalmente aceptable, indica relaciones moderadas entre las variables.
KMO por variable	A: 0.63, MI: 0.64, BPL: 0.68	Cada variable muestra una contribución moderada al modelo factorial.
Prueba de Bartlett	$X^2 = 353.37, p < 0.001$	Rechaza la hipótesis nula, justificando la aplicación del análisis factorial.

Fuente: Elaboración propia.

El AFE realizado con un enfoque de máxima verosimilitud y rotación Varimax permitió identificar un único factor subyacente que explica la varianza de los ítems analizados. Los resultados de las pruebas iniciales de adecuación fueron los siguientes:



Tabla 5.

Ítem	Carga factorial (ML1)	Varianza explicada (h ²)	Varianza no explicada (u ²)
A	0.67	0.46	0.54
MI	0.63	0.40	0.60
BPL	0.55	0.30	0.70

Fuente: Elaboración propia.

Las cargas factoriales, que oscilan entre 0.55 y 0.67, son moderadas y respaldan la contribución de cada ítem al factor latente. Sin embargo, parte de la varianza no está explicada por este factor único, lo que podría indicar la influencia de otras dimensiones no consideradas en el modelo.

Además, se realizó un Análisis Factorial Confirmatorio (AFC) para validar la estructura de un factor latente denominado "Cognitivo". Los índices de bondad de ajuste obtenidos son los siguientes:

Tabla 6.

Índice	valor	Criterios aceptados	Interpretación
CFI	1	> 0.95 = Excelente ajuste	El modelo ajusta perfectamente los datos.
TLI	1	> 0.95 = Excelente ajuste	Confirma la simplicidad y ajuste del modelo.
RMSEA	0	< 0.05 = Excelente ajuste	No hay discrepancia entre el modelo y los datos.
SRMR	0	< 0.08 = Buen ajuste	Los residuos entre las correlaciones predichas y observadas son mínimos.

Fuente: Elaboración propia.

Aunque los índices de ajuste sugieren un ajuste perfecto en este caso, es importante considerar que el valor de 1 en índices como el CFI y TLI, junto con un RMSEA de 0, podría ser indicativo de un modelo que ajusta bien, pero en algunos contextos puede también reflejar que el modelo es demasiado simple o que los datos son demasiado adecuados para el modelo.

Los resultados muestran que las cargas factoriales de los ítems son estadísticamente significativas:

Tabla 7.

Ítem	Carga factorial estandarizada
A	0.675
MI	0.631
BPL	0.548

Fuente: Elaboración propia.

Las cargas factoriales confirman que la variable A es la que más contribuye al factor cognitivo (0.675), mientras que BPL tiene la menor carga (0.548). Sin embargo, los valores indican una estructura unidimensional consistente con el AFE. A continuación, se presenta el modelo del factor cognitivo evaluado a través del TCG, considerando sus cargas factoriales calculadas en el AFC:

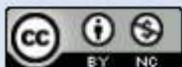
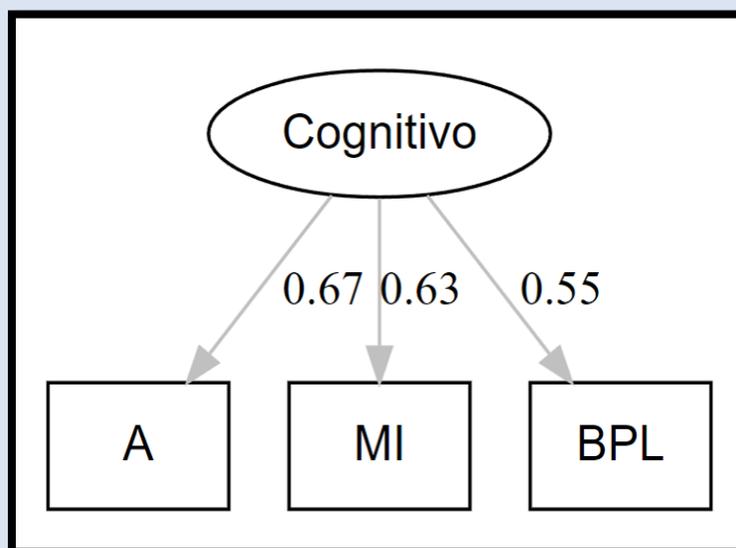


Figura 1.



Fuente: Elaboración propia. Generado en RStudio.

Estudio de validez del Test de Aprendizajes Formales (TAF)

El análisis factorial fue implementado para validar la estructura subyacente del TAF, con el objetivo de confirmar la existencia de una dimensión latente que explique las relaciones entre sus ítems.

Previamente al análisis factorial, se evaluó la adecuación de los datos mediante el índice de Kaiser-Meyer-Olkin (KMO) y la prueba de esfericidad de Bartlett, siguiendo los criterios establecidos por Garrido Bermúdez et al. (2023).

El KMO obtuvo un valor global de 0.83, indicando una relación adecuada entre las variables para el análisis factorial.

La prueba de esfericidad de Bartlett arrojó un estadístico chi-cuadrado de 1487.54 con 15 grados de libertad, alcanzando un nivel de significancia menor a 0.001. Esto permitió rechazar la hipótesis nula de ausencia de correlación entre las variables, confirmando que los datos son aptos para el análisis factorial.

Con base en estos resultados, se concluye que el conjunto de datos satisface los requisitos mínimos para el análisis factorial, respaldando la viabilidad de identificar una dimensión latente que sustente la estructura del TAF.

Tabla 8.

Prueba	Valor	Interpretación
KMO (Global)	0.83	Adecuado, indica relaciones fuertes entre las variables.
KMO por variable	BAM: 0.84, BALE: 0.85, CL: 0.79, OV: 0.84, OF: 0.85, VL: 0.84	Cada variable muestra una contribución adecuada al modelo factorial.
Prueba de Bartlett	$X^2 = 1487.54$, $p < 0.001$	Rechaza la hipótesis nula, justificando la aplicación del análisis factorial.

Fuente: Elaboración propia.

El AFE, realizado con un enfoque de máxima verosimilitud y rotación Varimax, permitió identificar un único factor subyacente que explica la varianza de los ítems analizados.

Tabla 9.

Ítem	Carga factorial (ML1)	Varianza explicada (h^2)	Varianza no explicada (u^2)
BAM	0.70	0.49	0.51
BALE	0.53	0.28	0.72
CL	0.83	0.69	0.31
OV	0.78	0.61	0.39
OF	0.73	0.53	0.47
VL		0.08	0.92

Fuente: Elaboración propia.

Las cargas factoriales respaldan la contribución de la mayoría de los ítems al factor latente, excepto para VL, cuya carga es 0. Esto podría justificar la exclusión de este ítem en futuras iteraciones del modelo.

Se realizó un AFC para validar la estructura del factor latente denominado "Habilidades". Los índices de bondad de ajuste obtenidos son los siguientes:

Tabla 10.

Índice	Valor	Criterios aceptados	Interpretación
CFI	0.948	> 0.95 = Excelente ajuste	Ajuste adecuado del modelo.
TLI	0.913	> 0.95 = Excelente ajuste	Confirma la simplicidad y ajuste del modelo.
RMSEA	0.109	< 0.05 = Excelente ajuste	Ajuste moderado; discrepan ligeramente las predicciones del modelo con los datos.
SRMR	0.046	< 0.08 = Buen ajuste	Residuos mínimos entre las correlaciones predichas y observadas.

Fuente: Elaboración propia.

Los resultados muestran que las cargas factoriales de los ítems son estadísticamente significativas:

Tabla 11.

Ítem	Carga factorial estandarizada
BAM	0.704
BALE	0.533
CL	0.831
OV	0.780
OF	0.725
VL	-0.282

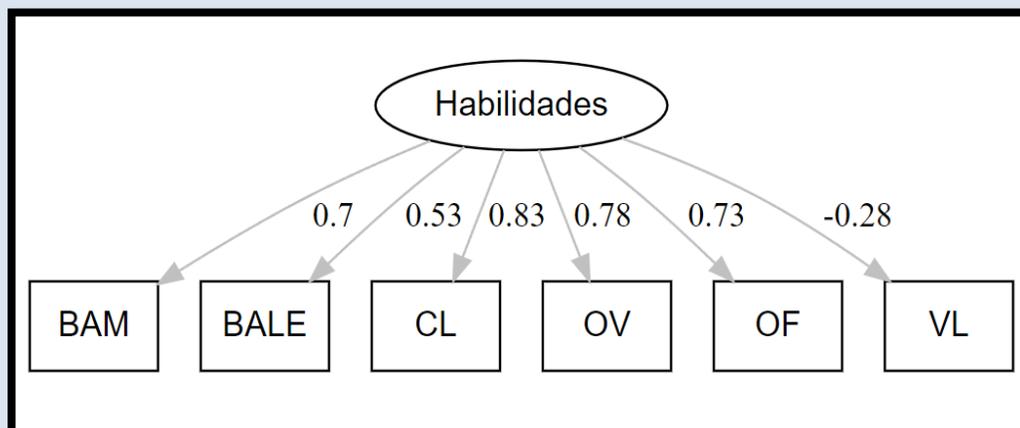
Fuente: Elaboración propia.

Las cargas factoriales confirman que el ítem CL es el que más contribuye al factor cognitivo (0.831), mientras que VL presenta un coeficiente negativo, sugiriendo su posible irrelevancia en la estructura latente. A pesar de ello, los resultados generales son consistentes con la estructura unidimensional identificada en el AFE.



A continuación, se presenta el modelo del factor habilidades académicas evaluado a través del TAF, considerando sus cargas factoriales calculadas en el AFC:

Figura 2.



Fuente: Elaboración propia. Generado en RStudio.

Estudio de validez de la Escala de Autopercepción del Aprendizaje (EAA)

Para determinar la viabilidad del análisis factorial, se calculó el índice de Kaiser-Meyer-Olkin (KMO) y se realizó la prueba de esfericidad de Bartlett.

Tabla 12.

Prueba	Valor	Interpretación
KMO (Global)	0.80	Indica correlaciones suficientemente fuertes.
KMO por variable	Motivacional: 0.88, Afectividad: 0.83, Instruccional: 0.78, Cognitiva: 0.77	Cada ítem contribuye adecuadamente al modelo factorial.
Prueba de Bartlett	$X^2 = 172.51, p < 0.001$	Rechaza la hipótesis nula, justificando la aplicación del análisis factorial.

Fuente: Elaboración propia.

Se realizó un AFE con método de máxima verosimilitud y rotación Varimax. Los resultados indicaron la presencia de un único factor subyacente que explica el 59% de la varianza.

Tabla 13.

Ítem	Carga factorial (ML1)	Varianza explicada (h ²)	Varianza no explicada (u ²)
Motivacional	0.61	0.37	0.63
Afectividad	0.76	0.58	0.42
Instruccional	0.82	0.67	0.33
Cognitiva	0.85	0.73	0.27

Fuente: Elaboración propia.

Las cargas factoriales, que oscilan entre 0.61 y 0.85, muestran que los ítems están relacionados con el factor latente identificado.



Se implementó un modelo de AFC con un factor latente denominado "Autopercepción del Aprendizaje", el cual se ajustó a los datos obtenidos de la EAA.

Tabla 14.

Índice	Valor	Criterios aceptados	Interpretación
CFI	1.000	> 0.95 = Excelente ajuste	El modelo ajusta perfectamente los datos.
TLI	1.024	> 0.95 = Excelente ajuste	Confirma la simplicidad y ajuste del modelo.
RMSEA	0.000	< 0.05 = Excelente ajuste	No hay discrepancia entre el modelo y los datos.
SRMR	0.011	< 0.08 = Buen ajuste	Los residuos entre las correlaciones predichas y observadas son mínimos.

Fuente: Elaboración propia.

Las cargas factoriales estandarizadas se presentan a continuación:

Tabla 15.

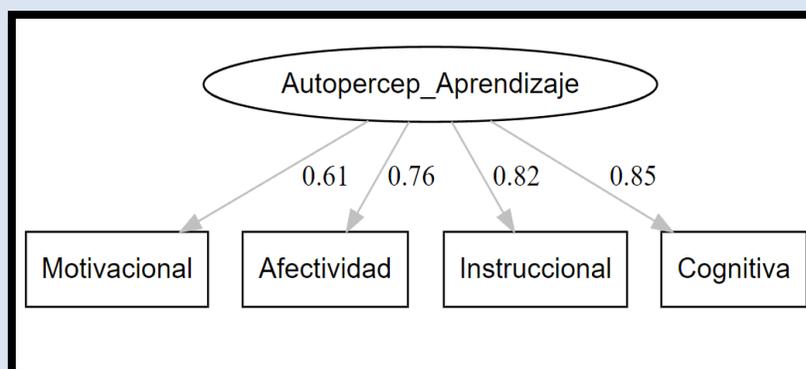
Ítem	Carga factorial estandarizada
Motivacional	0.607
Afectividad	0.759
Instrucciona	0.821
Cognitiva	0.855

Fuente: Elaboración propia.

Los resultados del AFE y AFC respaldan la estructura unidimensional de la EAA. Los indicadores de ajuste del modelo confirman que el factor "Autopercepción" explica de manera adecuada las relaciones entre los ítems.

A continuación, se presenta el modelo del factor Autopercepción del Aprendizaje evaluado a través de la EAA, considerando sus cargas factoriales calculadas en el AFC:

Figura 3.



Fuente: Elaboración propia. Generado en RStudio.

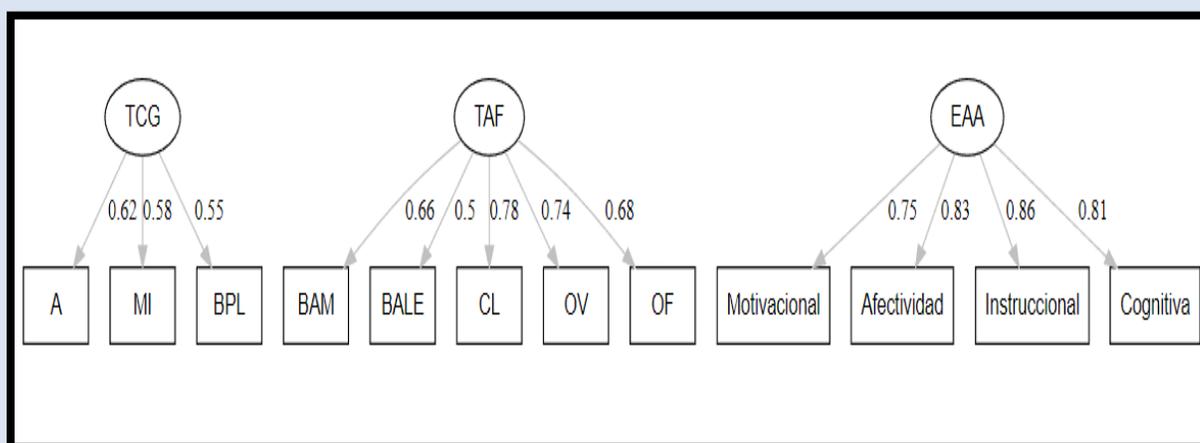
Análisis Factorial para la validez de la estructura tridimensional del TBA

Se realizó un AFE y AFC para evaluar la estructura dimensional de un conjunto de datos con 5321 observaciones, utilizando metodologías robustas para asegurar la validez psicométrica. El índice KMO global fue de 0.81, indicando una buena adecuación de la muestra. Los

valores individuales de adecuación muestral (MSA) oscilaron entre 0.63 y 0.88, lo que confirmó que los ítems son apropiados para el análisis factorial. La prueba de esfericidad de Bartlett fue significativa ($X^2 = 21444.503$, $p < 0.001$), indicando correlaciones suficientes entre los ítems.

En el AFE, se extrajeron tres factores mediante máxima verosimilitud y rotación Varimax. Las cargas factoriales estándar estuvieron en el rango de 0.50 a 0.86, lo que evidencia relaciones sólidas entre los ítems y los factores subyacentes. El índice SRMR fue de 0.01, sugiriendo un ajuste excelente. El AFC confirmó la estructura tridimensional propuesta, utilizando el modelo de ecuaciones estructurales. Los índices de ajuste globales fueron excelentes: RMSEA = 0.028 (IC 90%: 0.025-0.032) y CFI = 0.990, indicando un ajuste cercano al modelo ideal. Estos resultados son consistentes con una estructura factorial clara y estable.

Figura 4.



Fuente: Elaboración propia. Generado en RStudio.

Análisis de consistencia a través del Test Retest

La comparación entre los puntajes obtenidos en la primera aplicación del test (Test) y en la segunda aplicación (Retest) se presenta a continuación:

Tabla 16.

Estadístico	Test	Retest
Mínimo	-0.02566	-0.08
Primer cuartil	6.00000	6.00
Mediana	10.00000	10.50
Media	22.20522	22.80
Tercer cuartil	22.00000	21.00
Máximo	166.00000	152.00
Correlación (Pearson)	1.0000000	0.9920843
t-valor	51.201	-
Grados de libertad (df)	42	-
p-valor	< 2.2e-16	-
Intervalo de confianza 95%	[0.9854484, 0.9957006]	-
Correlación estimada	0.9920843	-

Fuente: Elaboración propia.

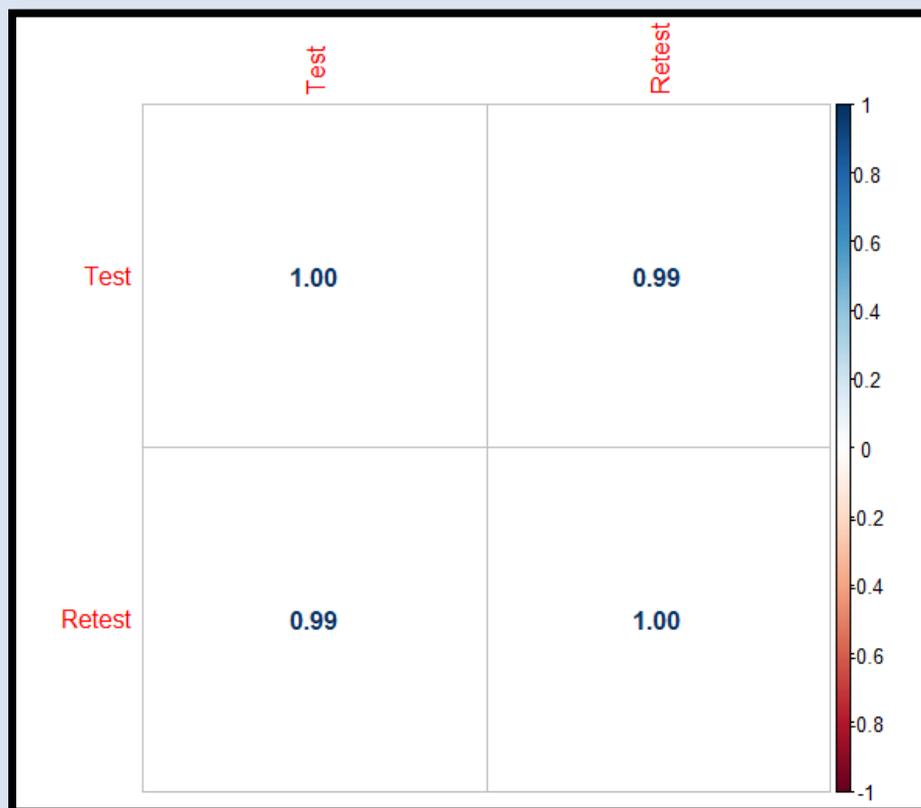


El análisis Test-Retest del TBA muestra una alta fiabilidad temporal en los resultados obtenidos. Los estadísticos descriptivos indican que, aunque existen algunas variaciones extremas en los puntajes (máximos de 166 y 152 para Test y Retest, respectivamente), la media y la mediana son consistentes entre ambas aplicaciones, sugiriendo una distribución relativamente estable de los datos.

La correlación de Pearson entre las dos aplicaciones es de 0.992, lo que refleja una correlación extremadamente alta y denota una fuerte consistencia entre los puntajes. Esta relación fue confirmada mediante una prueba de correlación de Pearson con un p-valor menor a $2.2e-16$, lo que indica que los resultados no son producto del azar. Además, el intervalo de confianza del 95% para la correlación (0.985 a 0.996) refuerza la confiabilidad del test a lo largo del tiempo, incluso con un intervalo de 2 semanas entre las aplicaciones.

El TBA demuestra ser un instrumento fiable para medir habilidades de aprendizaje, con un alto grado de estabilidad en los puntajes en un período corto de tiempo.

Figura 5.



Fuente: Elaboración propia. Generado en RStudio.

El gráfico refleja una relación fuerte de tipo directa, lo cual indica que los resultados del Retest crecen o disminuyen en función del crecimiento del Test. Además, se ajustó un modelo de regresión lineal para predecir Retest en función del Test. Los residuos del modelo oscilan entre -9.90 y 15.57, con un error estándar residual de 4.47.

El modelo presentó un intercepto de 1.79 ($p = 0.0287$) y un coeficiente para Test de 0.95 ($p < 0.001$), indicando una relación significativa entre las variables. El R cuadrado fue de 0.984, lo que sugiere que el modelo explica el 98.4% de la variabilidad de Retest. La

prueba F también resultó significativa ($F = 2621$, $p < 0.001$), confirmando la idoneidad del modelo.

Conclusión

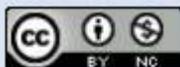
Los análisis matemáticos indican que el TBA es un instrumento estable e internamente consistente, dado que, en promedio, el alfa de Cronbach se encuentra entre 0.7 y 0.9, según intervalos de confianza del 95% (Celina y Campo, 2005). En cuanto a las variables que mide, el análisis factorial exploratorio permitió identificar factores que explican la estructura subyacente entre ellas (Méndez y Rondón, 2012), generando resultados coherentes con los modelos teóricos y posibilitando inferencias sobre la validez de los instrumentos y la tridimensionalidad de la batería. Los resultados del Test-Retest indican una predictibilidad del 98.4%, de acuerdo con el coeficiente de determinación (R^2).

El TBA cumple con todas las condiciones para constituirse como una batería de medición tridimensional capaz de obtener información relevante sobre el perfil de aprendizaje de las personas. Esto se sustenta en la ponderación del desarrollo cognitivo y las habilidades académicas y, además, en su función para estimar los cambios relativos de las variables ponderadas ($(T2-T1)/T1$), lo que permite calcular el potencial de aprendizaje por unidad de tiempo.

Dado que esta batería psicopedagógica no presenta un diseño curricular, reduce el sesgo en la medición de habilidades académicas. En términos cognitivos, al estar diseñada tanto para personas lectoras como no lectoras, los factores lingüísticos influyen en menor medida, lo que disminuye el sesgo socioeconómico típico de los instrumentos psicométricos. Esto contrasta con los instrumentos psicopedagógicos curriculares tradicionales, que, si bien proporcionan información relevante sobre los aprendizajes formales, presentan imprecisiones en la medición de habilidades, ya que evalúan a partir de conocimientos adquiridos.

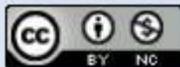
El TBA es una herramienta fundamental para los profesionales de las ciencias del aprendizaje, ya que permite estudiar perfiles de aprendizaje en distintos contextos, sea clínico, escolar, laboral, sociocomunitario u otro. Además, su informatización optimiza todos los procesos de análisis y mantiene los baremos actualizados con información poblacional creciente. Esto representa una diferencia significativa respecto a otros instrumentos de uso habitual, que suelen ser en formato lápiz y papel o físicos, y cuyas actualizaciones o nuevas versiones pueden tardar más de cinco años.

En cuanto a sus desventajas, el TBA no es un instrumento de diagnóstico automático, por lo que requiere de criterios profesionales para la inferencia diagnóstica de condiciones clínicas del neurodesarrollo, cognitivas y de aprendizaje. Respecto a los desafíos psicométricos, este instrumento se basa principalmente en la teoría clásica de la medición, por lo que su precisión depende directamente de la muestra poblacional utilizada para el cálculo de los niveles de desarrollo. No obstante, gracias a sus características tecnológicas, el TBA ofrece mayor factibilidad para avanzar hacia modelos psicométricos más complejos, lo que permitirá aumentar la precisión de los estudios cognitivos a partir de parámetros invariantes (Muñiz, 2010).



Referencias

- Bolívar López, J. M., & Rojas Velásquez, F. (2015). Estudio de la autopercepción y los estilos de aprendizaje como factores asociados al rendimiento académico en estudiantes universitarios. *Revista de Educación a Distancia (RED)*, 15(44). <https://revistas.um.es/red/article/view/237781>
- Celina Oviedo, H., & Campo Arias, A. (2005). Aproximación al uso del coeficiente alfa de Cronbach. *Revista Colombiana de Psiquiatría*, 34(4), 572-580. <https://www.redalyc.org/pdf/806/80634409.pdf>
- Chamorro, M. (2003). *La didáctica de la matemática para primaria*. Síntesis.
- Cuetos, F. (1996). *Psicología de la lectura*. Editorial Escuela Española.
- Cuetos, F. (2012). *Neurociencia del lenguaje: bases neurológicas e implicaciones clínicas*. Editorial Médica Panamericana.
- Fry, E. (1975). *Reading drills for speed and comprehension*. Jamestown Publishers.
- Garello, M. V. (2008). Autopercepción y aprendizaje autorregulado. En *XV Jornadas de Investigación y Cuarto Encuentro de Investigadores en Psicología del Mercosur*. Facultad de Psicología - Universidad de Buenos Aires. <https://www.aacademica.org/000-032/309>
- Garrido Bermúdez, E., Mena Rodríguez, H. Y., Zuluaga Arango, J. M., & Pérez Quintero, F. E. (2023). Proceso para validar un instrumento de investigación por medio de un análisis factorial. *UNACIENCIA*, 16(30), 61-73. <https://doi.org/10.35997/unaciencia.v16i30.724>
- International Test Commission & Association of Test Publishers. (2022). *Guidelines for technology-based assessment*. <https://www.intestcom.org/page/28>
- Maureira, C. F. (2017). *¿Qué es la inteligencia?* Bubok Publishing S.L.
- Méndez Martínez, C., & Rondón Sepúlveda, M. A. (2012). Introducción al análisis factorial exploratorio. *Revista Colombiana de Psiquiatría*, 41(1), 197-207. <https://www.redalyc.org/pdf/806/80624093014.pdf>
- Miller, G. A. (1956). El mágico número siete, más o menos dos: Algunos límites en nuestra capacidad para procesar la información. *Psychological Review*, 63(2), 81-97. <https://psychclassics.yorku.ca/Miller/>
- Morales, M. L. (1990). *Psicometría aplicada*. Trillas.
- Muñiz Fernández, J. (2010). Las teorías de los tests: Teoría clásica y teoría de respuesta a los ítems. *Papeles del Psicólogo: Revista del Colegio Oficial de Psicólogos*. <https://www.redalyc.org/articulo.oa?id=77812441006>
- Orozco Santa María, A. M., & García Ramírez, M. T. (2017). Autopercepción de habilidades de aprendizaje en ambientes virtuales. *CPU-e, Revista de Investigación Educativa*, 25, 144-167. <https://www.redalyc.org/pdf/2831/283152311007.pdf>
- Pérez, E. R., & Medrano, L. A. (2010). Análisis factorial exploratorio: Bases conceptuales y metodológicas. *Revista Argentina de Ciencias del Comportamiento (RACC)*, 2(1), 58-66. <https://dialnet.unirioja.es/servlet/articulo?codigo=3161108>



- San Martín Catalán, P. J., & Carrasco Cursach, J. F. (2022). Modelo psicopedagógico para la dinámica cognitiva del aprendizaje humano. *Revista Neuronum*, 8(1), 50-74. <https://dialnet.unirioja.es/servlet/articulo?codigo=9690956>
- Sousa, V. D., Driessnack, M., & Mendes, I. A. C. (2007). An overview of research designs relevant to nursing: Part 1: quantitative research designs. *Revista Latino-Americana De Enfermagem*, 15(3), 502-507. <https://doi.org/10.1590/S0104-11692007000300022>

