

Análisis psicométrico de la BEVTA en niños con problemas de aprendizaje

Denisse Lisette Manrique Millones

El presente estudio tuvo como objetivo principal el análisis psicométrico de la Batería de Evaluación Verbal para Trastornos de Aprendizaje (BEVTA), instrumento de evaluación psicopedagógica con una clara formulación teórica, diseñada y normada en Chile. Se trabajó con 447 participantes de 7 a 12 años, de ambos sexos, que habían sido evaluados con dicho instrumento en los años 1996 al 2004. La metodología utilizada fue de tipo descriptivo-psicométrico y el diseño de investigación fue el descriptivo comparativo. A partir de los resultados obtenidos se concluye que los cuatro subtest del instrumento analizado no presentan una elevada confiabilidad ni validez de constructo; sin embargo, sí se reporta validez relacionada con el criterio.

Validez / confiabilidad / problemas de aprendizaje

Psychometric analysis of the BEVTA in children with learning problems

The aim of the current study was to formulate a psychometric analysis of the Battery of Verbal Evaluation for Learning Disabilities (BEVTA), a psychopedagogical evaluation instrument with a plain framework designed in Chile. Participants were 447, with ages ranged from 7 to 12 years old, both male and female those were evaluated with the present battery from 1996 to 2004. The method was descriptive-psychometric with a comparative-descriptive design. The results showed that the four subtest of the battery neither have a high reliability nor construct validity but, on the other hand, presents validity related to the criterion.

Validity / reliability / learning disabilities

El avance de la investigación en psicometría confirma día a día que las diferencias individuales existen y son medibles. Las bases de la aplicación de pruebas se encuentran en la antigüedad, desde los exámenes que se hacían en China para el servicio civil, hasta la aplicación de exámenes en Grecia, para evaluar las habilidades físicas e intelectuales (Calleja, Nezira & Gómez-Perezmitré, 2001).

Es importante mencionar que la psicometría es la ciencia cuyo objeto es medir los aspectos psicológicos de una persona, como son: el conocimiento, las habilidades, las capacidades, las actitudes o la personalidad, asimismo se encarga de los procesos para la medida de los aspectos “psicológicos” de una persona, tales como conocimiento, habilidades, capacidades o personalidad. Las primeras pruebas psicométricas fueron construidas para medir la inteligencia y sus bases teóricas parten de la teoría de los tests y los métodos de elaboración de escalas (Magnusson, 1987).

Actualmente se debe analizar la problemática de los tests psicológicos, sus límites y alcances. En relación con ello, la falta de conocimientos acerca de la teoría y la técnica de las pruebas psicológicas aplicadas para evaluar los diversos aspectos del comportamiento humano, ha llevado a intentar soluciones inadecuadas en esta problemática. Entre ellas el uso de test elaborados en otros países, con patrones culturales

muy diferentes, sin la debida adaptación a nuestro contexto, así como también la aplicación indiscriminada de instrumentos que carecen de fundamentación científica.

El lenguaje es de suma importancia debido a que interviene en procesos básicos, como cognición, memoria, atención, etcétera, y con otras ramas del saber, como la neurofisiología –ya que esta explica ciertos trastornos del lenguaje sobre la base del sustrato fisiológico–; asimismo, la educación se ve estrechamente vinculada, puesto que esta desde siempre ha utilizado métodos de instrucción basados en el componente lingüístico.

Las características progresivas del desarrollo del lenguaje verbal en los diferentes niveles de edad, se adscriben a las etapas del desarrollo integral del niño, encontrándose estrechamente asociado a los siguientes aspectos (Sarafino, 1988):

- Al proceso de maduración del sistema nervioso, tanto al central (SNC) como al periférico, correlacionándose sus cambios progresivos con el desarrollo motor en general y con el aparato fonador en particular.
- Al desarrollo cognoscitivo que comprende desde la discriminación perceptual del lenguaje hablado hasta la función de los procesos de simbolización y el pensamiento.
- Al desarrollo socioemocional, que es el resultado de la influencia del medio sociocultural, de las interac-

ciones del niño y las influencias recíprocas.

En ese sentido, el modelo escolar de transmisión y evaluación de conocimientos, en nuestra cultura, es esencialmente verbal. La mayor parte de la información que se le brinda al niño en el salón de clases es expresada oralmente por los profesores, o por escrito mediante los textos.

LOS PROBLEMAS DE APRENDIZAJE

El problema del aprendizaje es un término general que describe problemas del aprendizaje específico. Un problema del aprendizaje puede causar que una persona tenga dificultades para aprender y usar ciertas destrezas. Las destrezas afectadas con mayor frecuencia son: leer, escribir, escuchar, hablar, razonar y desarrollar ejercicios de matemática. Los problemas de aprendizaje se hacen evidentes en los primeros años del periodo escolar, pues están directamente relacionados con materias a partir de las cuales se determina el correcto rendimiento académico.

Bravo (1980) ha clasificado las dificultades de aprendizaje en problemas generales de aprendizaje y trastornos específicos de aprendizaje.

PROBLEMAS GENERALES DE APRENDIZAJE

Se manifiesta un retardo general de todo el proceso de aprendizaje, observándose lentitud, desinterés, deficiencia en la

atención y concentración, que afectan el rendimiento global. Estas características se presentan en niños con un desarrollo normal y con inmadurez en el área cognitiva o verbal, lo que provocaría una lentitud para aprender.

TRASTORNO ESPECÍFICO DE APRENDIZAJE

Se manifiesta en niños con inteligencia normal o alrededor de lo normal, que carecen de alteraciones sensoriales o emocionales severas. Su ambiente sociocultural y educacional es satisfactorio. No logran un rendimiento escolar normal y presentan dificultades reiteradas en ciertas áreas del aprendizaje, funcionando bien en algunas y mal en otras. Estas dificultades dependen de alteraciones en el desarrollo y en la maduración psíquica y neurológica.

CONFIABILIDAD

Todo instrumento psicométrico se considera de valor cuando posee una medida consistente o *confiable*. Por ello, uno de los aspectos primordiales que será necesario determinar sobre la elaboración de una prueba es si resulta lo suficientemente confiable para medir aquello para lo que fue diseñada.

Hablando en términos estrictos, más que ser característica de una prueba la confiabilidad es una propiedad de las puntuaciones obtenidas cuando se administra la prueba a un grupo en particular de personas, en una ocasión particular y bajo condiciones específicas” (Thompson, 1994).

TEORÍA CLÁSICA DE LOS TEST

Se denomina Teoría Clásica de los Test (TCT) al conjunto de principios teóricos y métodos cuantitativos derivados de ellos, que fundamentan la construcción, aplicación, validación e interpretación de distintos tipos de test y que permiten derivar escalas estandarizadas aplicables a una población (Hambleton, 1994). Los principios en que se basa son relativamente simples y se aplican tanto a las pruebas de desempeño como a las de aptitud. Durante sus diferentes fases de desarrollo, se han elaborado procedimientos de análisis cuantitativo que han sido de gran utilidad, destacándose, en lo general, tres grandes etapas que se identifican por su objeto de interés primordial, así como por los métodos cuantitativos y tipos de análisis teóricos que utilizan.

VALIDEZ

Existe un sistema tradicional tripartito para clasificar los diferentes tipos de evidencia de la validez, que se utilizó en los estándares de 1985 y está muy arraigado en los estudios de la psicometría. No obstante, los estándares de 1999 abandonan en parte este sistema con el objetivo de presentar una representación más diversificada de los tipos de evidencia de la validez; es muy probable que el sistema tradicional no desaparezca pronto (Hogan, 2004).

La validez de contenido en gran medida tiene el mismo significado en ambos. La validez relacionada con el criterio del sistema tradicional corresponde a la categoría más reciente de “relaciones con otras variables”, sobre todo con la subcategoría de “relaciones entre prueba y criterio”. La distinción entre la validez concurrente y predictiva se establece tanto en los estándares de 1985 como en los de 1999, pero sin hacer subcategorías distintas. Aquí se presentan como subcategorías separadas, en virtud de que esta distinción se hace con mucha frecuencia en los manuales de las pruebas, los artículos de las publicaciones periódicas y en documentos similares.

OBJETIVOS DE LA INVESTIGACIÓN

Objetivo general

Analizar las propiedades psicométricas de la BEVTA, aplicado en un colegio especializado en problemas de aprendizaje.

Objetivos específicos

- Determinar y establecer las propiedades psicométricas del Test de Asimilación Verbal Inmediata (TAVI); del Test de Semejanzas Verbales (3-S); de la Prueba de Nominación de Conceptos Agrupados por Categorías (CAT-V) y de la Prueba de Seria-ciones Verbales (S-V) en el diagnóstico de problemas

de aprendizaje.

- Analizar la validez de constructo de las áreas de la BEVTA en el diagnóstico de problemas de aprendizaje.
- Determinar la validez relacionada con el criterio, a través de la validez predictiva, con base en la diferenciación del grupo conocido, de acuerdo con el grado de estudio de la Batería de Evaluación Verbal de Trastornos de Aprendizaje en el diagnóstico de problemas de aprendizaje.
- Determinar la confiabilidad de la Batería de Evaluación Verbal de Trastornos de Aprendizaje en el diagnóstico de problemas de aprendizaje.

METODOLOGÍA DE INVESTIGACIÓN

Diseño

El diseño de investigación utilizado es

el descriptivo comparativo. Este diseño parte de la consideración de dos o más investigaciones descriptivas simples (Sánchez, 2006). Esto es, recolectar información relevante en varias muestras con respecto a un mismo fenómeno y luego caracterizar este fenómeno sobre la base de la comparación de los datos recogidos, en la medida en que el análisis que se haga sea por sexo, grado y edad.

Participantes

Se consideró como participantes a todos los alumnos que habían sido examinados con la Batería de Evaluación Verbal de Trastornos de Aprendizaje en los años 1996-2004 en un colegio especializado en problemas de aprendizaje de Lima metropolitana.

Por la peculiaridad de la prueba los participantes fueron divididos en función de las edades (véase especificacio-

Tabla 1
Composición de los participantes por área evaluada

Subtest	Edades a evaluar	Participantes
Test de Semejanzas Verbales (3-S)	7 a 10 años	298
Test de Nominación de Conceptos	7 a 12 años	447
Agrupados por Categorías (CAT-V)	8 a 12 años	411
Test de Seriaciones Verbales (S-V)	8 a 12 años	411

nes en la tabla 1).

Instrumentos

El objetivo de la Batería de Exploración Verbal de Trastornos de Aprendizaje es establecer el funcionamiento de algunas áreas de rendimiento psicolingüístico que son relevantes en el rendimiento escolar básico.

Está compuesta por cuatro pruebas que permiten al investigador analizar áreas de rendimiento verbal. Tres fueron diseñadas por Luis Bravo Valdivieso y Arturo Pinto Guevara (1995): el TAVI, el S-V y el CAT-V, y el 3-S fue construido según el modelo de una prueba verbal de Sadek-Khalil (1968).

Esta batería se aplica a niños de 7 a 12 años de edad y abarca desde el segundo hasta el sexto grado de primaria. Asimismo, se aplica de manera individual y oral.

Procedimiento

Todos los datos fueron recogidos de los archivos del colegio. Cada alumno cuenta con un registro de su historia personal. Para realizar la presente investigación se revisaron dichas historias y se tomaron los aspectos considerados como necesarios para este trabajo. Se elaboró una ficha acumulativa para incluir los datos de cada niño, que consideró:

- Edad: de 7 a 12 años.
- Grado de escolaridad: de segundo a

sexto de primaria.

- Sexo: Femenino o masculino.
- Diagnósticos neurológicos: Estos datos fueron obtenidos de los antecedentes de cada niño. Básicamente, los diagnósticos que se repitieron en mayor número fueron: déficit de atención con hiperactividad, desorden de desarrollo del lenguaje expresivo, desorden específico del desarrollo motor, desorden del desarrollo de articulación verbal.
- Diagnósticos psicológicos, coeficiente intelectual total, verbal y de ejecución, obtenidos a través de las pruebas Wechsler, dependiendo de la edad de cada niño.
- Se utilizó la Batería de Exploración Verbal de Trastornos de Aprendizaje.

Posteriormente, para poder realizar los procedimientos necesarios, se trasladó dicha información a una base de datos. Esto permitió realizar los procedimientos estadísticos requeridos para el presente trabajo.

RESULTADOS

Confiabilidad de las áreas del BEVTA

Los criterios utilizados para evaluar la prueba fueron: KR-20 debe ser mayor de 0.80 (Kaplan & Saccuzzo, 2006).

Respecto del grado de confiabilidad en el Test de Asimilación Verbal Inmediata (TAVI), se obtiene un KR-20 de 0.56, que puede clasificarse como bajo,

pero que para fines del estudio no supera el criterio de 0.80 planteado por Kaplan y Sacuzzo (2006); por ello, se concluye que el instrumento no presenta confiabilidad por consistencia interna.

En cuanto a la confiabilidad en el test de Semejanzas Verbales (3-S), se obtiene un KR-20 de 0,68, que puede clasificarse como moderado. Estos resultados nos indican que la prueba alcanza una consistencia interna que no es aceptable para este test.

En el test de nominación de conceptos agrupados por categorías (CAT-V), se obtiene un KR-20 de 0.77, que puede clasificarse como aceptable. Estos resultados nos indican que se alcanza una inadecuada consistencia interna para este test.

Respecto del grado de confiabilidad en el Test de Seriaciones Verbales (S-V), se obtiene un KR-20 de 0.59, que puede clasificarse como bajo. Estos resultados indican que no se logra alcan-

zar una consistencia interna aceptable para este test.

Análisis de la validez de las áreas de la BEVTA

Al analizar las cuatro pruebas de la Batería se obtuvo como resultado lo siguiente:

En la tabla 2 se observa que existen correlaciones positivas entre los totales Tres S y CAT-V ($R=.61p<.01$) y los totales Tres S y S-V ($R=.43 p<.01$), y, finalmente, entre los totales CAT-V y S-V ($R=.55 p<.01$).

Además, se encuentran correlaciones inversas y bajas entre los totales TAVI y CAT-V ($-.13 p<.05$) y los totales TAVI y SV ($R= -.16 p< .01$).

Se concluye que los resultados no presentan validez de constructo ya que se observa asociación moderada y baja entre ellos.

Tabla 2

Validez de constructo a través de la matriz de correlaciones entre las áreas del BEVTA

Variable	Total TAVI	Total TRES-S	Total CAT-V
Total Tres S	-.07		
N	445		
Total CAT-V	-.13*	.61 **	
N	411	408	
Total S-V	-.16 **	.43 **	.55 **
N	411	408	411

* $p < .05$ ** $p < .01$

Los resultados en la tabla 3, sobre la matriz de correlaciones entre las áreas del BEVTA y el CI es que existen correlaciones positivas y bajas entre el total TAVI y CIE ($R = .21$ $p < .01$); asimismo, entre los totales TAVI y CI Total ($R = .24$ $p < .01$). Luego existen correlaciones positivas y bajas entre el los totales Tres S y CIV ($R = .26$ $p < .01$) y Tres S y CI Total ($R = .28$ $p < .01$).

Finalmente, existen correlaciones positivas y bajas entre los totales S-V y CIV ($R = .28$ $p < .01$); S-V y CIE ($R = .24$ $p < .01$) y S-V y CI TOTAL ($R = .28$ $p < .01$).

ANÁLISIS DE LA VALIDEZ PREDICTIVA

SOBRE LA BASE DEL MÉTODO DE LA DIFERENCIACIÓN DE GRUPOS

Respecto al análisis de la validez predictiva utilizando el método de la diferenciación de grupos, la comparación del TAVI por grado (véase tabla 4) se encuentra que existen diferencias significativas entre las edades evaluadas ($X^2(3) = 70.46$ $p < .001$), notándose además que en las comparaciones por pares, el grupo de segundo grado ($Mr = 87.67$) presenta valores menores que el de tercer grado ($Mr = 149.81$), cuarto grado ($Mr = 165.74$) y quinto grado ($Mr = 210.42$). También se encuentra que las puntuaciones se presentan en

Tabla 3
Matriz de correlación entre las áreas del BEVTA y el Coeficiente Intelectual (CI)

Variable	Total TAVI	Total TRES-S	Total CAT-V	Total S-V
CIV	.19 *	.26 **	.17	.28 **
N	160	158	125	125
CIE	.21 **	.19 *	.12	.24 **
N	160	158	125	125
CI Total	.24 **	.28 **	.18 *	.28 **
N	159	157	125	125

* $p < .05$ ** $p < .01$

Tabla 4
Análisis de la validez predictiva del TAVI a través de la comparación por grado

	2. ^{do} grado N = 76	3. ^{er} grado N = 80	4. ^{to} grado N = 89	5. ^{to} grado N = 53	X^2
	Media de rangos	Media de rangos	Media de rangos	Media de rangos	
Total TAVI	87.67abc	149.81a	165.74b	210.42c	70.46***

*** $p < .001$

Diferencias por pares a través de la prueba U de Mann-Whitney: a, b, c.

Tabla 5
Análisis de la validez predictiva del Tres S a través de la comparación por grado

	2. ^{do} grado N = 76	3. ^{er} grado N = 79	4. ^{to} grado N = 94	5. ^{to} grado N = 101	6. ^{to} grado N = 101	χ^2
	Media de rangos	Media de rangos	Media de rangos	Media de rangos	Media de rangos	
Total TRES-S	107.06abcd	180.90a	228.80b	270.76c	295.83d	115.08 ***

*** $p < .001$

Diferencias por pares a través de la prueba U de Mann-Whitney: a, b, c, d.

Tabla 6
Análisis de la validez predictiva del CAT-V a través de la comparación por grado

	3. ^{do} grado N = 79	4. ^{er} grado N = 94	5. ^{to} grado N = 101	6. ^{to} grado N = 98	χ^2
	Media de rangos	Media de rangos	Media de rangos	Media de rangos	
Total CAT-V	124.03	189.85	177.31	243.12	54.77***

*** $p < .001$

Diferencias por pares a través de la prueba U de Mann-Whitney: a, b, c.

forma ascendente por grado.

En conclusión, estos hallazgos indican el TAVI que presenta validez predictiva.

Respecto de la comparación del Tres S por grado (véase tabla 5), se encuentra que existen diferencias significativas entre las edades evaluadas ($X^2(3) = 115.080$ $p < .001$), notándose además que en las comparaciones por pares el grupo de segundo grado ($Mr = 107.06$) presenta valores menores que el de tercer grado ($Mr = 180.90$), cuar-

to grado ($Mr = 228.80$), quinto grado ($Mr = 270.76$) y sexto grado ($Mr = 295.83$). También se encuentra que las puntuaciones se presentan en forma ascendente por grado.

Respecto de la comparación del CAT-V por grado (véase tabla 6) se encuentra que existen diferencias significativas entre las edades evaluadas ($X^2(3) = 54.77$ $p < .001$), notándose además que en las comparaciones por pares el grupo de tercer grado ($Mr = 124.03$) presenta valores menores que los del cuarto grado ($Mr = 189.85$), los

Tabla 7
Análisis de la validez predictiva del S-V a través de la comparación por grado

	3. ^{do} grado N = 79	4. ^{er} grado N = 94	5. ^{to} grado N = 101	6. ^{to} grado N = 98	X ²
	Media de rangos	Media de rangos	Media de rangos	Media de rangos	
Total S-V	145.37abc	161.97a	188.29b	241.33c	42.591***

*** p < .001

Diferencias por pares a través de la prueba U de Mann-Whitney: a, b, c.

de quinto grado (Mr = 177.31) y los de sexto grado (Mr = 243.12). También se encuentra que las puntuaciones se presentan en forma ascendente por grado.

Respecto de la comparación del S-V por grado (véase tabla 7) se encuentra que existen diferencias significativas entre las edades evaluadas ($X^2(3) = 42.591$ p < .001), notándose además que en las comparaciones por pares el grupo de tercer grado (Mr = 145.37) presenta valores menores que los del cuarto grado (Mr = 161.97), los de quinto grado (Mr = 188.29) y los de sexto grado (Mr = 241.33). También se encuentra que las puntuaciones se presentan en forma ascendente por grado.

DISCUSIÓN

Estos hallazgos permiten afirmar que la batería de exploración verbal de trastornos de aprendizaje no posee índices aceptables de confiabilidad en cada uno de los test, pues, como sugiere Kaplan, estos deben ser mayores de 0.80.

Las razones por las cuales el TAVI y el S-V obtuvieron valores bajos podría deberse a que al trabajar con niños que presentan trastornos de aprendizaje, la muestra resulta muy homogénea, que es precisamente un elemento básico para la confiabilidad (Brown, 1986).

Asimismo, existirían fallas en la retención verbal inmediata, en la comprensión oral de oraciones de estructura simple y en la atención a estímulos verbales, incluso habría un déficit de verbalización de una respuesta. Asimismo, en el caso del Test de Seriaciones Verbales es probable que se vean afectadas las áreas de recepción y comprensión verbal, retención verbal de series con significado, percepción secuencial oral y atención verbal.

Esto constituye un elemento referencial para entender pedagógicamente el problema de la comprensión de lectura; asimismo, es posible que no se estén desarrollando estas capacidades y habilidades, necesarias para un adecuado manejo de la comprensión lectora,

en una temprana edad. Esta problemática se ha convertido actualmente en un inconveniente, debido a la estructura contextual en general, ya que está presente en todos los estratos socioeconómicos y en diferentes niveles educativos: primaria, secundaria y superior, repercutiendo directamente sobre ella (González, 1998).

Tomando como punto de partida el estudio de González (1998), quien investigó sobre estos aspectos, se evidenció que la distribución de frecuencias y por ende la categorización lectora de los estudiantes universitarios iniciales era tendencialmente muy baja. Asimismo, la cantidad de lectores deficientes con perspectivas pedagógicas no muy buenas era muy alta, emergiendo como una posibilidad real que la educación secundaria del país no logre ninguno de sus objetivos académicos, debiendo la universidad asumir la responsabilidad y trabajar en tareas de comprensión lectora, poniendo mucho énfasis en el primer ciclo universitario, como primer encuentro con la educación impartida por la escuela.

Con las distintas reformas educativas que se llevan a cabo a escala mundial, en los últimos años surge la inquietud sobre la falta de preparación académica de alumnos universitarios. Los maestros comentan sobre el gran número de alumnos carentes de hábitos de estudio; sin embargo, pareciera que no les falta la capacidad intelectual para el estudio, sino que no pueden en-

frentar las demandas del discurso académico ni del ritmo de trabajo (Azcoaga et al., 1979).

Asimismo, los resultados obtenidos en el Programa Internacional de Evaluación de Estudiantes (Programme for International Student Assessment, PISA), que se llevó a cabo en el 2004, teniendo como participantes a estudiantes de los países de la Organización para la Cooperación y el Desarrollo Económico (OCDE) indicaron que mientras el 10% de los estudiantes de países de la OCDE es calificado, a partir de los estándares de PISA, como un valioso capital humano que potencialmente podría liderar el desarrollo económico de sus países en una economía global. En la región latinoamericana es casi nulo el porcentaje de estudiantes que muestra este perfil. En esta región destaca el desempeño de México y Argentina, que presentan un 10% y 7% de estudiantes, respectivamente, que alcanzan como mínimo el nivel 4, que es el segundo nivel de dominio más alto definido por PISA, es decir, son capaces de responder preguntas complejas de la prueba de lectura que implican ubicar información incrustada y evaluar críticamente un texto.

En el caso del Perú, este porcentaje se reduce a menos de 1%, lo cual expresa una situación crítica de ausencia de estudiantes en nuestro sistema con posibilidades de constituirse en líderes dentro de un mundo globalizado y en la sociedad del conocimiento actual. Los

porcentajes tan bajos de estudiantes en los niveles de dominio superiores en la región, y de manera especial en el Perú, ponen de manifiesto que en nuestros países prácticamente no existe una élite intelectual consolidada capaz de insertarse en puestos de liderazgo internacional desde una economía global o que, en todo caso, esta es muy reducida y pertenece, casi en su totalidad, solo a los grupos socioeconómicos más favorecidos de la región.

De otro lado, alrededor del 80% de los estudiantes peruanos se desempeña por debajo del nivel 3; es decir, la gran mayoría de estudiantes peruanos de 15 años se encuentra ubicado en los niveles de dominio más bajos establecidos por PISA (Unidad de Medición de la Calidad Educativa, 2004).

En lo que concierne a la versión analizada de la BEVTA se aprecia que presenta un inadecuado funcionamiento psicométrico traducido en términos de validez de constructo, aunque respecto de la validez de criterio sí obtuvo un adecuado nivel de funcionamiento.

Estos resultados permiten afirmar que la prueba no presenta confiabilidad y que esta no cumple con los parámetros técnicos establecidos que responden a la teoría clásica de los test, que fundamenta la construcción, aplicación, validación e interpretación de distintos tipos de test y que permiten derivar escalas estandarizadas aplicables a una población, como lo propone Ham-

bleton (1994).

Los resultados obtenidos permiten aseverar que la BEVTA no presenta un adecuado desempeño psicométrico, en un contexto específico, tal como lo proponen Anastasi y Urbina (1998), Muñiz (1996) y Aiken (2003), y que en este caso corresponde a un colegio especializado en problemas de aprendizaje.

Para el análisis de la validez relacionada con el criterio se utilizó la validez predictiva sobre la base del método de la diferencia de grupos conocidos; para ello se tomó en cuenta el grado al que pertenecían los alumnos. Los resultados indican que en las comparaciones según grado educativo existen diferencias significativas. Se observa que las puntuaciones se presentan en forma ascendente por grado, apreciándose que ello ocurre con todos los subtest del BEVTA, lo cual indica que presentan indicadores de validez predictiva (Brown, 1986; Anastasi & Urbina, 1998).

En las escuelas de los países en desarrollo el nivel primario se establece, en teoría, en la forma de programas unitarios de seis a doce años como preparación para un examen final que determina la educación posterior, o como programas básicos que capacitan para una escuela intermedia, que a su vez está orientada hacia el examen final para el grado correspondiente.

De hecho, la apariencia de continui-

dad, progreso y el plazo fijo de años de estudio es engañoso. Usualmente, el primer año de la escuela resulta ser un fracaso tremendo para la mitad o más de los alumnos (Bowles, 1974).

Los que superan la primera selección pueden pasar todas las demás pruebas o fracasar y repetir cada prueba (como lo hacen las dos terceras partes de ellos). El siguiente punto de la selección, o sea el examen comprensivo, por lo general llega al terminar el sexto año.

Como resultado, es difícil para un alumno con problemas de aprendizaje aprender al ritmo prescrito. Quizás un estudiante de cada diez que entra a la escuela primaria pasará el examen final y tendrá la oportunidad de ingresar al siguiente nivel educativo.

Por ello, actualmente existen colegios especializados que brindan educación individualizada para proporcionar una alternativa de enseñanza a niños con estas dificultades en el aprendizaje; asimismo, a pesar de que estos niños presentan insuficiencias en los procesos de recepción, comprensión, retención, expresión y creatividad en el aprendizaje escolar (Bravo, 1999), poseen un nivel intelectual general normal que los coloca en un nivel ascendente, a pesar de estas adversidades.

REFERENCIAS

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC.: American Psychological Association.
- Aiken, L. (2003). *Tests psicológicos y evaluación*. 11.^a edición. México: Pearson Educación.
- Anastasi, A. & Urbina, S. (1998). *Tests psicológicos*. México: Prentice-Hall/Hispanoamericana, S.A.
- Azcoaga, J. E.; Derman, B. & Iglesias, P. A. (1979). *Alteraciones del aprendizaje escolar. Diagnóstico, fisiopatología y tratamiento*. Buenos Aires: Paidós.
- Bowles, S. & Gintis, H. (1983). La educación como escenario de las contradicciones en la reproducción de la relación capital-trabajo. *Revista Educación y Sociedad*, 2. Madrid, pp. 25-38.
- Bracken, B. A. & Barona, A. (1991). State of the art procedures for translating, validating and using psychoeducational tests in cross-cultural assessment. *School Psychology International*, 12, 119-132.
- Bravo, L. (1980). *Teorías sobre la dislexia y su enfoque científico*. Santiago: Universitaria.
- Bravo, L. (1994). *Dislexias y procesamiento cognitivo*. Santiago: Universitaria.

- Bravo, L. (1999). Lenguaje y dislexias. Enfoque cognitivo del retardo lector. 3.^a edición. México: Alfaomega.
- Bravo, L. & Pinto, A. (1995). *Pruebas psicopedagógicas de lenguaje y lectura*. BEVTA. Santiago: Pontificia Universidad Católica de Chile, Facultad de Educación.
- Brown, F. (1986). *Principios de la medición en psicología y educación*. México: El Manual Moderno.
- Calleja, Nazira & Gómez-Perezmitré, Gilda (2001). *Psicología social. Investigación y aplicaciones en México*. México: Fondo de Cultura Económica. Biblioteca de Psicología, Psiquiatría y Psicoanálisis.
- González, R. (1998). Comprensión lectora en estudiantes universitarios iniciales. *Persona, 1*. Lima: Universidad de Lima, pp. 43-65.
- Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment, 10*, 3, 229-244.
- Hambleton, R. K. & Kanjee, H. J. (1995). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hogan, T. (2004). *Pruebas psicológicas. Una introducción práctica*. México: El Manual Moderno.
- Kaplan, R. M. & Saccuzzo, D. P. (1993). *Medición psicológica*. México: Thomson.
- Kaplan, R. M. & Saccuzzo, D. P. (2006). *Pruebas psicológicas*. 6.^a edición. México: Thomson.
- Magnusson, D. (1987). *Teoría de los tests*. México: Editorial Trillas.
- Ministerio de Educación (2002). Programa Internacional para la Evaluación de Estudiantes de la OCDE (PISA) y la participación del Perú. *Boletín de la Unidad de Medición de la Calidad, 21*. Lima: Ministerio de Educación, Unidad de Medición de la Calidad.
- Ministerio de Educación (2004). “Una aproximación a la alfabetización lectora de los estudiantes peruanos de 15 años”. Versión preliminar. Lima: Ministerio de Educación, Unidad de Medición de la Calidad.
- Muñiz, J. (1996). *Psicometría*. Madrid: Universitas.
- Sánchez, H. (2006). *Metodología y diseños en la investigación científica*. Lima: Visión Universitaria.
- Sánchez, J.; Jonquera, L.; Muñoz, E. & Valenzuela, E. (2001). *Estimulación de estructuras espacio-temporales en niños ciegos mediante la interacción con ambientes virtuales basados en sonido especializado* [en línea]. <<http://lsm.dei.uc.pt/ribie/docfiles/txt2003729182343paper-333.pdf>>.
- Sarafino, E. (1988). *Desarrollo del ni-*