

Utilidad de las pruebas de aprovechamiento en español y matemáticas para la  
identificación de niños sobresalientes

Rubí Medrano Chan y Pedro Sánchez Escobedo

Facultad de Educación, Universidad Autónoma de Yucatán

## Resumen

Se diseñaron seis pruebas de conocimientos de Español y Matemáticas en nivel primaria, así como la determinación de la validez, confiabilidad y del uso como herramienta en la identificación de niños sobresalientes.

Se trabajó con 341 estudiantes de alto rendimiento de tercero, cuarto y quinto grado de primaria de escuelas públicas y privadas de la ciudad de Mérida. Para cada grado se diseñó una prueba de Español y una de Matemáticas. Las pruebas fueron elaboradas de acuerdo al currículo oficial y en cada una se utilizó contenidos del grado actual y del grado superior inmediato.

Cada una de las pruebas de Español estuvo conformada por 20 reactivos, 10 correspondientes al contenido del grado actual y 10 correspondientes al contenido del grado superior inmediato. Cada una de las pruebas de Matemáticas estuvo formada por 24 reactivos, 12 correspondientes al contenido del grado actual y 12 al contenido del grado superior inmediato.

Los resultados muestran que los reactivos de las pruebas requieren ser revisados para poder ser utilizados como herramienta de apoyo en la identificación de niños sobresalientes. Los resultados también muestran puntajes inferiores en las pruebas de Matemáticas, por lo que es necesario mayor trabajo futuro para obtener mejores niveles de confiabilidad y puntaje en las pruebas.

## Utilidad de las pruebas de aprovechamiento en español y matemáticas para la identificación de niños sobresalientes

Todos los campos del esfuerzo humano usan la medición en alguna forma, y cada uno tiene su propio conjunto de herramientas de medición y unidades de medida (Cohen y Swerdlik, 2001). En los últimos años, el creciente uso de pruebas estandarizadas ha causado una revolución en las políticas educativas mexicanas. El impacto de las pruebas de Evaluación Nacional de Logro Académico en Centros Escolares (ENLACE), ha sido sustancial, sobre todo en los incentivos y reconocimientos a las escuelas de alto rendimiento y se perfila como un hito en la historia reciente de la educación, por ser el primer indicador directo del aprendizaje del estudiante de educación básica que es atendido en el sistema educativo mexicano.

En educación superior, el Examen Nacional de Ingreso a la Educación Superior (EXANI) es utilizado para la admisión, acreditación y reconocimiento de diversos programas de estudios y; en educación especial, ha habido un incremento significativo en la utilización de pruebas psicológicas estandarizadas para evaluar las necesidades educativas especiales.

Los sistemas escolares deben contar con una batería de pruebas de aprovechamiento estandarizada, con el objeto de tomar decisiones en aspectos académicos de los estudiantes (Hogan, 2004). En México, hay todavía vacíos importantes en el área de medición y evaluación educativa, como es caso de los estudiantes altamente sobresalientes (EAS).

Por lo anterior, el objetivo de la presente investigación es diseñar, desarrollar y validar pruebas de rendimiento académico que sean útiles para apoyar el diagnóstico integral y la identificación oportuna del EAS en México, considerando el currículo de la

educación básica y el contexto nacional. Estas pruebas de aprovechamiento, evalúan logro escolar y no aptitud potencial, y están diseñados para alumnos considerados sobresalientes sin discapacidad.

*Detección del alumno altamente sobresaliente*

Para la identificación de los EAS, las pruebas de inteligencia han sido tradicionalmente el criterio para determinar la capacidad sobresaliente, pero no es el único criterio que se debe usar para la adecuada identificación. Por esta razón, se utilizan como complemento, diversos tipos de pruebas como las de aptitudes específicas. En México, la más utilizada es la SAGES-2 (Screening Assessment for Gifted Elementary and Middle School Student–Second Edition) que fue diseñada en Estados Unidos de América y que se ha estandarizado y validado para la población mexicana (Orta y Sánchez, 2003). Con relación a la estandarización, los resultados muestran que en cuanto a la administración y contenidos de la prueba, sólo fue posible sustentar la validez concurrente de la prueba SAGES-2 (4-S) y se reportaron coeficientes de confiabilidad relativamente bajos en lagunas de las escalas; en este sentido Sánchez y Acle (2003) a cargo de la estandarización en México, atribuyó estos hallazgos a que esta prueba derivaba del currículo de la educación básica Norteamericana, la cual es sustancialmente diferente al plan nacional de educación básica en México.

Al revisar los ítems en cuanto a contenido, muchos de estos incluyen información no contenida en el currículo oficial de la educación básica mexicana. Además, existen diferencias entre los resultados de procedimientos en México y Estados Unidos, por lo que se necesita investigar más acerca del tópico y diseñar pruebas que sean específicas para el contexto (Sánchez y Acle, 2003).

Marland (citado por Heward y Orlansky, 1998), menciona que la identificación de los niños con AS exige una comparación con los estándares normativos, o sea, aplicables a un contexto determinado. Callahan (cit. en Ohio State Department of Education, 2000), menciona que se necesitan de instrumentos válidos y confiables para la adecuada detección de las habilidades sobresalientes.

Respecto a los EAS Johnsen y Corn (2001), Sánchez, Cantón y Sevilla (1997) indican la importancia de la identificación de estos niños de manera confiable, temprana y objetiva para ser referidos a los programas que les ayuden a desarrollar tanto sus talentos como sus potencialidades.

Blanco (2001), menciona que es necesario utilizar pruebas de tipo objetivas y subjetivas como parte del proceso de identificación. Una de las pruebas objetivas es la administración de pruebas de rendimiento (actuación académica). En los primeros cursos de primaria, los mejores indicadores derivan de pruebas de lectura y matemáticas. Sin embargo, en México, hay pocas pruebas válidas y confiables estandarizadas de aprovechamiento y las utilizadas son algunas veces traducidas del inglés (Sánchez y Acle, 2003).

En México se han establecido las pruebas de estándares nacionales, que tienen el propósito de evaluar el sistema educativo y la calidad educativa más que el rendimiento de los estudiantes, además, las pruebas no fueron creadas de acuerdo a los planes y programas de estudio nacionales, razón por la cual se puede decir poco acerca del dominio de los contenidos curriculares que poseen los estudiantes (Backhoff y Martínez Rizo, 2004 y Backhoff, 2005, citados por Backhoff, Sánchez, Peón, Monroy y Tanamachi, 2006).

A pesar de que existen pruebas de inteligencia general y aptitudes específicas, que son usadas en la identificación de niños con aptitudes sobresalientes, no existe una prueba

de conocimientos estandarizada para la adecuada identificación. La importancia de este tipo de pruebas de conocimientos o rendimiento académico, es que los puntajes serían buenos indicadores de la aptitud sobresaliente. Además de ser uno de los primeros filtros que pueden estar al alcance de profesores y maestros, ya que la detección oportuna se realiza en la escuela elemental.

Por lo anterior, el diseñar pruebas de conocimientos de Español y Matemáticas de nivel primario, facilitará a los profesores la identificación oportuna de EAS y normar criterios para la ubicación avanzada o aceleración en su caso.

### Objetivos

1. Diseñar una batería de pruebas de conocimientos de español y matemáticas de nivel primaria.
2. Explorar la validez y confiabilidad de las pruebas.
3. Estimar el uso de las pruebas como una herramienta de identificación de niños con aptitudes sobresalientes.

### Metodología

El presente estudio es de corte cuantitativo, de tipo descriptivo y se basa en la teoría Clásica del desarrollo de pruebas y de instrumentos de medición y evaluación educativas.

### *Población*

La población son los alumnos de tercero, cuarto y quinto grado de primaria de escuelas públicas y privadas de la ciudad de Mérida. Se trabajó con estudiantes de primaria debido a que las pruebas están relacionadas a una identificación oportuna, además que en el nivel básico existe el Programa para niños con aptitudes sobresalientes. Únicamente se trabajó con niños de tercero, cuarto y quinto debido a la naturaleza de las pruebas.

*Muestra*

La muestra de alumnos se eligió por conglomerados, ya que se utilizaron grupos enteros de las 14 escuelas primarias seleccionadas. Para contestar las pruebas, de cada grupo o conglomerado, se seleccionaron a los alumnos que cumplieran con el criterio de tener un promedio superior a 9 puntos. Se eligió el criterio de la calificación, ya que el alto rendimiento académico está reflejado en las calificaciones. La calificación es uno de los indicadores para identificar a los niños potencialmente sobresalientes.

Las escuelas participantes se eligieron por conveniencia, considerando la división de zonas de la ciudad de Mérida, Yucatán México y abarcando todos los puntos geográficos de la ciudad, así como el acceso y voluntad de participar en el estudio y proveer de facilidades de acceso al investigador principal. Se eligieron 14 escuelas primarias: 11 públicas y 3 privadas.

De los 341 participantes, 124 (36.4%) son hombres y 217 (63.6%) son mujeres. La Tabla 1 muestra las frecuencias y porcentajes de género por cada grupo participante.

Tabla 1

*Frecuencias y porcentajes de género por cada grupo*

<i>Grado</i>	<i>Masculino</i>		<i>Femenino</i>		<i>Total</i>	
	<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>
Tercero	47	38.5	75	61.5	122	35.8
Cuarto	30	27.8	78	72.2	108	31.7
Quinto	47	42.3	64	57.7	111	32.6
Total	124	36.4	217	63.6	341	100

### Instrumentos

Se diseñaron seis instrumentos de conocimientos de Español y Matemáticas en nivel primaria, tres de Español y tres de Matemáticas. Para la elaboración de los instrumentos se utilizó el Plan y programa de educación básica primaria y los libros de texto gratuitos de la SEP. Los libros de texto gratuitos fueron analizados y se elaboró un listado de los contenidos, los cuales fueron utilizados para diseñar las pruebas.

Para cada grado se diseñó una prueba basada en el contenido del grado actual y con el contenido del grado superior inmediato, esto es, las pruebas de tercero tienen contenidos de tercero y cuarto, las pruebas de cuarto tienen contenido de cuarto y quinto y las de quinto tienen contenidos de quinto y sexto grado.

El propósito de realizar este tipo de prueba fue identificar los conocimientos, tanto del grado actual como del grado superior inmediato que tienen los alumnos de alto rendimiento, con el fin de estimar las probabilidades de éxito en el caso de ubicación adelantada (aceleración) del estudiante, bajo el supuesto de si un niño tiene los contenidos del grado superior, éste podrá acceder a la aceleración por grado.

Cada una de las pruebas de español está formada por 20 preguntas. Diez preguntas corresponden al contenido del grado actual y las otras diez al contenido del grado superior inmediato. Cada una de las preguntas tiene cuatro opciones de respuesta, una de las cuales es la correcta.

Cada una de las pruebas de Matemáticas está formada por 24 preguntas, 12 correspondientes al contenido del grado actual y las otras 12 correspondientes al contenido del grado superior inmediato. Cada una de las preguntas tiene cuatro opciones de respuesta, una de las cuales es la correcta (Ver Apéndice C).

Como criterio de validez externa se utilizó la prueba de Matrices de la escala Weschler para niños, versión IV (WISC-IV), la cual es esencialmente de razonamiento perceptual, en virtud de que se ha reconocido que las tareas de analogías de matrices constituyen buenas medidas de la inteligencia fluida y estimados confiables de la capacidad intelectual general (Brody, 1992; Raven, Raven y Court, 1998, citados por Weschler, 2007). Esta rutina trata de evaluar del procesamiento de la información y del razonamiento abstracto y contiene 35 fichas con matrices incompletas, en la que se le pide al sujeto que seleccione, entre cinco opciones, la respuesta que complete mejor la matriz.

Los estudios han mostrado una elevada correlación general de .72 entre las pruebas de analogías de matrices y las puntuaciones del Coeficiente Intelectual Total (CIT) de las y se les considera “relativamente libres de la influencia cultural y del lenguaje y no requieren de manipulación” (Weschler, 2007, p. 15).

Tabla 2

*Correlación de la subprueba de Matrices*

	<i>Español</i>	<i>Matemáticas</i>	<i>Matrices</i>
Español	-	.093	.093
Matemáticas		-	.022
Matrices			-

Como se puede observar la correlación es alta entre la escala de Matrices y las pruebas de Español; sin embargo, entre las pruebas de Matemáticas y la subprueba de Matrices no hay correlación significativa, hallazgo que por inesperado amerita investigación más amplia.

La confiabilidad se obtuvo a través del Alfa de Cronbach. En general, los coeficientes fueron relativamente bajos para este tipo de pruebas oscilando entre 0.44 y 0.68. Estos resultados se discuten más adelante.

Resultados

*Discrepancias del criterio por grado*

Una vez analizadas las pruebas de manera general, se procedió a realizar un análisis del contenido por grado y del contenido por grado superior en cada una de las seis pruebas. Esto es, cada una de las pruebas se dividió en dos: el contenido del grado y el contenido del grado superior inmediato. La tabla 3, resume los resultados de las pruebas.

Tabla 3

*Media y desviación estándar de los puntajes del contenido por grado actual y por grado superior*

<i>Pruebas</i>	<i>Español</i>				<i>Matemáticas</i>			
	<i>X̄ del grado actual (DS)</i>	<i>X̄ del grado superior (DS)</i>	<i>t</i>	<i>p</i>	<i>X̄ del grado actual (DS)</i>	<i>X̄ del grado superior (DS)</i>	<i>t</i>	<i>p</i>
Tercero	5.17 (1.73)	4.08 (1.55)	6.17	.000	7.27 (1.70)	4.49 (1.70)	15.08	.001
Cuarto	5.00 (1.72)	2.96 (1.45)	10.47	.000	4.12 (1.89)	5.99 (2.06)	9.35	.001
Quinto	3.94 (1.51)	3.93 (1.50)	.051	.959*	6.30 (2.06)	4.11 (2.10)	10.08	.001

Como se observa en la Tabla, no hay diferencias en los ítems de quinto y sexto grado, por lo que se podría inferir que los alumnos tienen los conocimientos del grado superior inmediato, o que en su defecto, los contenidos de ambos grados son muy similares. Esto requiere tanto de un análisis curricular como de la estructuración de las pruebas.

## Discusión

Los resultados de la administración de pruebas derivadas del currículo de primaria para tercero, cuarto y quinto grado ilustran las dificultades de evaluar el rendimiento escolar de los alumnos de primaria en México. En específico, hay tres hallazgos que invitan a la reflexión.

En primer lugar, los índices de confiabilidad fueron muy bajos y se alían a los reportados en investigaciones previas y pruebas similares, como la de Sánchez y Acle (2000) y las pruebas de ENLACE.

En segundo lugar, existe evidencia que vulnera la validez externa, en especial la ausencia de correlación de la escala de matemáticas con el criterio externo de la prueba de matrices del WISC-4. Esto es inquietante, sobre todo si se considera que los reactivos son derivados de los programas de estudio vigentes, a través de tablas de especificaciones y con las habilidades requeridas en la teoría general de las pruebas y que las pruebas fueron respondidas por alumnos de alto rendimiento.

En tercer lugar, no hay diferencias en los puntajes e ítems del quinto y sexto grado, lo que cuestiona la dificultad diferencial esperada en el sistema escolar mexicano.

Estos tres puntos invitan a reflexionar y a investigar los retos de la medición del logro escolar en México. Posibles áreas de exploración son: la falta de conocimiento y utilización adecuada de las pruebas académicas estandarizadas, el desapego a los contenidos curriculares, y desde luego, a problemas técnicos en elaboración de criterios y reactivos, el número de ellos y al grado de dificultad en su redacción.

Se sugiere continuar con la realización de investigaciones enfocadas a la atención de niños sobresalientes. De la misma manera investigar sobre la estandarización y normalización de pruebas que sean útiles para nuestro sistema educativo.

## Referencias

- Backhoff, E., Sánchez, A., Peón, M., Monroy, L. y Tanamachi, L. (2006). Diseño y desarrollo de los exámenes de la calidad y logro educativos. *Revista Mexicana de Investigación Educativa*, 11, 029, 617-638.
- Blanco, M. (2001). *Guía para la identificación y seguimiento de alumnos sobresalientes*. España: CISS PRAXIS.
- Cohen, R. y Swerdlik, M. (2004). *Pruebas y evaluación psicológicas. Introducción a las pruebas y a la medición*. (4ª. Ed.). México: McGraw-Hill.
- Heward, W. y Orlansky, M. (1998). *Programa de Educación Especial 1*. España: CEAC, S.A.
- Hogan, T. (2004). *Pruebas psicológicas. Una introducción práctica*. México: Manual Moderno.
- Johnsen, S. y Corn, A. (2001). *Sages-2 Screening Assessment for gifted elementary and middle school students*. (2a. Ed.) Austin, Texas: Pro-Ed.
- Ohio State Department of Education. (2000). Identification of Children who are gifted. *Disabilities and Gifted Education*. [En red] Recuperado de:  
<http://www.ode.state.oh.us>
- Orta, Y. y Sánchez, P. (2003). *Habilidades verbales de niños con habilidades sobresalientes: un estudio exploratorio preliminar en Yucatán*. Artículo no publicado, Universidad Autónoma de Ciudad Juárez y Universidad Autónoma de Yucatán.
- Sánchez, P. y Acle, G. (2003). *Validación y normas para la población del SAGES-2 mexicana*. México: UNAM.
- Weschler, D. (2007). *Escala de inteligencia para niños-IV. Manual técnico*. México: El Manual Moderno.