

DISEÑO, OPERACIÓN Y EVALUACIÓN DEL PROGRAMA PARA EL ANÁLISIS GRÁFICO DE ÍTEMS (PAGI®), PARA HABILITAR A DOCENTES EN SERVICIO EN LA MEJORA DE SUS PRUEBAS DE AULA

LUIS ÁNGEL CONTRERAS NIÑO / GUADALUPE DE LOS SANTOS LÁZARO / ERICK URÍAS LUZANILLA
Universidad Autónoma de Baja California

RESUMEN: Aunque la evaluación del aprendizaje es una actividad cotidiana en las aulas, existe evidencia de que en México los profesores tienen una competencia limitada para evaluar el logro de sus estudiantes de manera justa, válida y congruente con lo que enseñan. Lo anterior justificó una intervención para apoyarlos, a fin de mejorar las pruebas de aula que desarrollan como parte de su trabajo habitual. Se consideró que la técnica del análisis gráfico de ítems constituye un enfoque que permite al docente analizar de manera simple, directa y visualmente la forma en que sus estudiantes responden ítems de respuesta seleccionada. En consecuencia, el propósito del estudio fue diseñar un programa de cómputo, denominado PAGI® que facilitara a los docentes conocer y mejorar las características que definen la cali-

dad técnica de ítems de opción múltiple y respuesta alterna que utilizan en sus pruebas, mediante el análisis gráfico de ítems. Tras su desarrollo, se habilitó a 33 profesores de diversos niveles educativos en el manejo del software, mediante un curso-taller que tuvo una duración de cuatro horas. Los resultados muestran que los docentes fueron capaces de manejar el programa y de emitir juicios razonados y congruentes con las gráficas que produjo PAGI® sobre los ítems de sus pruebas de aula. En consecuencia, se considera que el programa representa una aportación relevante para la formación y la actualización de profesores en el área de la evaluación del aprendizaje.

PALABRAS CLAVE: Análisis gráfico de ítems, Pruebas de aula, Actualización docente.

Introducción

La evaluación del aprendizaje es inherente al proceso educativo. Sin embargo, existe evidencia de que los profesores tienen una capacidad limitada para calificar a sus estudiantes de manera justa, válida, confiable y congruente con lo que enseñan (Shepard, 2006). En México, ello puede explicarse porque el currículum para la formación inicial de profesores privilegia contenidos teóricos, por lo que los futuros docentes no tienen competencias necesarias para evaluar el aprendizaje con calidad (Martínez, Maya y Zenteno, 1996). Además, la actualización que reciben no es suficiente para lograr un desempeño eficaz (Lozano, 2007).

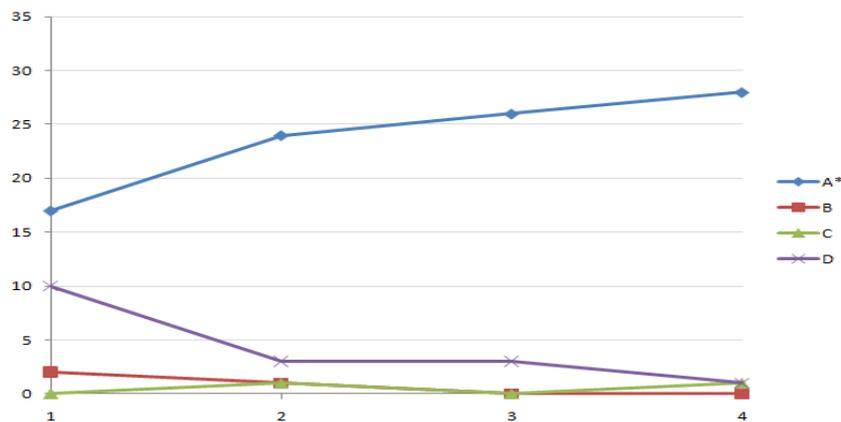
Lo anterior justificó una intervención para apoyarlos, de manera que puedan mejorar las pruebas que desarrollan como parte de su trabajo cotidiano. Así, este estudio buscó desarrollar un programa de cómputo que facilitara a los docentes adquirir y utilizar conocimientos y habilidades elementales de psicometría, y que les permitiera saber si los ítems de tipo objetivo que utilizan en sus pruebas tienen dificultad apropiada, buena discriminación y un funcionamiento adecuado de los distractores.

Se consideró que la técnica del Análisis Gráfico de Ítems (AGI) desarrollada por Batenburg y Laros (2002), es un enfoque que permite al docente analizar de manera simple, directa y visual cómo responden sus estudiantes ítems de respuesta seleccionada. Haladyna (2004) considera que una representación gráfica permite observar variaciones en las tendencias de la respuesta correcta y los distractores en un ítem de respuesta seleccionada, respecto a la calificación en la prueba. Así, el análisis de ítems resulta más significativo e interpretable y pueden comprenderlo fácilmente personas ajenas a la psicometría, sin antecedentes estadísticos necesarios para calcular los índices tradicionales de los parámetros del ítem.

En consecuencia, el objetivo general del estudio fue desarrollar, operar y evaluar un programa de cómputo que permitiera a profesores en servicio, mejorar sus pruebas de aula mediante el uso de la técnica del análisis gráfico de ítems.

Cabe señalar que el análisis gráfico de ítems parte del supuesto de que quienes obtienen una puntuación máxima en una prueba, tienen alta probabilidad de seleccionar la opción correcta en cada ítem, y que la elección de opciones falsas o distractores disminuye al aumentar la puntuación total. Lo anterior puede observarse en la Figura 1.

Figura 1. Frecuencias de respuestas a la opción correcta y a los distractores, por grupos de habilidad, en un ítem de opción múltiple con cuatro opciones



* Se trata de la respuesta correcta

En la figura 1, se observa que el eje Y muestra la frecuencia de examinados que respondieron el ítem. En el eje X se observan cuatro grupos de habilidad que se formaron a partir de la calificación que obtuvieron los alumnos en la prueba: grupo 1, bajo; grupo 2, medio bajo; grupo 3, medio alto; y grupo 4, alto (con las calificaciones más altas en el examen). En este caso, en cada grupo de habilidad hay 30 examinados, porque el ítem fue aplicado a una muestra de 120 estudiantes.

Las líneas que se observan en el eje de coordenadas son las curvas de respuestas al ítem. Se observa que la curva de la respuesta correcta (A) aumenta a medida que incrementa la habilidad. Por otro lado, las líneas que están marcadas como B, C y D son llamados distractores, y para ser efectivos sólo deben elegirlos los alumnos que no dominan el contenido.

El ítem que se ilustra en la figura 1 se considera idealmente bueno, porque la curva de la respuesta correcta aumenta conforme incrementa la habilidad y la elección de los distractores disminuye a medida que incrementa la habilidad.

Método

Participantes

En el estudio participaron 21 profesores en servicio de secundaria de Ensenada, Baja California, con diversas especialidades, y 12 profesores de bachillerato y de universidades de Tampico y la Ciudad de México. Para participar en el estudio evaluativo los docentes debieron cumplir con los siguientes requisitos: traer una laptop y un archivo en Excel© de Microsoft con las respuestas brutas que dieron 32 o más alumnos a por lo menos 10 ítems de una prueba de aula aplicada previamente.

Materiales y procedimientos

Para habilitar a los docentes, se desarrollaron y aplicaron los siguientes materiales e instrumentos:

- Programa para el Análisis Gráfico de Ítems (PAGI®). Como la técnica del análisis gráfico de ítems se desarrolló en el contexto de la evaluación del aprendizaje a gran escala, fue necesario adaptarla a las características y condiciones de la evaluación en el aula. El problema era garantizar que con pocos estudiantes se podían observar claramente los cambios de tendencia en la curvas de la respuesta co-

recta y de los distractores. Para ello, con el apoyo de una base de datos auténtica derivada de la aplicación a gran escala de la prueba EFCyE elaborada por la Universidad Autónoma de Baja California, se hicieron variar sistemáticamente el número de grupos de habilidad (2, 3, 4 y 5) en que podrían dividirse las puntuaciones en dicha prueba, así como el número total de examinados (10, 15, 20, 25, 30, 35, 40, 45, 50) que se podrían asignar a cada grupo de habilidad. Tras los correspondientes ensayos se observó que, cuando se representaron las puntuaciones en la prueba en 4 grupos de habilidad, y se asignó un mínimo de ocho examinados a cada grupo de habilidad (32 casos en total), era posible observar con claridad las tendencias de las respuestas al ítem. Posteriormente se desarrolló el programa de cómputo PAgI® para que los docentes utilizaran dicha técnica en sus evaluaciones de aula.

- Manual del participante. Contiene información conceptual y metodológica sobre aspectos elementales de la teoría clásica de los tests, el análisis gráfico de ítems y la interface gráfica del programa de cómputo PAgI®.

De manera complementaria, se desarrollaron: a) un cuestionario para evaluar la operación del programa de cómputo, mismo que se estructuró con 28 ítems tipo Likert cada uno de ellos con cinco opciones de respuesta, para manifestar el grado de acuerdo respecto a afirmaciones correspondientes al entrenamiento proporcionado y al uso del programa de cómputo; y b) una guía para apoyar la interpretación de las gráficas producidas por el programa PAgI® sobre la calidad técnica de los ítems.

Tras elaborar estos materiales e instrumentos se procedió a habilitar a los docentes en el manejo del programa PAgI®. Para ello, se instrumentaron dos cursos – taller (uno para docentes de secundaria y otro para docentes de bachillerato y educación superior), cada uno de ellos en una sesión de cuatro horas. Primero se revisaron conceptos necesarios para la interpretación gráfica de los ítems y para entender la interface gráfica de PAgI®. Después cada participante utilizó el programa y analizó las gráficas que produjo PAgI® sobre sus ítems, hasta que fue capaz de registrar en la guía un dictamen razonado sobre la calidad técnica de cada uno de sus ítems.

Posteriormente, se aplicó el cuestionario para evaluar la operación de PAgI®, el entrenamiento y los materiales elaborados.

Resultados

El principal resultado del estudio fueron los propios materiales e instrumentos que fueron desarrollados para propiciar que los docentes pudieran determinar y analizar los parámetros de sus ítems; es decir, el propio programa PAGI®, el manual para poder utilizarlo, la escala Likert para evaluar su operación y la guía para apoyar la interpretación de las gráficas que produjo el programa sobre los ítems.

El cuestionario para evaluar la operación mostró evidencias de calidad técnica. En cuanto a confiabilidad, se obtuvo un coeficiente Alfa de 0.88 y, la media de correlación ítem – total fue de 0.53. Así, se procedió a analizar los resultados de su aplicación. La opinión manifestada por los profesores fue muy favorable en general. El promedio de adscripción a los ítems fue de 4.57 en la escala Likert del 1 al 5. Respecto a las tres dimensiones evaluadas, el promedio más alto (4.64 en la escala del 1 al 5) correspondió a la dimensión Instrucción, cuyos ítems evaluaban el dominio del contenido, la planeación didáctica y la actuación de la instrucción; seguida de Materiales de apoyo, que incluían el programa de cómputo PAGI®, el manual del participante y el formato para evaluar los ítems, con 4.51; y Operación de la capacitación, que incluía el impacto percibido de lo aprendido en su futura práctica docente, y el desarrollo de las actividades de capacitación, con 4.50. El aspecto en que estuvieron menos de acuerdo los profesores, en esta última dimensión evaluada, fue el Desarrollo de actividades (4.37), principalmente porque consideraron que el tiempo de la capacitación fue insuficiente.

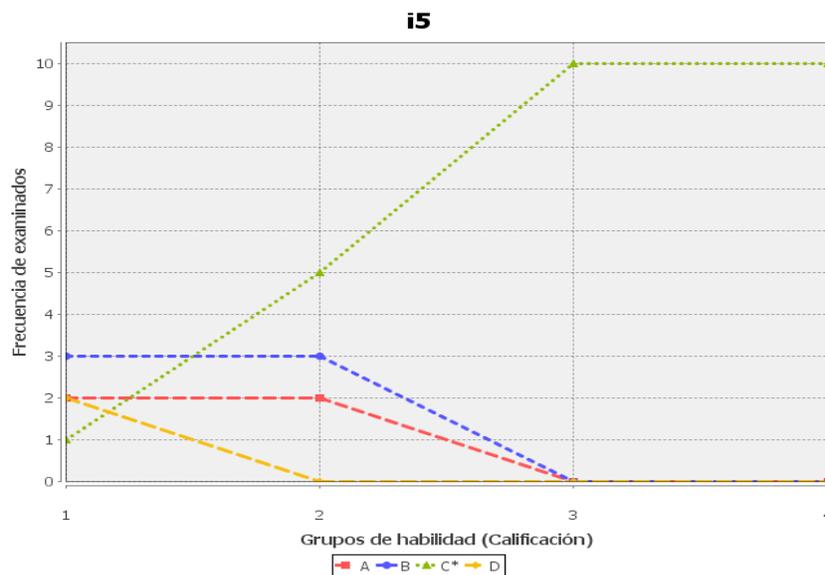
Respecto a los juicios que plasmaron los docentes en la guía para interpretar el comportamiento de sus ítems, se analizó su correspondencia con cada gráfico que produjo PAGI®. En general, se observó que los docentes formularon comentarios significativos y congruentes con las gráficas que produjo el software. En particular, fueron congruentes la identificación de distractores poco funcionales, la estimación de la dificultad y el poder discriminativo del ítem, así como el dictamen general sobre eliminar el ítem o sobre qué aspectos modificar en el ítem a fin de mejorarlo. Lo anterior se ilustra en la Tabla 1, la cual presenta los comentarios que formuló en la guía uno de los participantes sobre la ejecución de sus estudiantes, en uno de los ítems evaluados; así como en la figura 2, la cual muestra la correspondiente gráfica que produjo el programa PAGI® sobre dicho ítem evaluado por el participante.

Tabla 1. Comentarios del participante número 29 que evaluó el ítem 5 de su prueba

Profesor	Ítem	Opción	Comentarios
29	5	a	Solo lo contestan los de los grupos 1 y 2, por lo demás está bien
		b	Solo lo contestan los de los grupos 1 y 2
		c*	Muestra pendiente positiva, a la alza
		d	Buen distractor pues solo lo eligen los del grupo 1
		Juicio final sobre el ítem	Ítem algo difícil que discrimina bien. Es buen reactivo

* Se trata de la respuesta correcta

Figura 2. Gráfica producida por PAGI® a partir de las respuestas al ítem 5 del participante 29



Conclusiones y recomendaciones

Los resultados obtenidos permiten concluir que se cumplieron los objetivos propuestos. Así, fue posible desarrollar, operar y evaluar el programa de cómputo para habilitar a docentes en servicio de modo que puedan mejorar sus pruebas objetivas mediante el análisis gráfico de ítems. Los siguientes argumentos apoyan esta conclusión:

- Los profesores reportaron aprendizajes significativos de conocimientos y habilidades básicas de psicometría, necesarios para el análisis gráfico de sus ítems.

- En opinión de los participantes, PAGI® les permitió identificar fortalezas y debilidades de sus ítems, según los indicadores psicométricos que definen su calidad técnica, así como formas específicas de mejorarlos. Sus participaciones durante la capacitación y los productos que generaron tras analizar sus ítems y llenar la guía, confirman esta opinión.
- El instrumento de evaluación arrojó datos que muestran su calidad psicométrica y por ello se puede afirmar que en opinión de los docentes los materiales diseñados y la instrucción propiciaron aprendizajes significativos que fueron pretendidos.

No obstante, el estudio tiene varias limitaciones. Entre ellas, puede señalarse que el análisis gráfico de ítems es una técnica que incluye aspectos como análisis del sesgo en ítems y la construcción de versiones paralelas de una prueba, que no se incluyeron en el trabajo por razones temporales. Por ello, se sugiere rediseñar PAGI® y el manual para incluir tales aspectos, lo que hará más útil y versátil el programa para los docentes.

También se requiere utilizar el programa de cómputo con un número mayor de profesores y de otros niveles y modalidades educativos, antes de pensar en la posibilidad utilizarlo de manera más extensiva tanto en la formación inicial, como en la actualización de los profesores en servicio.

Bibliografía

- Batenburg, T. V. & Laros, J. A. (2002). Graphical Analysis of Items. *Education research and evaluation*, 8(3), 319 – 333.
- Haladyna, M. (2004). Developing and validating multiple choice test items. Consultado el 15 de Noviembre de 2009 en: <http://books.google.com.mx/books>
- Lozano, M. E. (2007). Formación docente: El servicio de asesoría académica a las escuelas y el trayecto formativo. En *Memorias del IX Congreso Nacional de Investigación Educativa* [Disco compacto] Yucatán, México: Universidad Autónoma de Yucatán.
- Martínez, M., Maya, C. y Zenteno, E. (1996). Hacia una evaluación institucional de la Educación Normal. En *Memorias del II Foro de Evaluación Educativa*. Saltillo, Coahuila, México: Centro Nacional de Evaluación para la Educación Superior.
- Shepard, L. (2006). La evaluación en el Aula. En Robert I. Brennan (Eds), *Identification of Mild Handicaps* (pp545- 572). Bolder, Colorado, E.E.U.U. (Reimpreso en México por Instituto Nacional para la Evaluación de la Educación, Ed. 2006, México. Consultado el 30 de octubre de 2008 en: <http://www.inee.edu.mx/>.