VOLUMEN 3 NÚMERO 1 2015

Revista Internacional de

Educación y Aprendizaje

Análisis de la estructura cognitiva del área de habilidades cuantitativas del EXHCOBA mediante el modelo LLTM de Fisher

JUAN CARLOS PÉREZ MORÁN, NORMA LARRAZOLO REYNA, EDUARDO BACKHOFF ESCUDERO, GUANER ROJAS

COMMON GROUND

SOBRELAEDUCACION.COM

Análisis de la estructura cognitiva del área de habilidades cuantitativas del EXHCOBA mediante el modelo LLTM de Fisher

Juan Carlos Pérez Morán, Instituto Nacional para la Evaluación de la Educación, México
Norma Larrazolo Reyna, Universidad Autónoma de Baja California, México
Eduardo Backhoff Escudero, Instituto Nacional para la Evaluación de la Educación, México
Guaner Rojas, Universidad de Costa Rica, Costa Rica

Resumen: *En el presente estudio se integran aspectos de la psicología cognitiva y la psicometría (Rupp & Mislevy, 2006) para el análisis de las fuentes de dificultad del área de Habilidades Cuantitativas (HC) del Examen de Habilidades y Conocimientos Básicos (EXHCOBA) (Backhoff, Tirado, & Larrazolo, 2001), mediante el Modelo Logístico Lineal de Rasgo Latente (LLTM, por sus siglas en inglés) desarrollado por Fischer (1973). Para alcanzar dicho propósito, un grupo de expertos propuso catorce operaciones para resolver los ítems del área HC apoyándose en los resultados de reportes verbales concurrentes y retrospectivos realizados a un grupo de estudiantes de educación secundaria. Además, se analizó la estructura cognitiva determinada por los expertos y se evaluó la dimensionalidad y el ajuste de los ítems a los modelos Rasch y LLTM. Los resultados mostraron que los datos ajustan moderadamente al modelo LLTM con una matriz Q de catorce atributos y que un refinamiento de las operaciones para resolver los ítems podría explicar mejor sus fuentes de dificultad y el rendimiento de los estudiantes en la prueba.*

Palabras clave: *modelo Rasch, estructura cognitiva, matriz Q, examen de admisión*

Abstract: *The present study integrates aspects of cognitive psychology and psychometrics (Rupp & Mislevy, 2006) for the analysis of the sources of difficulty in the area of quantitative skills (HC, for its acronym in Spanish) of the Basic Skills and Knowledge Test (EXHCOBA, for its acronym in Spanish) (Backhoff, Tirado, & Larrazolo, 2001), using the Linear Logistic Test Model (LLTM) developed by Fischer (1973). To achieve this purpose, a group of experts proposed fourteen operations to solve items in the HC area, based on the results of retrospective and concurrent verbal report,s obtained from a group of students of secondary education. In addition, the cognitive structure determined by experts was analyzed, and the dimensionality and the adjustment of the items to Rasch model and LLTM was assessed. The results showed that the data fit moderately to the LLTM with a Q-matrix of fourteen attributes and, also that, a refinement of the operations to resolve the items could explain better the sources of difficulty and the performance of examinees in the test.*

Keywords: *Rasch Model, Cognitve Structure, Q-Matrix, Admission Test*

Introducción

Conocer las características de los atributos cognitivos latentes requeridos para producir respuestas correctas en los ítems de las pruebas educativas, puede ser muy útil en diferentes contextos y procesos dentro del ámbito educativo. Uno de estos procesos que se beneficia de forma directa y casi inmediata es el desarrollo de instrumentos de medición educativa. Por otra parte, algunos de los procesos educativos que también se benefician al conocer las características de los atributos cognitivos que subyacen a las puntuaciones de las pruebas son la realimentación y operación del currículum, el diagnóstico de las estrategias de enseñanza (Abida, Azeem & Bashir, s.f.) y la mejor comprensión del proceso de aprendizaje (Cortada de Kohan, 2000). Las estructuras cognitivas han sido analizadas en anteriores investigaciones a través de diferentes enfoques que integran la psicología cognitiva y el análisis psicométrico (p. ej., DiBe-

© Common Ground. J.C. Pérez Morán *et al.* Todos los derechos reservados.
Permisos: soporte@commongroundpublishing.com

llo, Stout & Roussos, 1995; Embretson, 1984, 1995; Fischer, 1973; Tatsuoka, 1985; Romero, 2010). En el contexto de la medición con el modelo de Rasch, la prueba del modelo logístico lineal se utiliza con frecuencia para predecir la dificultad de los ítems ante los atributos cognitivos (p. ej., Embretson & Wetzel, 1987; Fischer, 1973; Spada, 1977; Spada & May, 1982).

Modelo Logístico Lineal de Rasgo Latente (LLTM)

Por su parte, los modelos componenciales de la Teoría de Respuesta al ítem (TRI) facilitan a la Psicometría la medición de procesos estudiados por la Psicología Cognitiva (Embretson, 1994). Algunos de estos modelos son el reflejo de un gran esfuerzo (véase Mislevy, 1995) por mezclar modelos ampliados de la TRI y modelos cognitivos basándose en una descomposición lineal de $\boldsymbol{\beta_j}$ o $\boldsymbol{\theta_i}$. Un ejemplo de ello es el Modelo Logístico Lineal de Rasgo Latente (LLTM, por sus siglas en inglés) de Fischer (1973, 1995). Dicho modelo, permite estimar la dificultad de los ítems y la contribución de los diferentes componentes establecidos por las teorías cognitivas, y decidir si estos son significativos. La descripción detallada del fundamento psicométrico de este modelo se puede encontrar en las obras de Fischer y Molenaar (1995), y Van der Linden y Hambleton (1997).

De manera particular, en el análisis con el LLTM, $\boldsymbol{\beta_j}$ se reescribe como una combinación lineal de K parámetros básicos $\boldsymbol{\eta_k}$ con pesos $\boldsymbol{q_{jk}}$ y lógito

$$P_j(\theta_i) = \theta_i - \sum_{k=1}^{K} q_{jk}\eta_k \qquad (1)$$

donde Q= $[q_{jk}]$ es una matriz generalmente obtenida a priori basada en un análisis de los ítems sobre los atributos cognitivos necesarios para resolverlos, y η_k es la contribución del atributo K a la dificultad de los ítems de dicho atributo (Junker & Sijtsma, 2001).

Otros modelos de la TRI como los modelos compensatorios multidimensionales (véase, Adams, Wilson & Wang, 1997) siguen la tradición del análisis factorial. En dichos modelos, se descompone el parámetro θ_i de unidimensionalidad en una combinación lineal del ítem-dependiente a los rasgos subyacentes con lógito:

$$P_j(\theta_i) = \sum_{k=1}^{K} B_{jk}\ \theta_{ik} - B_j \qquad (2)$$

Los modelos compensatorios de la TRI, como es el caso de los modelos de análisis factorial, pueden ser sensibles a componentes relativamente grandes que presentan variación en θ. Sin embargo, estos modelos generalmente no están diseñados para distinguir los componentes más finos de variación entre los estudiantes, que suelen ser de interés en la evaluación cognitiva. Por otra parte, el modelo LLTM puede ser sensible a estos componentes más finos de variación entre los ítems, pero tiene la particularidad de no estar diseñado para ser sensible a los componentes de la variación entre los estudiantes (Junker & Sijtsma, 2001).

De forma específica, el LLTM se obtiene del modelo de Rasch dicotómico:

$$P(x = 1|\theta,\delta_1) = \frac{e^{(\theta-\delta_1)}}{1 + e^{(\theta-\delta_1)}}\ , \qquad (3)$$

en el cual el parámetro de dificultad del ítem δ_1 se descompone de forma lineal en los parámetros de dificultad de cada uno de los componentes (Fischer & Molenaar, 1995):

$$\delta_i = \sum_{k=1}^{p} q_{ik}\ \alpha_k + c \qquad (4)$$

En la ecuación (4) δ_1 es el parámetro de dificultad del i-ésimo ítem en el modelo de Rasch; α_k, k=1,…, p son los parámetros básicos del LLTM y corresponden a las dificultades de cada

componente k; q_{ik} son los pesos dados de los parámetros básicos α_k, representando la complejidad correspondiente al ítem i en el componente k-ésimo; y c es una constante de normalización. Cabe señalar que la aplicación del modelo LLTM mediante la ecuación (4) tiene sentido sólo si el modelo de Rasch se ajusta suficientemente bien a los datos. Por otra parte, si los componentes considerados explican de manera exhaustiva las diferencias entre los ítems, se deberían recuperar a través de la ecuación (4) estimaciones δ_1 similares a las obtenidas directamente del modelo de Rasch, lo que implicaría una alta correlación entre los parámetros estimados bajo ambos modelos.

Ahora bien, para cada ítem i, se debe definir a priori un vector de pesos q_i = (q_{i1}, q..., q_{lk} .... q_{im}) en el cual k es la ocurrencia del atributo o componente en la respuesta. La matriz formada por todos los vectores de los ítems compone la matriz Q, comúnmente compuesta de unos ($q_{ik} = 1$) y ceros ($q_{ik} = 0$). Tal composición, se estructura en función de la presencia o ausencia de una determinada operación cognitiva en la resolución de un ítem. A partir de dicha matriz y de las respuestas de los sujetos, se estiman los valores α_k, que sirven para calcular el parámetro de dificultad δ de la formula (4).

Por otra parte, la aplicación del LLTM requiere de un modelo cognitivo que aporte un marco de referencia para la explicación de los requerimientos de procesamiento cognitivo de los ítems y de la variabilidad de la dificultad entre los mismos; de tal forma que la adecuación del LLTM se fundamenta en la plausibilidad y validez sustantiva del modelo cognitivo, representado formalmente en la matriz Q. La validación de la matriz Q es un requerimiento previo, sin el cual la estimación de los parámetros no tiene valor, por lo que las estimaciones de δ_1 deben contener la información valida que refleje fielmente el modo en que los individuos resuelven los ítems (Fisher, 1995). En definitiva, la pertinencia del LLTM depende estrechamente de la adecuación de la matriz Q, en referencia a la teoría cognitiva subyacente a los ítems de la prueba.

Propósito del estudio

El propósito principal del estudio fue analizar los procesos cognitivos que utilizan los estudiantes para resolver los ítems del área de Habilidades Cuantitativas (HC) del EXHCOBA y en base a ello, estructurar una matriz Q para aplicar el LLTM (Fischer, 1973). Como objetivo secundario, se propuso ilustrar un modelo metodológico con enfoque inductivo (*top-down)* para la obtención de evidencias de validez del proceso y de la estructura cognitiva, que puede ser muy útil para investigadores y profesionales en el campo del desarrollo de pruebas psicológicas y educativas.

Método

Se adaptó y aplicó un modelo metodológico tomando en cuenta dos propuestas teóricas: a) los estudios cognitivos de análisis por expertos y reportes verbales (Ericsson & Simon, 1993; Leighton & Gierl, 2007), que desde una perspectiva inductiva (Messick, 1989) o enfoque top-down (Chi, 1994; Ericsson & Simon, 1993; Leighton & Gierl, 2007) pueden permitir la identificación de las operaciones que subyacen a las puntuaciones de los ítems y b) el modelo para el desarrollo de test y evaluaciones cognitivas propuesto por Embretson (1994).

Muestra y recolección de datos

Es necesario aclarar para los términos del estudio que hay tres tipos de participantes. El primer tipo son aquellos que constituyen la muestra de 2801 graduados del nivel medio superior en México que contestaron la totalidad de ítems del EXHCOBA, de los cuales el 52% son hombres y el 48% son mujeres. Dichos participantes respondieron todos los ítems del EXHCOBA durante un proceso de selección real para el ingreso a la educación superior. El segundo tipo de participantes, fue constituido por ocho estudiantes de tercer grado de secundaria (cuatro hombres y

cuatro mujeres de alto y bajo rendimiento), a quienes se les aplicó los reportes verbales concurrentes y retrospectivos. Por su parte, el tercer tipo de participantes fue constituido por un grupo de expertos profesores del área de matemáticas y desarrolladores de pruebas psicológicas y educativas, quienes realizaron el análisis de los procesos de respuesta de los estudiantes ante los ítems de la prueba y la especificación de la estructura cognitiva del área HC del EXHCOBA.

Instrumento

Para propósitos del estudio, se seleccionaron los treinta ítems del área de HC del EXHCOBA, que es utilizado para la selección de estudiantes que desean ingresar a diferentes niveles educativos (bachillerato, licenciatura y posgrado) en México (Backhoff, Tirado, & Larrazolo, 2001). En la actualidad, el examen cuenta con seis versiones fijas en formato computarizado, que se encuentran integradas por 310 ítems de opción múltiple de los cuales se le aplican al estudiante de bachillerato 190 de ellos, dependiendo de la carrera a la que desean ingresar (véase Tabla 1).

Tabla 1. Estructura nomológica del EXHCOBA en su versión de ítems de opción múltiple

Nivel	Sección	Área	No. de ítems
Primaria	Habilidades básicas (60 ítems)	Habilidades verbales	30
		Habilidades cuantitativas	**30**
Secundaria	Conocimientos básicos (70 ítems)	Español	15
		Matemáticas	15
		Ciencias naturales	20
		Ciencias sociales	20
Media superior	Conocimientos básicos para la especialidad (60 ítems, 20 por cada una de las tres asignaturas según el tipo de examen)	Matemáticas para el cálculo	20
		Matemáticas para la estadística	20
		Física	20
		Química	20
		Biología	20
		Ciencias sociales	20
		Humanidades	20
		Lenguaje	20
		Cs. económico-administrativas	20
		Total de ítems del examen	**310**
		Total de ítems que contesta el estudiantes	**190**

Fuente: Adaptado de Backhoff, Tirado y Larrazolo (2001).“Ponderación diferencial para mejorar la validez de una prueba de ingreso a la universidad”. Revista Electrónica de Investigación Educativa 3(1). Consultado el 20 de febrero de 2013 en: http://redie.ens.uabc.mx/vol3no1/contenido-tirado.html.

Procedimiento

El estudio se desarrolló según el modelo metodológico adaptado en tres fases: 1) desarrollo, pilotaje y aplicación del estudio cognitivo; 2) análisis de los procesos de respuesta de los estudiantes ante los ítems de la prueba, para la especificación de la estructura cognitiva; y 3) aplicación y evaluación del modelo componencial.

Durante la primera fase, el trabajo se enfocó en conseguir la construcción de la teoría sustantiva que subyace a los procesos de respuesta de los treinta ítems del área HC del EXHCOBA.

Dicha meta, requirió de la aplicación de reportes verbales (Ericsson & Simon, 1993; Leighton & Gierl, 2007) tanto concurrentes como retrospectivos. En específico, cada estudiante fue sometido a la técnica concurrente de pensamiento en voz alta para capturar, a través de audio y video con ayuda del programa CANTASYA STUDIO 5, las verbalizaciones y el sendero del indicador del ratón en la pantalla de la computadora. Para la técnica retrospectiva, cada uno de los participantes respondió un protocolo de salida en donde se les preguntó sobre su proceso de respuesta.

Para la segunda fase del estudio, tres expertos del área de matemáticas (profesores de primaria, secundaria y licenciatura de la asignatura en matemáticas) y un experto en estudios cognitivos, con base en el análisis de los datos recopilados durante los reportes verbales, elaboraron la estructura cognitiva del área HC del EXHCOBA. Para el análisis cognitivo, los expertos partieron de una perspectiva inductiva (Messick, 1989) o enfoque top-down (Chi, 1997; Ericsson & Simon, 1993; Leighton & Gierl, 2007) lo que les permitió, sin necesidad de tener una estructura cognitiva a priori, la identificación de las operaciones que subyacen a las respuestas de los estudiantes.

En la tercera fase, se estudió la fiabilidad y dimensionalidad, la calibración y ajuste del modelo de Rasch, y la aplicación del modelo LLTM. En cuanto el análisis de la dimensionalidad, se utilizó el Modelo de Análisis Factorial de Fraser (1988). En dicho análisis, el ajuste del modelo se valoró por medio del índice de Tanaka (T), el cual fue obtenido por medio del programa NOHARM (Fraser, 1988). Por su parte, para la calibración, ajuste del modelo de Rasch y aplicación del modelo LLTM se utilizaron el programa libre R 2.15.1. (R Development Core Team, 2006) y el programa LPCM Win 1.0 de Fischer y Ponocny-Seliger (1998).

Resultados

Operaciones cognitivas

El análisis cognitivo realizado por los expertos sobre los procesos de respuesta de los estudiantes ante los ítems resultó en una estructura cognitiva de 14 atributos. Algunas de las operaciones cognitivas que requieren los estudiantes para responder a los ítems del área HC son: comprender problemas matemáticos contextualizados, emplear operaciones aritméticas y manejar expresiones algebraicas (véase Tabla 2).

Tabla 2. Atributos cognitivos del área de HC del EXHCOBA

Código	Atributos cognitivos	Definición de los atributos
AT6	Comprender problemas matemáticos contextualizados	Involucra la comprensión de problemas matemáticos planteados en lenguaje común y contextualizados
AT1	Emplear operaciones aritméticas	Representa las habilidades básicas de aritmética y sucesiones de números
AT2	Manejar sucesiones aritméticas	Involucra la habilidad para desarrollar sucesión de números reales.
AT3	Manejo de expresiones algebraicas	Involucra la habilidad para substituir los valores en expresiones algebraicas y solucionarlas
AT4	Comprender el sistema decimal (dígitos)	Incluye la comprensión básica de los nombres de los dígitos dentro del sistema decimal y su ubicación
AT5	Representar modelos exponenciales	involucra la habilidad para representar las reglas básicas de las operaciones exponenciales
AT7	Manejar fracciones	Involucra la habilidad de comprender y resolver fracciones y de su equivalencia con números decimales

Código	Atributos cognitivos	Definición de los atributos
AT8	Fraccionar y visualizar figuras geométricas	Considera la habilidad de fraccionar y visualizar figuras geométricas
AT9	Aplicar suma, resta y multiplicación de fracciones	Involucra la habilidad de aplicar la suma, resta, y multiplicación de fracciones
AT10	Aplicar la regla de tres simple	Incluye la aplicación del modelo aritmético de la regla de tres simple
AT11	Representar el modelo matemático del área	Involucra la aplicación de fórmulas para obtener el área de una figura geométrica
AT12	Manejar sucesiones aritméticas	Involucra el conocimiento de las conversiones básicas y su aplicación
AT13	Sumar los ángulos de un triangulo	Implica el conocimiento de la suma de los ángulos de un triangulo
AT14	Aplicar las reglas básicas de la probabilidad	Incluye la comprensión de las reglas básicas de probabilidad

Fuente: Adaptado de Pérez, J.C. "Análisis del aspecto sustantivo de la validez de constructo de una prueba de Habilidades Cuantitativas". Tesis doctoral, p. 26, Ensenada: Instituto de Investigación y Desarrollo Educativo, UABC, 2013.Fuente(s): Información adaptada de Apellido del autor, año de publicación.

Para ejemplificar el contraste entre los atributos cognitivos y los ítems del área HC, se puede ver el ejemplo del ítem número 20 (véase Figura 1), el cual requiere para su resolución, primero, comprender problemas matemáticos contextualizados (AT1), luego, emplear operaciones aritméticas (AT2) y, finalmente, representar el modelo matemático del área (AT11). La matriz Q (30X14) que une a cada ítem de la prueba con los atributos que subyacen a estos se presenta en la Tabla 3. Obsérvese, que el ítem 20 tiene un 1 en las casillas correspondientes a los atributos que teóricamente le subyacen y 0 en el resto de las casillas.

Fiabilidad y dimensionalidad

La fiabilidad de la prueba estimada mediante el coeficiente alpha de Cronbach arrojó un valor de 0,880. La media del índice de dificultad de los ítems (p) fue de 0,54 y la media de las correlaciones biseriales-puntuales (Rpbis) resultó ser igual a 0,61. En cuanto a la dimensionalidad del área HC del EXHCOBA, se puede observar en la última columna de la Tabla 3 los resultados del análisis del Modelo de Análisis Factorial de Fraser. Obsérvese que las cargas factoriales son todas positivas y grandes (el menor valor corresponde al ítem 6 con un valor de 0,426), indicando que todos los ítems contribuyen adecuadamente a la medida del rasgo latente. Además, se obtuvo un índice de Tanaka de 0.9864, lo cual confirma que los 30 ítems del test miden una sola dimensión. Al respecto, algunos investigadores, como Gierl, Tan y Wang (2005), mencionan que, aunque no hay normas de interpretación generalmente aceptadas para dicho índice, se ha propuesto que valores por arriba de 0,9 indican un buen ajuste (Rouse, Finger & Butcher, 1999), cumpliéndose así el supuesto de unidimensionalidad.

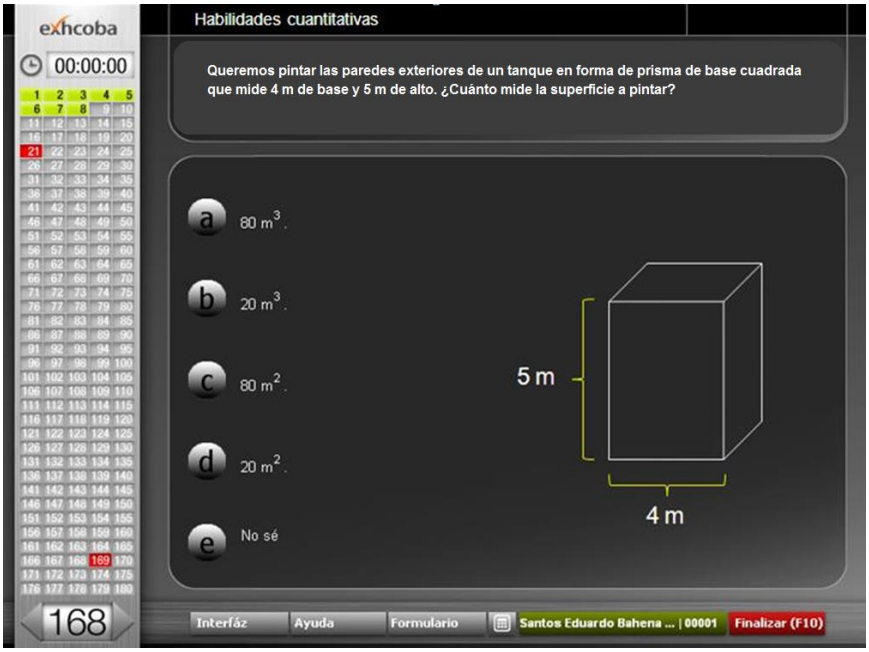

Figura 1. Presentación en pantalla del ítem 20 del área HC del EXHCOBA
Fuente: Elaboración propia a partir de la pantalla del interfaz EXHCOBA (Backhoff & Larrazolo, 2012).

Tabla 3. Matriz Q con 14 atributos y 30 ítems del área HC del EXHCOBA

	Atributos														
Ítem	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**Cargas factoriales**
01	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0.509
02	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0.642
03	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0.557
04	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0.617
05	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0.817
06	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0.426
07	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0.566
08	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0.659
09	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0.959
10	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0.630
11	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0.789
12	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0.946
13	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1.000
14	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0.945
15	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0.853
16	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0.587
17	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1.000
18	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0.703
19	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1.000

20	**1**	**1**	0	0	0	0	0	0	0	0	**1**	0	0	0	1.000
21	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0.736
22	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0.604
23	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0.691
24	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0.655
25	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0.774
26	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0.604
27	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.918
28	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0.904
29	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0.643
30	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0.756

Fuente: Adaptado de Pérez, J.C. "Análisis del aspecto sustantivo de la validez de constructo de una prueba de Habilidades Cuantitativas". Tesis doctoral, p. 152, Ensenada: Instituto de Investigación y Desarrollo Educativo, UABC, 2013.

Ajuste al modelo de Rasch y LLTM

En cuanto al ajuste del modelo LLTM para los 30 ítems del área HC, la correlación de Pearson entre las dificultades estimadas por los dos modelos es moderada (0,604 significativa al nivel .01) indicando un ajuste moderado del LLTM, esto se observa también en la gráfica correspondiente en la Figura 2. Por su parte, el estadístico CLR resultó significativo (gl=29; X^2=827.367), por lo tanto no se confirma el ajuste (véase Fischer, 1995). Esto no es sorprendente dado el rigor divulgado de la prueba CLR en la literatura (Fischer, 1995, p. 147). De tal forma que, aunque el ajuste visual y la correlación sean adecuados, es frecuente que el CLR resulte significativo. En la Tabla 4 se detalla, para cada uno de los 30 ítems, la estimación de los parámetros con el LLTM y con el modelo de Rasch. Así mismo, se presenta el error estándar de esta estimación y el valor Anderson-Darling X^2.

Tabla 4. Comparación de los parámetros de dificultad del modelo LLTM y del modelo de Rasch de los ítems de HC del EXHCOBA

No. de ítem	**RASCH**					**LLTM**	
	δi	**SE(δi)**	**X^2**	**p**		**δi**	**SE(δi)**
1	0.375	0.043	3208.126	0.000	*	0.021	0.026
2	1.156	0.047	2679.194	0.927		1.594	0.061
3	-0.55	0.043	3494.702	0.000	*	0.528	0.043
4	1.006	0.046	2741.596	0.727		0.840	0.035
5	-0.613	0.043	2874.761	0.121		0.840	0.035
6	2.021	0.058	3142.886	0.000	*	0.861	0.047
7	-0.225	0.042	3350.812	0.000	*	0.528	0.043
8	-0.524	0.043	3108.573	0.000	*	-0.031	0.057
9	-0.47	0.042	2323.536	1.000		0.457	0.047
10	1.101	0.047	3135.814	0.000	*	0.445	0.050
11	-0.234	0.042	2635.049	0.981		0.436	0.033
12	0.674	0.044	1834.346	1.000		0.928	0.045
13	-0.311	0.042	2220.251	1.000		0.928	0.045

No. de ítem	RASCH					LLTM	
	δi	SE(δi)	X^2	p		δi	SE(δi)
14	0.023	0.042	2255.841	1.000		0.497	0.057
15	0.593	0.044	2091.906	1.000		0.850	0.055
16	0.85	0.045	2791.835	0.471		0.861	0.047
17	-1.085	0.045	2610.022	0.992		0.861	0.047
18	-0.798	0.043	3190.346	0.000	*	-0.003	0.039
19	-1.55	0.048	2650.447	0.968		-0.628	0.045
20	-1.414	0.047	2797.163	0.442		-0.628	0.045
21	-0.038	0.042	2795.346	0.452		-0.124	0.056
22	-0.383	0.042	3363.689	0.000	*	0.105	0.057
23	0.271	0.043	2714.154	0.835		0.528	0.043
24	0.228	0.043	2843.678	0.223		-0.003	0.039
25	0.126	0.042	2541.894	1.000		0.433	0.062
26	-1.124	0.045	3753.892	0.000	*	-0.003	0.039
27	0.262	0.043	2030.827	1.000		0.728	0.057
28	-0.903	0.044	2851.752	0.192		0.265	0.047
29	1.736	0.053	2370.339	1.000		0.763	0.046
30	-0.199	0.042	2638.296	0.978		0.784	0.056

(*) $p<0.01$.

Fuente: Adaptado de Pérez, J.C. "Análisis del aspecto sustantivo de la validez de constructo de una prueba de Habilidades Cuantitativas". Tesis doctoral, pp.170-180, Ensenada: Instituto de Investigación y Desarrollo Educativo, UABC, 2013.

Aplicación del modelo LLTM

En la Tabla 5 pueden observarse las estimaciones de los parámetros básicos para cada atributo $(\boldsymbol{\alpha_k})$, sus respectivos errores estándares y valores mínimos y máximos obtenidos mediante el programa libre R 2.15.1. (R Development Core Team, 2006) y el programa LPCM Win 1.0 de Fischer y Ponocny-Seliger (1998). Los atributos que introducen mayor dificultad a los ítems del área HC son: manejo de sucesiones aritméticas (AT2), manejo de operaciones algebraicas (AT3), comprensión del sistema decimal con dígitos (AT4) y suma de ángulos de un triángulo (AT13). En particular, el AT4 no se esperaba como una fuente de dificultad alta. Sin embargo, dicho atributo juega un papel importante en la complejidad de uno de los ítems más difícil (ítem 29) de la prueba pero, además, forma parte de la estructura cognitiva de los ítems 3 y 7, los cuales tiene problemas de ajuste. Con respecto a los atributos cognitivos restantes, cinco de ellos contribuyen a la dificultad moderada de los ítems (parámetros básicos positivos) y cinco a la facilidad (parámetros básicos negativos).

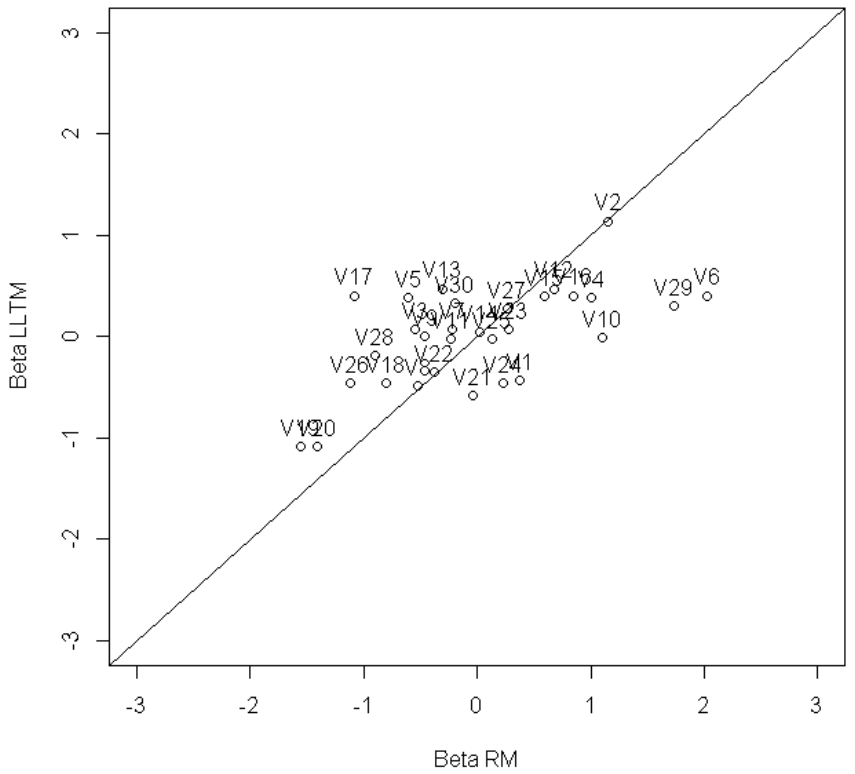

Figura 2. Ajuste gráfico del LLTM
Fuente: Elaboración propia a partir de la base de datos del EXHCOBA (Backhoff & Larrazolo, 2012).

Tabla 5. Parámetros básicos del LLTM

Atributos (AT)	α_k	$SE\ (\alpha_k)$	Min.	Max.
1	0.021	0.026	-0.030	0.071
2	1.574	0.057	1.462	1.685
3	0.519	0.036	0.448	0.590
4	0.840	0.035	0.772	0.909
5	-0.019	0.063	-0.142	0.104
6	-0.012	0.031	-0.072	0.049
7	0.436	0.033	0.372	0.500
8	0.492	0.036	0.422	0.563
9	-0.431	0.051	-0.530	-0.332
10	-0.012	0.037	-0.084	0.059
11	-0.637	0.041	-0.718	-0.556
12	0.096	0.055	-0.011	0.203

Atributos (AT)	α_k	$SE\ (\alpha_k)$	Min.	Max.
13	0.728	0.057	0.615	0.840
14	0.256	0.034	0.189	0.322

Fuente: Adaptado de Pérez, J.C. "Análisis del aspecto sustantivo de la validez de constructo de una prueba de Habilidades Cuantitativas". Tesis doctoral, pp.183, Ensenada: Instituto de Investigación y Desarrollo Educativo, UABC, 2013.

Conclusiones

En el presente estudio se analizaron los procesos cognitivos que utilizan los estudiantes para resolver los ítems del área HC del EXHCOBA y con base en ello se estructuró una matriz Q para aplicar el LLTM (Fischer, 1973). De manera particular, se analizaron las operaciones cognitivas activadas por los estudiantes para responder los ítems de la prueba a través de reportes verbales y el análisis de expertos (Ericsson & Simon, 1993; Leighton & Gierl, 2007) desde una perspectiva inductiva (Messick, 1989). Además, se analizaron las fuentes de dificultad de los ítems mediante la aplicación del modelo LLTM en el cual, según Prieto y Delgado (1999), la dificultad del ítem se considera un indicador de la complejidad cognitiva requerida para resolver correctamente los reactivos.

Por su parte, el uso de los reportes verbales y luego el análisis de expertos para elaborar la estructura cognitiva, que representara lo más fiel posible los atributos que utilizan los estudiantes para contestar los ítems de la prueba, resultó una estrategia con muchas ventajas. Una de las principales, fue el mapeo a profundidad de los procesos utilizados por el estudiante para responder los ítems y la facilidad para utilizar dichos datos en la elaboración de la matriz Q.

En lo que respecta a la aplicación del LLTM, se puede decir que dicho modelo permitió conocer la contribución de las operaciones cognitivas estructuradas en la matriz Q a la dificultad de los ítems (Romero, Ponsoda & Ximénez, 2006). También se encontró que un refinamiento de los catorce atributos propuestos, mediante los reportes verbales y el análisis de expertos, es necesario para explicar mejor el rendimiento de los estudiantes en la prueba. Lo anterior indica la necesidad de una mayor elaboración del modelo cognitivo y la consideración de estrategias alternativas de respuesta.

Por otro lado, el modelo LLTM también presentó algunas limitaciones: a) problemas de adecuación ante una prueba que se identificó por medio de los reportes verbales y el análisis de expertos que mide más de un rasgo latente, b) la dificultad para cumplir con el requisito previo de unidimensionalidad, c) la falta de una asíntota inferior y de un parámetro de discriminación a partir del contenido del ítem (véase Revuelta & Ponsoda, 1998), y d) el carácter compensatorio lineal de la descomposición de los parámetros de dificultad (véase Romero, Ponsoda & Ximénez, 2006). En conclusión, en los resultados y durante el análisis de la estructura cognitiva de la prueba HC del EXHCOBA, se muestran las principales técnicas empleadas en el proceso, así como las ventajas y limitaciones del LLTM como modelo componencial.

REFERENCIAS

Abida, K., Azeem, M. & Bashir, M. (s.f.). Assessing Students' Math Proficiency Using Multiple-choice and Short Constructed Response Item Formats. *International Journal of Technology, Knowledge and Society, 7*, 135-150.

Adams, R., Wilson, M., & Wang, W-C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement, 21*, 1-24.

Backhoff, E. & Larrazolo, N. "Evaluación de competencias escolares". Trabajo presentado en el I Foro Iberoamericano de Evaluación Educativa de la Red Iberoamericana de Medición y Evaluación de Sistemas Educativos. 7 de noviembre de 2012, baja California. Consultado en: http://www.ustream.tv/recorded/26785327

Backhoff, E., Tirado, F. & Larrazolo, N. (2001). Ponderación diferencial de reactivos para mejorar la validez de una prueba de ingreso a la universidad. *Revista Electrónica de Investigación Educativa, 3*(1). Consultado el 20 de febrero de 2013 en: http://redie.ens.uabc.mx/vol3no1/contenido-tirado.html

Chi, M. (1994). *Analyzing the content of verbal data to represent knowledge: A practical guide.* Unpublished manuscript, Learning Research and Development Center, University of Pittsburgh.

Cortada de Kohan, N. (2000). Importancia de la investigación psicométrica. *Revista Latinoamericana de Psicología, 34*(3), 229-240.

DiBello, L., Stout, W., & Roussos, L. (1995). Unified cognitive/psychometric diagnostic assessment likelihood-based classification techniques. En P. D. Nichols, S. F. Chipman & R. L. Brennan (Eds.), *Cognitively diagnostic assessment,* (pp. 361-389). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Embretson, S. (1984). A general latent trait model for response processes. *Psychometrika, 49*(2), 175-186.

— (1994). Applications of cognitive design systems to test development. En C.R. Reynolds (Ed.), *Cognitive assessment. A multidisciplinary perspective* (pp. 107-135). New York: Plenum.

Embretson, S. E., & Wetzel, C. D. (1987). Component latent trait models for paragraph comprehension tests. *Applied Psychological Measurement, 11*(2), 175-193.

Ericsson, K., & Simon, H. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: The MIT Press.

Fischer, G. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica, 37*, 359-374.

— (1995). The linear logistic test model. En G. H. Fisher, & I. W. Molenaar (Eds.), *Rasch models: Foundations, recent developments, and applications*. New York: Springer-Verlag.

Fischer, G. & Molenaar, I. (1995). *Rasch Models. Foundations, Recent Developments and Applications*. Ed: Springer-Verlag. New York.

Fischer G. & Ponocny-Seliger, E. (1997). LPCM-WIN Program. IEE ProGAMMA. Groningen.

Fraser, C. (1988). *NOHARM: Computer software and manual.* Australia: Author.

Gierl, M., Tan, X. & Wang, Ch. (2005*). Identifying content and cognitive dimensions on the SAT*. Research Report nº 2005-11. College Board.

Junker, B., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement, 25*(3), 258-272.

Leighton, J., & Gierl, M. (2007). Verbal reports as data for cognitive diagnostic assessment. En J. P. Leighton y M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education. Theory and applications,* (pp. 146–172). Cambridge, MA: Cambridge University Press.

Messick, S. (1989). Validity. En R. L. Linn (Ed.), *Educational measurement (3rd ed.),* (pp. 13-103). New York, NY, England: Macmillan Publishing Co, Inc.

Mislevy, R. (1995). Probability-based inference in cognitive diagnosis. En P. D. Nichols, S. F. Chipman y R. L. Brennan (Eds.), *Cognitively diagnostic assessment*, (pp. 43-71). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Pérez, J.C. (2013). Análisis del aspecto sustantivo de la validez de constructo de una prueba de Habilidades Cuantitativas. Tesis doctoral. Ensenada, Instituto de Investigación y Desarrollo Educativo, UABC, 2013.

Prieto, G & Delgado, A. (1999). Medición cognitiva de las aptitudes. En J. Olea, V. Ponsoda & G. Prieto (Eds.), *Test informatizados: Fundamentos y aplicaciones*. Madrid: Pirámide.

Revuelta, J. & Ponsoda, V. (1998). Un test adaptativo informatizado de análisis lógico basado en la generación automática de ítems. *Psicothema, 10*, 753-760.

Romero, S. Propiedades y aplicaciones del método de las distancias mínimo-cuadráticas (LSDM) para la validación y análisis de atributos cognitivos. Tesis, Universidad Europea de Madrid, 2010.

Romero, S.J., Ponsoda, V. & Ximénez, C. (2006). Validación de la estructura cognitiva del test de signos mediante modelos de ecuaciones estructurales. *Psicothema, 18*, 835-840.

Rouse, S.V., Finger, M.S. & Butcher, J.N. (1999). Advances in clinical personality measurement: An item response theory analysis of the MMPI – PSY- 5 scales. *Journal of Personality Assessment, 72*, 282-307.

Rupp, A. & Mislevy, R. (2006). Cognitive Foundations of Structured Item Response Models. En J.P. Leighton y M.J. Gierl (Eds.), *Cognitive Diagnostic Assessment: Theories and Applications.* Cambridge: Cambridge University Press.

Spada, H. (1977). Logistic models of learning and thought. En H. Spada, & W. F. Kempf (Eds.), *Structural models of thinking and learning*, (pp. 227-262).

Spada, H. & May, R. (1982). The linear logistic test model and its application in educational research. En D. Spearrit (Ed.), *The improvement of measurement in Education and Psychology*, (pp. 67-84). Hawthorne, Victoria: Australian council for educational research.

Tatsuoka, K. (1985). A probabilistic model for diagnosing misconceptions by the pattern classification approach. *Journal of Educational Statistics*, *10*(1), 55-73.

Van der Linden, W. & Hambleton, R. (1997). *Handbook of modern item response theory*. New York: Springer.

SOBRE LOS AUTORES

Juan Carlos Pérez Morán: Director de área de la Junta de Gobierno del Instituto Nacional para la Evaluación de la Educación (INEE); Doctor en Ciencias Educativas (DCE) y Maestro en Ciencias Educativas (MCE) por el Instituto de Investigación y Desarrollo Educativo (IIDE) de la Universidad Autónoma de Baja California (UABC); Investigador adjunto del proyecto Examen de Habilidades y Conocimientos Básicos (EXHCOBA); licenciado en Psicología por la Facultad de Medicina y Psicología (FMP) de la UABC; profesor de asignaturas en psicología de la educación, metodología de la investigación en psicología, revisión de artículos científicos y psicometría a nivel licenciatura y posgrado; jefe de la academia técnico-metodológica de la FMP y; asesor psicológico y educativo privado.

Norma Larrazolo Reyna: Investigadora y docente en el IIDE de la UABC. Trabajos de investigación en el área de evaluación educativa, docente de la Maestría de Ciencias Educativas, coordinadora de EXHCOBA y EXEDII, ambos proyectos de financiamiento externo.

Eduardo Backhoff Escudero: Consejo de la Junta de Gobierno del Instituto Nacional para la Evaluación Educativa (INEE). Doctor en Educación y miembro del Sistema Nacional de Inves-

tigadores de México. Entre los cargos ocupados, destacan los siguientes: profesor investigador de la Universidad Autónoma de Baja California, profesor de la Universidad Nacional Autónoma de México (UNAM), director del Instituto de Investigación y Desarrollo Educativo (UABC); director de Pruebas y Medición del Instituto Nacional para la Evaluación de la Educación (México); Editor Científico de la Revista Electrónica de Investigación Educativa (REDIE); miembro asesor de los proyectos PISA y TALIS (OCDE). Ha impartido más de 70 cursos de licenciatura y posgrado. Ha publicado cerca de 70 artículos de investigación, 15 capítulos de libro y es autor (o coautor) de 14 libros. Ha presentado más de 90 ponencias en congresos nacionales e internacionales. Es miembro de comités editoriales de cinco revistas científicas. Su campo de interés es el desarrollo y validación de pruebas de aprendizaje de gran escala, y la evaluación asistida por computadora. Sus proyectos más recientes de investigación tienen que ver con los problemas de traducción de las pruebas nacionales e internacionales.

Guaner Rojas: Investigador y docente en el Instituto de Investigaciones Psicológicas de la Universidad de Costa Rica. Obtuvo su Doctorado y Máster Metodología de las Ciencias del Comportamiento y de la Salud en la Universidad Autónoma de Madrid y la Universidad Complutense de Madrid, respectivamente. Su Licenciatura en Enseñanza de la Matemática por la Universidad de Costa Rica. La investigación de Guaner Rojas se enfoca en el modelado de procesos de respuesta a ítems con los modelos de diagnóstico cognitivo. Sus intereses de investigación se centran en investigación metodológica: psicometría, medición y evaluación con pruebas psicoeducativas.

La ***Revista Internacional de Educación y Aprendizaje*** pretende promover la investigación, invitar al diálogo y construir un conjunto de conocimientos sobre la naturaleza y el futuro del aprendizaje. La revista proporciona un foro para cualquier persona interesada por la educación en todas sus etapas y formas, desde la infancia, los colegios, la enseñanza superior y el aprendizaje continuo a lo largo de toda la vida, así como en todos de sus ámbitos, desde el hogar, la escuela, la universidad o el lugar de trabajo.

La revista resulta de interés para académicos, investigadores, profesores, estudiantes de educación superior y educadores, así como administradores y gestores de educación.

La ***Revista Internacional de Enseñanza y Aprendizaje*** es una revista académica sujeta a revisión por pares., y acepta artículos escritos en español y portugués.

ISSN 2255-453X