

FICHA TECNICA PARA PRESENTACIÓN DE DOCUMENTOS

1ER CONGRESO LATINAMERICANO DE MEDICIÓN Y EVALUACIÓN EDUCACIONAL (COLMEE)

Centro de Medición MIDE UC, Pontificia Universidad Católica de Chile

Marzo, 2012

REFERENCIA

Moreira-Mora, T.E. (marzo 2012). Evidencias de constructo de la Prueba de Aptitud Académica del Instituto Tecnológico de Costa Rica. Presentado en 1er Congreso Latinoamericano de Medición y Evaluación Educacional (COLMEE), Santiago, Chile.

INFORMACIÓN CONTACTO

TITULO	Evidencias de constructo de la Prueba de Aptitud Académica del Instituto Tecnológico de Costa Rica
AUTORES	Tania Elena Moreira Mora
DIRECCION AUTOR PRINCIPAL (calle, región, ciudad, país)	Centroamérica, Costa Rica, San José, Tibás.
INSTITUCIÓN (AUTOR PRINCIPAL)	Instituto Tecnológico de Costa Rica
CORREO ELECTRÓNICO (AUTOR PRINCIPAL)	tmoreira.costarica@gmail.com
TELEFONO (AUTOR PRINCIPAL)	(506) 25 50 26 75

INFORMACIÓN BIBLIOGRÁFICA

REVISTA/MEDIO DE PUBLICACIÓN	Inédito.
FECHA DE PUBLICACIÓN	Inédito.
EJE TEMÁTICO	Validación de pruebas educativas.
PREGUNTA/PROBLEMA QUE RESUELVE	¿Cuál es el grado de asociación de modelo de habilidades matemáticas y verbales de la Prueba de Aptitud Académica con la habilidad de razonamiento general?
REFERENCIAS BIBLIOGRÁFICAS	American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999). Standards for educational and psychological testing. Washington, Estados Unidos de América: Autor.
	Batista, J.M., Coenders, G. y Alonso, J. (2004). Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud. Revista Medicina Clínica, 122(1), 21-27. Recuperado de http://www.elsevier.es/es/node/2184538
	Cea, M.A. (2002). Análisis multivariable. Teoría y práctica en la investigación social. Madrid, España: Editorial Síntesis

	<p>Gierl, M.J., Tan, X. & Wang, Ch. (2005). Identifying content and cognitive dimensions on the SAT. College Board Research Report N° 2005-11. Recuperado 2 de marzo del 2009 de https://professionals.collegeboard.com/profdownload/pdf/CBR0511.pdf</p>
	<p>Kobrin, J.L. & Kimmel, E.W (2006). Test development and technical information on the writing section of the SAT reasoning test. College Board. Recuperado el 8 de enero del 2009 de http://www.clepprepcenter.net/research/pdf/RN-25.pdf</p>
	<p>Lawrence, I.M., Rigol, G.W., Van Essen, T. & Jackson, C.A. (2003). A historical perspective on the content of the SAT. College Board. Recuperado el 15 de enero del 2009 de http://professionals.collegeboard.com/profdownload/pdf/031306researchreport2_23504.pdf</p>
	<p>Messick, Samuel (1995). Standards of validity and the validity of standards in performance assessment. Educational Measurement: Issues and practice, 14 (2), pp. 5-24. Boston, Estados Unidos de América: Blackwell Publising.</p>
	<p>Milewski, G.B, Johnsen, D., Glazer, N. & Kubota, M. (2005). A survey to evaluate the Alignment of the New SAT®. Writing and Critical Reading Sections to Curricula and Instructional Practices (Research report N° 2005-1). College Board. Recuperado el 27 de abril del 2009 de http://professionals.collegeboard.com/profdownload/pdf/05792RDCBReport05-1_050420.pdf</p>
	<p>Pardo, A. y Ruiz, M.A. (2002). SPSS 11. Guía para el análisis de datos. Madrid, España: McGraw Hill.</p>
	<p>Young, J. W. (2001). Differential validity, differential prediction, and college admission testing: A comprehensive review and analysis (Reporte N° 2001-6). College Board. Recuperado el 16 de enero del 2009 de http://professionals.collegeboard.com/profdownload/pdf/differential_validity_10539.pdf</p>

DESCRIPCIÓN DOCUMENTO (4 PÁGINAS MÁXIMO)

Introducción

Con este estudio se busca obtener evidencias de la relación de cada una de las habilidades matemáticas y verbales, medidas en la prueba de aptitud académica (PAA) del Instituto del Tecnológico de Costa Rica (TEC), con su constructo y responder a la interrogante de ¿cuál es el grado de asociación de modelo de habilidades matemáticas y verbales de la Prueba de Aptitud Académica con la habilidad de razonamiento general?

Desde la perspectiva del modelo clásico se pueden presentar evidencias relacionadas a la validez de contenido, de constructo y de criterio, entre otras. Cuanto más evidencias de validez se obtengan de una prueba, más sólido es el diseño y más apropiados el uso de sus resultados. Además, desde la perspectiva unificada propuesta por Messick, sólo existen diferentes formas para encontrar evidencias de validez. Según este autor, la validez encierra "...un juicio integrado y evaluativo del grado en que la evidencia empírica y las razones teóricas apoyan lo adecuado y lo apropiado de las interpretaciones y las acciones basadas en las puntuaciones de las pruebas u otras formas de evaluación" (Messick, 1995, p. 5). Es más, conforme con los estándares para las pruebas educativas y psicológicas establecidos por la American Educational Research Association (AERA), la American Psychological Association (APA) y la National Council on Measurement in Education (NCME), el proceso de validación involucra la acumulación de evidencia que proporciona una base científica para la interpretación de las puntuaciones de la prueba y la relevancia (1999).

En general, los estudios concernientes a las PAA se han centrado en dos vertientes. Una es la psicométrica fundamentada en los estándares para pruebas psicológicas y educativas establecidas por la AERA, la APA y el NCME que competen a la calificación, confiabilidad y validez, procesos administrativos y de construcción (Kobrin & Schmidt, 2005; Lawrence, Rigor, Van Essen & Jackson, 2003; Young, 2001). La segunda está más orientada hacia los modelos cognitivos y la conceptualización del constructo, las habilidades o dimensiones cognoscitivas de estas pruebas (Gierl, Tan & Wang, 2005; Milewski, Johnsen, Glazer & Kubota, 2005).

En congruencia con esta segunda vertiente, se busca especificar un modelo básico de habilidades y aptitudes académicas, acorde con el perfil del estudiantado de primer ingreso del TEC, mediante una tabla de especificaciones para garantizar una muestra representativa de las habilidades verbales y matemáticas. Este modelo propuesto puede ser conceptualizado como un proceso científico de contrastación de hipótesis que permite contrastar la validez ajustando un modelo que lo asuma y diagnosticando su bondad de ajuste (Batista, Coenders y Alonso, 2004). Precisamente, en este estudio se pretende contrastar las habilidades verbales y matemáticas hipotetizadas en este modelo de medición. De acuerdo con Cea (2002), este contraste se puede generar a partir de una teoría, de generalizaciones empíricas o de la estructura latente que el investigador espera encontrar en los datos.

Objetivos

1. Determinar el grado de asociación del modelo de habilidades matemáticas y verbales con el constructo de razonamiento general medido en la Prueba de Aptitud Académica del Tecnológico de Costa Rica.
2. Comprobar el grado de ajuste del modelo de habilidades matemáticas y verbales mediante un análisis factorial confirmatorio.

Método

El enfoque es cuantitativo por el interés de comprobar el grado de ajuste del modelo de habilidades matemáticas y verbales de la PAA con el razonamiento general, por tal razón, este es un estudio transversal y descriptivo explicativo.

La primera fase de esta investigación fue descriptiva y se focalizó en la construcción de la matriz de las especificaciones (2009) con base en los datos recolectados mediante la técnica de encuesta, aplicada a la población de docentes y a una muestra por conglomerados del estudiantado que estaba matriculado (550), al menos, en el cuarto nivel del plan de estudios en el TEC, cuyos resultados se resumen en la siguiente tabla.

Tabla 1

Distribución porcentual de las habilidades matemáticas y verbales, según consulta a docentes y estudiantes del Tecnológico de Costa Rica. 2009

Habilidades matemáticas	Consulta	Habilidades verbales	Consulta
HM1: Aplicación de operaciones básicas	91,5%	HV1: Comprensión de lectura en textos cortos	84,8%
HM2: Razonamiento deductivo/inductivo	90,9%	HV2: Comprensión de lectura en textos largos	84,6%
HM3: Interpretación de gráficas	89,2%	HV3: Razonamiento con analogías y silogismos	84,0%
HM4: Resolución de problemas	88,3%	HV4: Comprensión de léxico	81,2%
HM5: Razonamiento con figuras	87,5%	HV5: Comprensión del contenido, la estructura y función de oraciones	78,9%
HM6: Razonamiento probabilístico	84,6%	HV6: Identificación de estilos discursivos	77,3%

En la segunda parte de esta investigación, se realiza el análisis multivariado, utilizando el paquete estadístico LISREL (Linear Structural Relations), versión 8.54, conforme el siguiente procedimiento:

1. Identificar del modelo: La regla de los grados de libertad (gl) es la más frecuente y se calcula con la fórmula $k * (k+1)/2$. Según el número de indicadores matemáticos (6) y verbales (4) incluidos en la ecuación, se tiene un total de 55 elementos conocidos en la matriz de covariancia y se ha especificado un total de 10 parámetros. Al restar este número, el modelo queda con 45 grados de libertad; por lo tanto, es un modelo sobre identificado.
2. Evaluar la identificación del modelo: Una recomendación para tratar que todos los parámetros estén identificados es utilizar al menos tres indicadores por variable latente e igualar la métrica de cada variable latente con uno de sus indicadores, lo que se consigue fijando arbitrariamente al valor 1 el peso de uno de los indicadores (Cea, 2002). En el caso de este modelo, se incluye 10 indicadores para el constructo de PAA y a la habilidad HM1 se le asignó el peso de 1.

3. Seleccionar las medidas, basado en la recolección, descripción y depuración de datos: Las mediciones se basaron en las puntuaciones de los ítems de la PAA del 2010, las cuales se agruparon según la categorización de las habilidades matemáticas y verbales.
4. Estimar el modelo con el método de máxima verosimilitud: Es la técnica más empleada en los programas informáticos, entre estos, el LISREL.
5. Comprobar el ajuste del modelo con los siguientes índices:
 - Goodness of Fit Index (GFI) = 0,94. Con este valor se comprueba que el ajuste global del modelo es aceptable, al superar el valor de referencia 0,90.
 - Normed Fit Index (NFI) = 0.95. Este índice de ajuste compara el modelo propuesto con el modelo especificado de las habilidades y el ajuste es aceptable al superar el valor recomendado de 0,90.
 - Adjusted Goodness of Fit Index (AGFI) = 0.91. Este índice de ajuste de parsimonia, se ajusta al valor deseado de 0,90.

Una vez concluida esta etapa de la estimación del modelo AFC, se interpretan las relaciones de los indicadores con la variable latente medida en la PAA.

Resultados

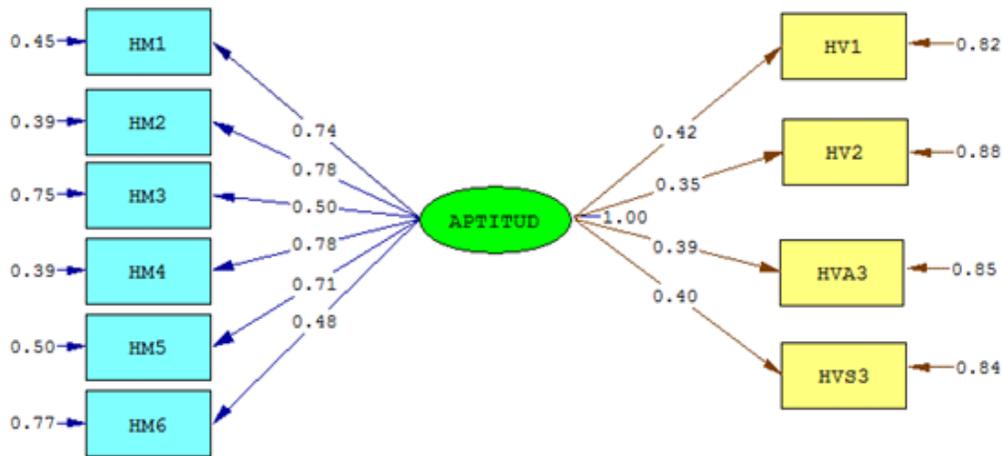
En general, los modelos de ecuaciones estructurales (Structural Equation Modeling, SEM) buscan ajustar las covarianzas entre las variables y minimizar la diferencia entre las covarianzas observadas de la muestra y las covarianzas estimadas por el modelo estructural. Por esta razón es que se les llama modelos de estructura de covarianza (Pardo y Ruiz, 2002). En cuanto a las correlaciones múltiples cuadradas para la ecuación estructural, la variancia explicada por los indicadores matemáticos y verbales incluidos en este modelo fueron:

HM1	HM2	HM3	HM4	HM5	HM6	HV1	HV2	HVS3	HVA3
0.55	0.61	0.25	0.61	0.50	0.23	0.18	0.12	0.16	0.15

De acuerdo con estos datos, las habilidades matemáticas que más explican la variancia en el constructo de PAA son razonamiento deductivo e inductivo (HM2) y la resolución de problemas (HM4), en tanto que, la comprensión de lectura en textos cortos (HV1) es el indicador verbal que más aporta al modelo que, en general, tales habilidades fueron bajas. Estos datos se complementan con el gráfico de la solución estandarizada, donde se muestran el grado de asociación y los errores de medición.

Con el diagrama de sendas se facilita la interpretación de los resultados al basarse en unidades estandarizadas, que favorecen la comparabilidad entre los coeficientes de correlación. Según el análisis las habilidades matemáticas con el mayor grado de asociación a la variable latente exógena fueron razonamiento deductivo/inductivo (HM2) y Resolución de problemas (HM4), ambas con un coeficiente igual a 0,78; en tanto, que en las habilidades verbales tienden a disminuir, siendo el indicador HV1, comprensión de lectura en textos cortos, el más alto (0,42). Este modelo incluye los indicadores observables, los errores que afectan a las mediciones y las relaciones entre el constructo, específicamente, se tiene:

- Variables observadas exógenas (λ): HM1, HM2, HM3...
- Variable latente exógena (ξ): Aptitud de razonamiento
- Errores de medición en variables observadas exógenas (δ): De 0,39 a 0,88
- Coeficientes de correlación entre la variable latente con sus respectivos indicadores (λ): Los valores oscilan entre 0,35 y 0,78.



Chi-Square=3711.42, df=35, P-value=0.00000, RMSEA=0.094

Este modelo de ecuaciones estructurales proporcionó evidencias de la existencia de factores comunes que explicaron la variabilidad de un conjunto indicadores (HMx y HVx) y la de un factor general (aptitud) que explica la variabilidad de estos factores comunes.

Conclusiones

La validación debe ser un proceso permanente y cuanto más evidencias de validez se obtengan de una prueba, más sólido es el diseño y más razonable los usos y las interpretaciones de sus resultados, especialmente, por sus altas consecuencias sociales. Por ello, se debe disponer de diversas fuentes de evidencia de validez, como la proporcionada por la tabla de especificaciones concernientes con el dominio del contenido que, en conjunto con las evidencias del AFC, comprobó el grado de asociación entre las variables observadas exógenas (x) de las habilidades matemáticas y verbales con el constructo de la habilidad de razonamiento, variable latente exógena (ξ). Por lo demás, mostró un buen ajuste con el índice de RMSEA (inferior a 0,10), un ajuste global del modelo propuesto (GFI), de ajuste incremental (NFI) y de ajuste de parsimonia (AGFI). En suma, esta investigación inició con la propuesta de un modelo de indicadores matemáticos y verbales, sustentado en una tabla de especificaciones, medidos en un conjunto de ítems en la PAA, y concluyó con la estimación del AFC. Con estos resultados, se ha

comprobado que el modelo propuesto es uno de los posibles modelos que se ajustan, controlando los efectos del error de medida sobre los coeficientes estructurales, a la medición de la habilidad de razonamiento de la PAA de esta Universidad.