

Validación de un modelo de cuatro estrategias de resolución de ítems de razonamiento en una prueba estandarizada de selección*

Validation of a Four-Strategies Model for Solving Reasoning Items in a Standardized Selection Test

Karol Jiménez Alfaro, Guaner Rojas Rojas, Armel Brizuela Rodríguez
y Nelson Pérez Rojas
Universidad de Costa Rica, Costa Rica

Resumen

La premisa de este estudio consiste en que un modelo cognitivo permite a las personas mejorar el desempeño en la resolución de preguntas de una prueba estandarizada donde el uso de estrategias tiene un papel crucial. El propósito de la investigación fue validar un modelo cognitivo con cuatro estrategias definidas por jueces expertos con base en los procesos de respuesta que subyacen a ítems de la Prueba de Aptitud Académica de la Universidad de Costa Rica (UCR). Se realizaron ocho entrevistas semiestructuradas a estudiantes de la UCR, quienes estaban cursando el primer año en esta institución y se aplicaron las técnicas de reporte verbal para obtener evidencias de los procesos de respuesta de los ítems. Se analizaron los reportes para comprobar la correspondencia entre el marco de las estrategias definidas previamente por jueces expertos y las respuestas dadas por los participantes. Los resultados indicaron que los participantes siguieron las estrategias propuestas para la solución de las situaciones planteadas por los jueces; por tanto, los ítems son indicadores de los procesos subyacentes a estas estrategias. Estos hallazgos abren la posibilidad de implementar investigaciones con atributos presentes en cada estrategia propuesta de este estudio, que posibiliten predecir las puntuaciones de la prueba en el rendimiento académico en la UCR.

Palabras clave: prueba de admisión, procesos de respuesta, reporte verbal, razonamiento en contexto matemático, modelo cognitivo

*Esta investigación fue financiada por la Universidad de Costa Rica mediante la Vicerrectoría de Investigación y el Programa de la Prueba de Aptitud Académica.

Karol Jiménez Alfaro, Guaner Rojas Rojas, Armel Brizuela Rodríguez y Nelson Pérez Rojas; investigadores en el Programa de la Prueba de Aptitud Académica de la Universidad de Costa Rica, Costa Rica.

La correspondencia en relación con este artículo se dirige a Karol Jiménez Alfaro, investigadora en el Programa de la Prueba de Aptitud Académica de la Universidad de Costa Rica, Costa Rica. E-mail: karol.jimenez@ucr.ac.cr

K. JIMÉNEZ, G. ROJAS, A. BRISUELA, N. PÉREZ

Abstract

The premise of this study is that a cognitive model can increase students' performance for solving items on a standardized test, where strategies play a crucial role. The purpose of the study is to validate a cognitive model with four strategies defined by expert judges based on the response processes that underlie the items of the University of Costa Rica (UCR) Academic Aptitude Test. Eight semi-structured interviews were conducted with first-year students at UCR and the techniques of verbal reporting were applied to gather evidence of the items' response processes. The reports were analyzed to verify the correspondence between the framework of the strategies previously defined by expert judges and the answers given by the participants. In light of the results, it was concluded that the participants followed the proposed strategies for solving the situations posed, and, therefore, the items are indicators of the processes underlying these strategies. The results open the possibility of implementing research with attributes present in each strategy proposed in this study, which will allow predicting test scores in academic performance at the UCR.

Keywords: admission test, response processes, verbal report, reasoning in mathematical context, cognitive model

La estructura curricular en la secundaria costarricense (tercer ciclo y ciclo diversificado) en el área de matemática se focaliza en la búsqueda del fortalecimiento de mayores capacidades cognoscitivas en el estudiantado para abordar los retos de una sociedad moderna, que demanda en las personas la necesidad de altas habilidades de razonamiento y capacidades mentales (Ministerio de Educación Pública, 2012). El empleo y el fortalecimiento de estas habilidades implican un desafío a los estudiantes en la resolución de situaciones educativas complejas, o situaciones problema en el marco del desarrollo del pensamiento y actitudes para enfrentar los retos de la ciencia, la tecnología y la vida.

En Costa Rica, una vez que las personas concluyen el ciclo diversificado de educación secundaria y desean continuar con sus estudios en una universidad pública, se encuentran con diversas opciones de educación superior. Quienes desean ingresar a la Universidad de Costa Rica (UCR) o a la Universidad Nacional de Costa Rica (UNA), deben obtener un promedio de admisión cuyo resultado es la combinación del promedio de las notas de aprovechamiento en el ciclo diversificado y del resultado obtenido en la Prueba de Aptitud Académica (PAA). La PAA evalúa habilidades de razonamiento general en contextos verbales y matemáticos. El razonamiento se considera como la capacidad de las personas para establecer relaciones entre ideas y obtener conclusiones; es decir, se trata del proceso de elaboración de conclusiones, que permite informar sobre los esfuerzos realizados por la persona en la resolución de un problema y en la toma de decisiones que le permite llegar a un objetivo (Leighton & Sternberg, 2004).

En la PAA, las habilidades de razonamiento se evalúan con el planteamiento de situaciones construidas a partir de las experiencias en las que cotidianamente podrían estar implicadas las personas. Las situaciones incorporadas en la PAA se centran en experiencias educativas, personales, de trabajo y en circunstancias imaginarias. Además, el planteamiento de las situaciones se combina con las nociones de conceptos básicos del currículo matemático de la educación primaria y secundaria.

Martínez, Hernández y Hernández (2006) plantean que en la evaluación educativa se logra observar lo que las personas dicen o hacen en circunstancias particulares y se intenta inferir lo que conocen

o son capaces de hacer en general. En este sentido, la puntuación obtenida por una persona en la PAA podría considerarse como una representación parcial de las capacidades cognoscitivas para abordar los retos de una sociedad moderna (Juan-Espinosa, 1997). En varios estudios (Jiménez & Morales, 2009-2010; Molina, Gallardo, & Cordero, 2011; Montero, Villalobos, & Valverde, 2007; Rojas, 2013) se han obtenido evidencias sobre el poder predictivo de la Prueba, los cuales indican que la nota en la PAA y el promedio de presentación, entendido como porcentaje de respuestas correctas y la puntuación estandarizada con media de 500 y desviación de 100, respectivamente, predicen por igual, y de manera efectiva, el logro académico de las personas que ingresan a la UCR.

Un estudio previo de Brizuela, Jiménez, Pérez y Rojas (2016) permitió condujo a la obtención de un aporte a la comprensión de los procesos de respuesta utilizados para contestar los ítems de razonamiento en contextos verbales al emplear técnicas que evocaran los procesos que seguían los estudiantes cuando resolvían las preguntas. Otro estudio de Cerdas y Montero (2016) generó una primera aproximación de una tabla de especificaciones para los ítems de razonamiento en contexto matemático con base en el criterio de los expertos y en el modelo de Rasch (Bond & Fox, 2015). Sin embargo, el trabajo no contempló el análisis de reportes verbales sobre las estrategias que utilizan los examinados en la solución de estos ítems.

El presente estudio responde a la necesidad de continuar con el aporte sustancial sobre el entendimiento de las estrategias utilizadas por las personas que contestan los ítems que conforman la PAA, para lo cual los reportes verbales en voz alta han sido de gran ayuda internacional (Castillo & Padilla, 2013; Cromley & Azevedo, 2006; Cui & Roduta, 2013; Ercikan et al., 2010; Leighton & Gierl, 2007a, 2007b; Leighton, 2004; Taylor & Dionne, 2000) y bajo la perspectiva contemporánea (Borsboom, Mellenbergh & Heerden, 2004; Cronbach & Meehl, 1955; Embretson, 1996; Embretson & Gorin, 2001; Gorin, 2007; Markus & Borsboom, 2013; Messick, 1995) sobre la validez en el ámbito de las pruebas psicológicas y educativas (American Educational Research Association [AERA], American Psychological Association [APA] & National Council on Measurement in Education [NCME], 2014). Por tanto, esta es una investigación inicial y exploratoria que intenta contestar a la interrogante sobre si el modelo cognitivo constituido por cuatro estrategias definidas por jueces expertos es representativo de los procesos de respuesta, que subyacen a los ítems de razonamiento en contexto matemático de la PAA, para lo cual se emplea la técnica de pensamiento en voz alta. En este sentido, se entabla la siguiente pregunta: ¿Existe alguna evidencia que respalde un modelo de estrategias requeridas para resolver los ítems de razonamiento en contexto matemático de la Prueba de Aptitud Académica de la Universidad de Costa Rica?

Método

Participantes

El criterio para determinar el número de participantes fue el de saturación, el cual se refiere al momento en el que la recopilación de reportes verbales en voz alta no aporta estrategias de resolución nuevas (Hays & Singh, 2011). Por lo tanto, el muestreo utilizado en la presente investigación fue de tipo no probabilístico y a conveniencia. Se seleccionó a un grupo de ocho estudiantes de la UCR, quienes estaban cursando el primer año en esta institución. Los participantes se seleccionaron por su condición de haber contestado correctamente, al menos, el 90 % de las preguntas de la PAA y por provenir de una

institución pública de educación secundaria. La condición de selección de participantes con una proporción alta de respuestas correctas se realizó con la intención de que los estudiantes fueran exitosos tanto en la resolución de los ítems como en el reporte en voz alta de las estrategias de resolución. La elección de las personas de instituciones públicas correspondió a la proporción mayoritaria de aspirantes provenientes de colegios públicos que ingresa a la UCR (Programa Estado de la Nación, 2013). Cada participante fue contactado por vía telefónica y en el momento del inicio de las entrevistas leyeron y firmaron un consentimiento informado de participación respaldado por el Comité de Ética de la UCR.

Instrumentos

Se eligieron 15 ítems del banco de ítems de razonamiento en contexto matemático para conformar un test para la entrevista. La estructura de los ítems es de escogencia única con un encabezado y cinco alternativas de respuesta de las cuales solo una es la correcta. Los ítems de la PAA contienen situaciones referidas al medio o ambiente construido a partir de experiencias a las que comúnmente se enfrentan los aspirantes. En estas experiencias educativas, personales, trabajo, imaginarias, entre otras, se espera que los aspirantes se involucren con las nociones de conceptos básicos de matemática de la tabla 1. Los ítems utilizados forman parte del banco de ítems que se utilizan regularmente para el ensamblaje de la PAA usada en el proceso de admisión a la UCR y la UNA, por lo cual son de naturaleza confidencial y no pueden ser incluidos en este trabajo.

Tabla 1

Nociones de conceptos básicos de matemática para la PAA

Código	Descripción
1	Operaciones aritméticas con números naturales, enteros, fraccionarios, con expansión decimal
2	Mínimo común múltiplo y máximo común divisor
3	Relaciones numéricas y de orden
4	Conteo intuitivo
5	Permutaciones
6	Geometría (área y perímetro de polígonos, circunferencia)
7	Sucesiones (numéricas, de objetos, de personas)
8	Análisis de datos y probabilidad: información presentada en forma tabular, mínimo, máximo, promedio, eventos más probables, igualmente probables o menos probables
9	Notación de expresiones algebraicas
10	Relaciones algebraicas de orden y de equivalencia (mayor, menor, igualdad, mayor o igual, menor o igual)
11	Unidades de medida (masa, longitud, capacidad, tiempo, monetarias)

Procedimientos

Con el propósito de explorar los procesos involucrados en la solución de los ítems de razonamiento en contexto matemático de la PAA, se realizó primero una revisión bibliográfica sobre la resolución de problemas en matemática y en tareas de razonamiento (Alda & Hernández, 1998; Echenique, 2006; Leighton & Sternberg, 2004; Polya, 1987; Rico, 2006). A partir del criterio experto de los investigadores, quienes han trabajado tanto en el diseño la PAA durante los últimos 10 años como en los aportes teóricos de acercamiento al constructo que se mide en esta prueba, se elaboró un sistema de categorías que respondiera a las estrategias específicas requeridas para resolver los ítems de razonamiento en contexto matemático de esta prueba. Luego, se examinaron 60 ítems del banco que exhibían una calidad técnica alta según los criterios de los parámetros de discriminación y de dificultad del modelo logístico de dos parámetros de teoría de respuesta al ítem (Baker & Kim, 2004). En la escala numérica continua de -3 a 3, los ítems presentaron valores alrededor de 1 para la discriminación y superiores a 0 en la dificultad.

Posteriormente, de los 60 ítems analizados se eligieron 15 representativos de las siguientes categorías: generalización, verificación, indagación y representación. Estas categorías fueron definidas por el equipo de investigación de la Prueba, quienes las consideran inherentes a estrategias de resolución de ítems con situaciones contextuales deseables en la medición del razonamiento. Las definiciones de las categorías se incluyen en la tabla 2.

Tabla 2

Categorías de análisis de los ítems

Categoría	Definición
Generalización	Los ítems clasificados en esta categoría presentan en su enunciado expresiones numéricas, algebraicas con descripciones o geométricas para las que se cumple un patrón que debe ser descubierto por la persona que resuelve el problema. La validez de una conclusión depende de la relación que existe entre la veracidad de las premisas dadas en el enunciado y la información que se puede tener acerca de la situación. El patrón de algunos ítems podría formar parte de las opciones de respuesta. En otros casos, el patrón se determina mediante algún término, ley, regla o caso específico de una secuencia.
Representación	La situación del ítem permite esbozar la información de forma inmediata mediante un lenguaje gráfico, simbólico o verbalizado. Se espera que la persona transforme esos esbozos en caminos que le faciliten producir un desglose y organización más simple de la información, y la guíen para encontrar la solución de la situación.

Continúa...

K. JIMÉNEZ, G. ROJAS, A. BRISUELA, N. PÉREZ

Indagación	En el enunciado del ítem se presentan condiciones que se deben cumplir, y que deben ser manipuladas para construir una o varias situaciones posibles o casos de lo que sucedería si se asumen valores y operaciones creadas por la persona o sugeridos por las mismas opciones posibles de respuesta, sin tener que efectuar una verificación de cada una. Los casos creados por la persona le permiten conjeturar sus propias hipótesis para llegar a la respuesta.
Verificación	Los ítems clasificados en esta categoría presentan en el enunciado información que guía la creación de uno o varios casos que deben ser revisados o analizados para confirmar su veracidad. En algunas situaciones, la respuesta a un ítem requiere un análisis y una verificación de cada opción de respuesta. En el análisis, la persona que resuelve identifica la respuesta correcta mediante una conexión entre el enunciado y las opciones. Para llegar a una adecuada solución de la situación, se escoge una opción de respuesta que no invalida o contradice la información dada en el planteamiento de la situación. En otros casos, en el enunciado o en las opciones de respuesta aparecen proposiciones y se requiere asumir lo propuesto como cierto para encontrar, al menos, un ejemplo que invalide la proposición para así descartarla.

Los participantes respondieron a una entrevista en la oficina del Programa de la PAA de la UCR. Cada participante realizó la entrevista individualmente y en momentos distintos. Un único juez aplicó todas las entrevistas. Se elaboró una guía de entrevista con preguntas como las siguientes: ¿Cómo llegó a esa respuesta?, ¿cómo lo resolvió?, ¿comenzó por analizar las opciones? A cada participante se le entregaron los 15 ítems (en un orden diferente y al azar para cada uno) para que los resolviera en voz alta, esto es, para que reportara todo lo que pensaba mientras intentaba seleccionar una respuesta (reporte verbal concurrente) y un folleto de hojas en blanco por si requerían escribir y organizar la información del ítem (reporte escrito). Una vez elegida la opción, a los participantes se les planteaban las preguntas mencionadas anteriormente. En la parte final de la entrevista, se obtuvo información sobre las estrategias empleadas mediante un reporte verbal retrospectivo.

Codificación

Se procedió con las transcripciones de las entrevistas y el análisis de las respuestas dadas por cada estudiante. Finalmente, dos de los autores codificaron las entrevistas de manera individual e independiente. En cada uno de los 15 ítems, se evidenciaba que los estudiantes lo resolvían según lo esperado si empleaban una estrategia relacionada con la categoría a la que fue inicialmente asignado cada ítem. La codificación (0 = no resolvió según lo esperado, 1/2 = se acercó a la estrategia esperada utilizando algunos elementos y 1 = se utilizó la estrategia esperada). Así, se corroboró que la clasificación hecha inicialmente por los jueces expertos estuviera en concordancia con la forma en que los estudiantes resolvieron los ítems de razonamiento en contexto matemático de la PAA. Con el propósito de poder evaluar los niveles de acuerdo entre los jueces, según las categorías de resolución del ítem, se utilizó el coeficiente de kappa de Cohen cuya interpretación se realizó con los lineamientos de Landis y Koch (1977): pobre (menor a 0), ligero (0 – 0.20), justo (0.21 – 0.40), moderado (0.41 – 0.60), sustancial (0.61 – 0.80) y casi perfecto (0.81–1). Los análisis se realizaron con el paquete denominado *Various Coefficients of Interrater Reliability and Agreement* ('irr') (Gamer, Lemon, Fellows, & Singh, 2015) del programa R (R Core Team, 2013).

Resultados

El coeficiente kappa de Cohen, expuesto en la tabla 3, muestra la concordancia entre los jueces que clasificaron en forma individual e independiente el conjunto de ítems seleccionados según las categorías. El coeficiente kappa de Cohen, para todas las categorías, resultó, al menos, moderado según los criterios proporcionados por Landis y Koch (1977).

Tabla 3
Concordancia entre codificadores

Categoría	Kappa de Cohen
Generalización	.689
Representación	.586
Indagación	.497
Verificación	.438
Total	.545

En concordancia con los parámetros de Landis y Koch (1977), las cuatro categorías de la Tabla 2 se confirman mediante el acuerdo entre jueces. Los jueces tuvieron un acuerdo moderado en Representación, Indagación y Verificación. Los expertos también presentaron un acuerdo sustancial en la categoría de Generalización. A continuación, la definición de cada categoría se complementa con ejemplos de las frases empleadas por los participantes en las entrevistas, las cuales permitieron a los jueces expertos la ratificación del uso de las estrategias esperadas para la resolución de los ítems.

Según la tabla 2, los ítems de la categoría Generalización presentan en el enunciado expresiones que cumplen una regla o patrón que debe ser generada por la persona que resuelve el problema. Los participantes encontraron la ley que rige las expresiones del enunciado y expresaron frases que confirman el uso de la Generalización.

- 1) ...ir descomponiendo en anteriores, según las reglas que vienen aquí.
- 2) ...tengo que observar el patrón que sigue y trato de analizar esto para lo que continúa.
- 3) ...la secuencia es que [expresión], sí, la regla va a ser [expresión].

En la categoría Representación se esperaba que las personas organizaran la información sobre la situación planteada en un lenguaje gráfico, simbólico o verbalizado de manera que determinaran un camino guiado hacia la solución de la situación:

- 4) ... pongo todos los datos primero entonces para relacionarlos y los voy viendo más fácil que viéndolos de la pregunta.

K. JIMÉNEZ, G. ROJAS, A. BRISUELA, N. PÉREZ

5) ... se acomoda solito, como que la cabeza sabe qué hacer con cada dato, ya sabe dónde poner las cosas para que funcionen.

6) ...primero acomodar los datos que tenía y calcular los datos que se podían sacar de primera mano, es como una tabla de restas.

Al analizar los reportes escritos de los entrevistados, se observó, en el caso de esta categoría, que algunos dibujaban tablas, esquemas o expresiones con símbolos que representaban equivalencias o relaciones entre ellas.

Los ítems categorizados como de Indagación tienen en el enunciado varias condiciones que intentan exigir de los examinados la creación de una o varias posibles situaciones o casos para conjeturar sus propias hipótesis que les guíaran a la respuesta:

7) ... supóngase que [afirmación] empiezo a evaluar opciones.

8) ...entonces realmente la opción es [respuesta elegida] ya que [afirmación] no puede ser igual a... porque [afirmación], suponga entonces que se tiene...

9) me basé en la condición que dice que [expresión], me fui directamente a la [respuesta elegida], si la [respuesta elegida] fuera mentira, significaría que...

10) ...con los datos que tengo siempre reviso que las opciones digamos sean descartables o sean correctas.

Finalmente, en los ítems clasificados en la categoría Verificación los entrevistados realizaron y revisaron una comprobación de uno o varios casos creados de forma natural y guiada a partir de la información del enunciado. En otros casos, las personas mostraron el proceso de solución de los ítems mediante una conexión entre el enunciado y las opciones de respuesta. Esta vía de aproximación a la respuesta correcta se ejecuta con la confirmación de una opción que no invalida o contradice la información dada en el planteamiento de la situación.

11) ...fui probando, planteé que eran los datos y fui probando cada opción, buscando la posibilidad de que esa opción se cumpliera.

12) ...probar casos, creo que es como contraejemplos. (los casos se refieren a las opciones de respuesta dadas).

13) ...como las opciones, fui viendo casos, si tengo que buscar un caso en el que esto se cumpla y sea correcto a lo que diga en el encabezado.

14) ...fui probando cada una de las opciones del valor de [expresión], para ver si existía una solución o un resultado para [expresión] que se complementara y que estuviera dentro del rango permitido.

Discusión

En el estudio presentado fue posible implementar una metodología de clasificación y comprobación de categorías presentes en los ítems de la PAA mediante el uso de los reportes verbales de estudiantes con calificaciones altas en la PAA y podían expresar en voz alta con mayor precisión los procesos de resolución de las preguntas. Los resultados sugieren que la técnica y los procedimientos constituyen

indicadores para recabar información de los participantes en una eventual aplicación de la metodología a otros test que requieran de la confirmación de estrategias de resolución en ítems diseñados para medir capacidades cognoscitivas.

La taxonomía construida para la PAA se verificó en función de las frases codificadas que potencialmente cumplían con las definiciones teóricas de las categorías de Generalización, Representación, Indagación y Verificación. Las frases permitieron a los jueces expertos una aproximación muy cercana a lo que verdaderamente expresaban los participantes y lo que se esperaba que estuviera involucrado en cada ítem. Sin embargo, los resultados deben interpretarse con cautela debido a que la muestra de estudiantes se consideró a conveniencia y con un perfil de éxito alto en las respuestas correctas en la PAA. Esta selección ofreció información de los procesos de respuesta, pero limitó conocer la generalización a personas con menor porcentaje de respuestas correctas en la prueba, por lo que una posible línea de investigación emergente podría centrarse en la ampliación y la especificación de subprocesos en las categorías con estudiantes de distinto perfil de ingreso la universidad.

Los resultados mostrados también abren la posibilidad de implementar estudios cuasiex-perimentales (Montero, Rojas, Negrín, & Francis, 2015) y de predicción de las puntuaciones de la PAA en el rendimiento académico en la UCR (Jiménez & Morales, 2009-2010; Rojas, 2013). En el estudio experimental de Montero et al. (2015) podría modificarse la capacitación con una intervención que propicie a las personas con diferentes perfiles de desempeño en resolución de ítems apropiarse del manejo guiado de los cuatro procesos de manera que se podría analizar si la intervención tiene un efecto positivo sobre el incremento de las puntuaciones en la PAA. En los trabajos de Jiménez y Morales (2009-2010) y Rojas (2013) se podrían incluir las categorías como variables independientes en diferentes modelos de regresión lineal múltiple. También, las definiciones de la tabla 2 podrían guiar una revisión futura de bancos de ítems y la confección de nuevos ítems que se apeguen al marco teórico de las categorías.

Referencias

- American Educational Research Association (AERA), American Psychological Association (APA) y National Council on Measurement in Education (NCME) (2014). *Standards for Educational and Psychological Testing*. Washington, Estados Unidos: American Educational Research Association.
- Alda, F., & Hernández, M. D. (Enero, 1988). Resolución de problemas. *Cuadernos de Pedagogía*, 265, 28-32.
- Baker, F.B, & Kim, S.H. (2004). *Item Response Theory: Parameters Estimation Techniques*. New York: Marcel Dekker, Inc.
- Bond, T., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences*. New York: Routledge.
- Borsboom, D., Mellenbergh, G.J., & Heerden, J. (2004). The Concept of Validity. *Psychological Review*, 111(4), 1061-1071.
- Brizuela, A., Jiménez, K., Pérez, N. & Rojas, G. (2016). Autorreportes verbales en voz alta para la identificación de procesos de razonamiento en pruebas estandarizadas. *Revista Costarricense de Psicología*, 35(1), 17-30.

K. JIMÉNEZ, G. ROJAS, A. BRISUELA, N. PÉREZ

- Castillo, M., & Padilla, J. (2013). How cognitive interviewing can provide validity evidence of the response processes to scale items. *Social indicators research*, 114(3), 963-975.
- Cerdas, D., & Montero, E. (2016). Uso del modelo de Rasch para la construcción de tablas de especificaciones: propuesta metodológica aplicada a una prueba de selección universitaria. *Actualidades Investigativas en Educación*, 17(1). 1-16.
- Cromley, J., & Azevedo, R. (2006). Self-report of reading comprehension strategies: What are we measuring? *Metacognition and Learning*, 1(3), 229-247.
- Cronbach, L.J., & Meehl, P.E. (1955). Construct Validity in Psychological Test. *Psychological Bulletin*, 52(2), 281-302.
- Cui, Y., & Roduta, M. (2013). Validating Student Score Inferences with Person-Fit Statistic and Verbal Reports: A Person-Fit Study for Cognitive Diagnostic Assessment. *Educational Measurement: Issues and Practice*, 32(1), 34-42.
- Echenique, I. (2006). *Matemáticas resolución de problemas*. Recuperado de <http://dpto.educacion.navarra.es/publicaciones/pdf/matematicas.pdf>
- Embretson, S., & Gorin, J. (2001). Improving Construct Validity Cognitive Psychology Principles. *Journal of Educational Measurement*, 38(4), 343-368.
- Embretson, S. E. (1996). The New Rules of Measurement. *Psychological Assessment*, 8(4), 341-349.
- Ercikan, K., Arim, R., Law, D., Domene, J., Gagnon, F., & Lacroix, S. (2010). Application of think aloud protocols for examining and confirming sources of differential item functioning identified by expert reviews. *Educational Measurement: Issues and Practice*, 29(2), 24-35.
- Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2015). *Package irr: Various Coefficients of Interrater Reliability and Agreement* (versión 0.84). Recuperado de <https://cran.r-project.org/web/packages/irr/irr.pdf>
- Gorin, J. S. (2007). Reconsidering Issues in Validity Theory. *Educational Researcher*, 36(8), 456-462.
- Hays, D. & Singh, A. (2011). *Qualitative inquiry in clinical and educational settings*. Estados Unidos: Guilford Press.
- Jiménez, K., & Morales, E. (2009-2010). Validez predictiva del Promedio de Admisión de la Universidad de Costa Rica y sus componentes. *Actualidades en psicología*, 23-24 (110-111), 21-55.
- Juan-Espinosa, M. d. (1997). *Geografía de la Inteligencia Humana*. Madrid: Pirámide.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
- Leighton, J. (2004). Avoiding misconception, misuse, and missed opportunities: The collection of verbal reports in educational achievement testing. *Educational Measurement: Issues and Practice*, 23(4), 6-15.
- Leighton, J., & Gierl, M. (2007a). Verbal Reports as Data for Cognitive Diagnostic Assessment. En J. Leighton y M. Gierl (Eds.), *Cognitive Diagnostic Assessment for Education* (pp. 146-172). Estados Unidos: Cambridge University Press.

- Leighton, J., & Gierl, M. (2007b). Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' thinking processes. *Educational Measurement: Issues and Practice*, 26(2), 3-16.
- Leighton, J. P., & Sternberg, R. J., (2004). *The Nature of Reasoning*. United States: Cambridge University Press.
- Markus, K., & Boorsboom, D. (2013). *Frontiers of Test Validity: Measurement, Causation, and Meaning*. Estados Unidos: Routledge.
- Martínez Arias, M. R., Hernández Lloreda, M. J., & Hernández Lloreda, M. V. (2006). *Psicometría*. Madrid: Alianza Editorial.
- Messick, S. (1995). Validity of Psychological Assessment. *American Psychologist*, 50(9), 741-749.
- Ministerio de Educación Pública (2012). *Reforma curricular en ética, estética y ciudadanía: Programas de estudio de matemática*. Recuperado de <http://www.mep.go.cr/sites/default/files/programadeestudio/programas/matematica.pdf>
- Molina, M., Gallardo, E., & Cordero, R. (2011). *Tiempo para graduarse en la Universidad de Costa Rica: Factores Socioeconómicos y académicos asociados* [Reporte de investigación del proyecto B1075]. Costa Rica: Universidad de Costa Rica.
- Montero, E., Rojas, G., Negrín, M., & Francis, S. (2015). Efectos de los puntajes de la prueba de admisión de la Universidad de Costa Rica: una aproximación bayesiana. *Actualidades en Psicología*, 29 (119), 115-130.
- Montero, E., Villalobos, J., & Valverde, A. (2007). Factores institucionales, pedagógicos, psicosociales y sociodemográficos asociados al rendimiento académico en la Universidad de Costa Rica: Un análisis multinivel. *RELIEVE*, 13(2), 215-234.
- Polya, G. (1987). *Cómo plantear y resolver problemas*. México: Trillas, S.A.
- Programa Estado de la Nación (2013). *Cuarto informe Estado de la Educación*. Costa Rica: Programa Estado de la Nación. Recuperado de <http://www.estadonacion.or.cr/informe-iv-estado-educacion?highlight=WyJjdWFydG8iXQ==>
- R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: <http://www.R-project.org/>.
- Rico, L. (2006). Marco teórico de evaluación en PISA sobre matemática y resolución de problemas. *Revista de Educación*, extraordinario 2006, 275-294.
- Rojas, L. (2013). Validez predictiva de los componentes del promedio de admisión a la Universidad de Costa Rica utilizando el género y el tipo de colegio como variables control. *Actualidades Investigativas en Educación*, 13(1), 45-69.
- Taylor, K., & Dionne, J. (2000). Accessing problem-solving strategy knowledge: The complementary use of concurrent verbal protocols and retrospective debriefing. *Journal of Educational Psychology*, 92(3), 413-425.

K. JIMÉNEZ, G. ROJAS, A. BRISUELA, N. PÉREZ

Recibido: 17 de marzo de 2017

Revisión recibida: 20 de noviembre de 2017

Aceptado: 25 de abril de 2018

Sobre los autores:

Karol Jiménez Alfaro se tituló como Máster en Matemática con énfasis en Matemática Educativa y es investigadora en el Programa de la Prueba de Aptitud Académica de la Universidad de Costa Rica.

Guaner Rojas Rojas se graduó como Doctor en Metodología de las Ciencias del Comportamiento y de la Salud con énfasis en medición y especialidad en Modelos de Diagnóstico Cognitivo de la Universidad Autónoma de Madrid. Labora como investigador en el Programa de la Prueba de Aptitud Académica de la Universidad de Costa Rica.

Armel Brizuela Rodríguez obtuvo el Máster en Investigación Psicológica y labora como investigador en el Programa de la Prueba de Aptitud Académica de la Universidad de Costa Rica.

Nelson Pérez Rojas es Bachiller en Filología Española e investigador en el Programa de la Prueba de Aptitud Académica de la Universidad de Costa Rica.

Publicado en línea: 28 de junio de 2018