



Revista Educación
ISSN: 0379-7082
ISSN: 2215-2644
revedu@gmail.com
Universidad de Costa Rica
Costa Rica

Validación de una prueba de comprensión lectora para estudiantes universitarios

Brizuela Rodríguez, Armel; Pérez Rojas, Nelson; Rojas Rojas, Guaner

Validación de una prueba de comprensión lectora para estudiantes universitarios

Revista Educación, vol. 44, núm. 1, 2020

Universidad de Costa Rica, Costa Rica

Disponible en: <https://www.redalyc.org/articulo.oa?id=44060092005>

DOI: <https://doi.org/10.15517/revedu.v44i1.34983>



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivar 3.0 Internacional.


Validación de una prueba de comprensión lectora para estudiantes universitarios

A Validation of a Reading Comprehension Test for University Students

Armel Brizuela Rodríguez

Universidad de Costa Rica, Costa Rica

armel.brizuelarodriguez@ucr.ac.cr

 <http://orcid.org/0000-0002-0185-8952>

DOI: <https://doi.org/10.15517/revedu.v44i1.34983>

Redalyc: <https://www.redalyc.org/articulo.oa?id=44060092005>

Nelson Pérez Rojas

Universidad de Costa Rica, Costa Rica

nelson.perezrojas@ucr.ac.cr

 <http://orcid.org/0000-0001-8929-3249>

Guaner Rojas Rojas

Universidad de Costa Rica, Costa Rica

guaner.rojas@ucr.ac.cr

 <http://orcid.org/0000-0002-3064-9631>

Recepción: 14 Marzo 2019

Aprobación: 19 Agosto 2019

RESUMEN:

El objetivo del presente artículo es brindar evidencias de validez y de confiabilidad sobre una prueba nueva para medir el nivel de comprensión lectora de estudiantes universitarios. Para ello se empleó un diseño de investigación observacional y transversal, mediante la cual se aplicó la prueba a un grupo de 244 estudiantes de primer ingreso a una carrera universitaria. Los resultados fueron analizados con base en la Teoría Clásica de los Test y en el modelo de Rasch, así como también se estimó el coeficiente de correlación de Pearson entre las puntuaciones en la prueba y un curso universitario de primer año de carrera. El principal hallazgo de este estudio es la creación de una prueba de 44 ítems que permite generar puntuaciones con una confiabilidad de 0,74 y que presentan una correlación estadísticamente significativa ($p < 0,000$) de 0,21 [0,09, 0,33] con las notas de estudiantes en un curso de carrera de primer año. Con base en los análisis efectuados, se concluye que la prueba podría ser útil para detectar estudiantes de primer ingreso a la educación superior que podrían presentar problemas de comprensión lectora y, por ende, ver disminuido su rendimiento académico. En este sentido, la prueba podría servir para propósitos de *screening*, es decir, para coadyuvar en la elaboración de planes remediales, implementación de intervenciones educativas y desarrollo de unidades didácticas orientadas al fortalecimiento de la comprensión lectora de los(as) estudiantes universitarios de primer ingreso.

PALABRAS CLAVE: Evaluación del estudiante, Estudiante universitario, Prueba de respuesta múltiple, Lectura, Comprensión.

ABSTRACT:

The goal of this article is to provide evidence as to the validity and reliability of a new Reading Comprehension test aimed at college students which is based on an observational and cross-cutting research design. The test was taken by 244 students in the first year of their major. The results were analyzed according to the Classical Test Theory, the Rasch Model and a Pearson correlation coefficient to compare the scores obtained with first year courses taken by students for their Major. As a result of the study, a 44-item was elaborated with a reliability scale of 0.74 and a statistically significant correlation scores of 0.21 [0.09, 0.33] ($p < 0.000$) for student grades in first-year courses. It can be concluded that this test could be useful to detect reading comprehension difficulties among first-year university students which could potentially negatively impact their academic performance. The test could also serve as a screening tool in the development of remedial programs to strengthen reading comprehension among these students.

KEYWORDS: Student Evaluation, University Student, Multiple Choice Test, Reading, Comprehension.

INTRODUCCIÓN

La lectura de textos es uno de los mecanismos más importantes para adquirir conocimientos en la universidad: entre los deberes del estudiantado está la lectura de documentos sociales, científicos y tecnológicos; por tanto, las dificultades para comprender textos podrían afectar la asimilación de los nuevos contenidos por parte de la población estudiantil y su rendimiento académico. Además, los textos preuniversitarios tienen características discursivas distintas a los textos universitarios: en estos últimos, el empleo de léxico técnico, la gramática compleja, la variedad de superestructuras y las exigencias en la interacción de los contenidos requieren de un mayor desarrollo de las habilidades de comprensión lectora.

Por otra parte, la educación universitaria tiene como propósito el desarrollo individual, social y profesional del alumnado; en consecuencia, la importancia de la lectura no se limita al ámbito académico. Por ello, ha habido un interés creciente en que la población estudiantil sea capaz de comprender la información relevante sobre los hechos que afectan el mundo y el contexto inmediato: economía, política, sucesos nacionales e internacionales, divulgación científica, crisis humanitarias, u otros eventos similares. Asimismo, en el ámbito laboral, la lectura no solamente es necesaria para que las personas profesionales actualicen sus conocimientos, sino que también es uno de los medios para garantizar la calidad de los productos, los servicios y las interacciones humanas, especialmente cuando la comunicación social cada vez cuenta con canales más diversos, la información es más abundante y se requiere mayor pericia para identificar su pertinencia y su veracidad.

Así las cosas, evaluar el desempeño del estudiantado universitario en las tareas asociadas a la comprensión lectora con el propósito de establecer soluciones, en caso de que sea necesario, es una medida responsable en el accionar de las entidades que tienen a cargo la formación de ciudadanía y profesionales.

No obstante, las habilidades de comprensión lectora de estudiantes de universidad no han sido tan investigadas como las de la población infantil (Landi, 2010), lo cual ha contribuido a que en la actualidad existan más pruebas para medir habilidades de bajo nivel (reconocimiento de palabras, velocidad lectora, conciencia fonológica, etc.) que las disponibles para medir procesos cognitivos de alto nivel asociados a la comprensión de textos complejos académicos.

En el ámbito hispano hablante, existen aproximadamente cuarenta instrumentos orientados a la medición de la comprensión lectora para diferentes edades, con formatos de respuesta de toda índole y diseñados para propósitos educativos, clínicos, de selección, etc. (Cruz, 2018). Asimismo, existen diversas técnicas que se han utilizado para medir la comprensión lectora, tales como la evocación o el recuerdo libre, preguntas de sondeo, preguntas de respuesta abierta y cuestionarios de respuesta seleccionada (Pérez, 2005). Especial mención merece la prueba de comprensión lectora desarrollada por la doctora Violeta Tapia en 1977 (López, 2010), pues es una de las pocas que ha sido diseñada para población universitaria. En dicha prueba se miden las siguientes habilidades: identificación de información sobre hechos específicos, definición del significado de las palabras, identificación de la idea central del texto, interpretación de los hechos, inferencia sobre el autor, inferencia sobre el contenido del fragmento y rotulación (López, 2010). Otra prueba que merece una mención aparte es *CompLec*, la cual fue desarrollada por Lloréns et ál. (2011) y tiene como población meta a adolescentes.

En el ámbito no hispanohablante, existe una gran variedad de pruebas estandarizadas de comprensión lectora, como el *TOEFL* o uno de los subtest del *SAT*. Existen otras pruebas de aplicación más restringida, como el *Nelson Denny Reading Test* o la prueba *Woodcock-Johnson III Diagnostic Reading Battery* (también existe una versión en español, la *Woodcock-Muñoz*). Asimismo, existen pruebas educativas en las que se incluye un componente de comprensión lectora y están diseñadas para evaluar esta habilidad en relación con las habilidades adquiridas mediante la educación formal, tales como *PIRLS* (Progress in International Reading Literacy Study), *PISA* (Programme for International Student Assessment), *PIAAC* (Programme for the International Assessment of Adult Competencies), *EECL* (Estudio Europeo de Competencia

Lingüística), *NEPS* (National Educational Panel Study), entre otras. Para conocer con cierto detalle en qué consisten este tipo de pruebas, puede consultarse el trabajo de Reyes, Castillo, Zúñiga y Llarena (2009).

Ante esta carencia de pruebas de comprensión lectora en español orientadas específicamente a estudiantes universitarios, se hace necesario desarrollar un instrumento que permita identificar las posibles debilidades de los(as) estudiantes de primer ingreso a la educación superior. Además de las carencias de pruebas orientadas a estudiantes universitarios, es importante agregar que la mayoría ha sido diseñada sin un modelo de medición subyacente, por lo que el siguiente estudio presenta algunos resultados iniciales de la validación de una prueba estandarizada para estudiantes universitarios primer ingreso a la educación superior.

ANTECEDENTES TEÓRICOS SOBRE LA COMPRENSIÓN LECTORA

Según la Organización para la Cooperación y el Desarrollo Económico (2000), la capacidad de asignarle significado a los textos permite cumplir objetivos propios e incentivar la participación de los individuos en el nivel social. La comprensión lectora es una habilidad central que las y los estudiantes deben poseer para adquirir conocimientos y habilidades en todas las disciplinas académicas que integran la educación superior (Leighton y Gierl, 2011). No obstante, en los países latinoamericanos, el rendimiento de las y los estudiantes en cuanto a la comprensión lectora ha mostrado ser sumamente bajo de acuerdo con los resultados obtenidos en las pruebas PISA (Montero, Rojas, Zamora y Rodino, 2011).

La lectura es una actividad compleja que ha impactado al ser humano individual y culturalmente, pues es un proceso que implica una serie de sistemas representacionales y procesos cognitivos fundamentales en todo tipo de aprendizajes (Berko, Vellutino y Wolf, 1999). Por lo tanto, la lectura es una interacción entre el lector y el texto (Solé, 2009) que requiere la capacidad de comprender un texto escrito (Camps y Colomer, 1996).

Según Cassany (2006), la lectura puede ser vista desde tres concepciones distintas: la lingüística, en la cual la atención está fijada en el texto como poseedor del significado; la psicolingüística, según la cual el peso del análisis recae sobre los procesos de lectura por parte del lector (leer también requiere habilidades cognitivas, como crear hipótesis, verificarlas o reformularlas); o social, cuyo interés está centrado en los textos como productos sociales y su recepción en un determinado contexto.

Desde las perspectivas psicolingüística y pedagógica, tradicionalmente los modelos de comprensión de textos han sido divididos en dos tipos (Camps y Colomer, 1996; Berko et ál., 1999; Solé, 2009): procesamiento ascendente y procesamiento descendente. En el primero, el lector procesa las unidades lingüísticas menores (letras, palabras, frases y oraciones) hasta llegar a la unidad mayor (el texto). En el segundo, el lector le da significado al texto a partir de sus conocimientos previos y la activación de procesos cognitivos. Al respecto, Camps y Colomer (1996) afirman que estos conocimientos sobre el procesamiento de la información son de suma importancia para comprender la lectura, dado que el procesamiento descendente permite que el lector resuelva ambigüedades y escoja interpretaciones posibles para el texto.

Si se entiende la lectura como proceso que combina tanto el procesamiento ascendente como el descendente, es claro que el texto genera expectativas en el lector en distintos niveles, y este intentará satisfacerlas durante la lectura: la información de un nivel inferior se propaga y sirve de *input* al siguiente nivel, como en el procesamiento ascendente; de manera simultánea, se crean expectativas en el nivel semántico sobre la globalidad del texto que deben ser verificadas, es decir, también se da un procesamiento descendente (Solé, 2009). Por una parte, el lector es un participante activo encargado del análisis y el procesamiento del texto; el lector pretende cumplir con un objetivo de lectura, lo cual repercutirá en el significado que resulte de la actividad (Colomer y Camps, 1996; Day y Bamford, 1998; Alderson, 2000; Solé, 2009;). Por otra parte, los aportes efectuados por el texto están determinados por distintas secuencias: el lector se encontrará con información diferente si lee una novela o un periódico, lo cual también modificará sus expectativas y sus objetivos de lectura.

En el proceso de comprensión de la lectura, el análisis de los signos grafonémicos presentes en el mensaje activa la representación semántica previamente almacenada de la palabra correspondiente a esos signos. La representación activada permanece en la memoria operativa durante un tiempo limitado, lo cual permite su integración con el significado de otras palabras del mensaje. Así, la inferencia se daría mediante la activación conjunta de varios significados que proceden de distintas partes del mensaje y de los conocimientos previos del lector (Gutiérrez, 1999).

La comprensión lectora, como proceso cognitivo, se ubica a medio camino entre los procesos de percepción, por un lado, y los de pensamiento y resolución de problemas, por otro. Normalmente, las personas leen y entienden un texto de manera análoga a como lo harían con una imagen, sin un planeamiento elaborado o el uso de estrategias de resolución de problemas. Sin embargo, cuando este proceso se encuentra con escollos, el lector se convierte en un solucionador de problemas que debe esforzarse por comprender el texto que lee. En este sentido, la comprensión se refiere a la construcción automática de significado y se transforma en un proceso consciente y activo de resolución de problemas cuando el texto plantea desafíos de mayor peso al lector (Kintsch, 2013).

La comprensión lectora depende del funcionamiento e integración de muchos procesos cognitivos. Para entender una oración, es necesario procesar visualmente las palabras individuales, así como identificar y acceder a sus representaciones fonológicas, ortográficas y semánticas para posteriormente construir el significado global de toda la oración. De manera análoga, para comprender un conjunto de oraciones que componen un texto, el lector necesita procesar y conectar entre sí la información individual que transmite cada oración. Para que dicho proceso sea exitoso, muchos factores entran en juego, como las características del lector, las propiedades del texto y las exigencias de la tarea específica de lectura (Kendeou, van den Broek, Helder y Karlsson, 2014).

La diferencia entre la comprensión de un texto compuesto por varias oraciones y la de una simple lista de oraciones es la siguiente: para que ocurra la primera se requiere que el lector detecte las relaciones conceptuales entre las oraciones de las que se compone el texto. Por lo tanto, un proceso exitoso de comprensión lectora da lugar a una red semántica de nodos y conexiones en la mente del lector (Van den Broek, 2012).

Los procesos de comprensión que permiten formarse una representación mental de un texto ocurren en distintos niveles: léxico, sintáctico y discursivo. Además de los aspectos estrictamente lingüísticos mencionados, el reto de comprender un texto adecuadamente demanda un mínimo de conocimientos sobre el contenido de este, estrategias de lectura adecuadas al tipo de texto, y a los objetivos del lector, y un nivel de motivación apropiado.

La complejidad de la comprensión lectora se evidencia a partir de los diversos modelos teóricos que describen los procesos cognitivos y lingüísticos que subyacen a esta. Algunos modelos enfatizan las representaciones mentales que los lectores construyen como resultado de la comprensión de las palabras, oraciones y sus respectivas relaciones dentro de un texto. Por otra parte, otros modelos se enfocan en la adquisición de las habilidades que se requieren para entender un texto. No obstante, a pesar de que los modelos se especialicen en diferentes aspectos de la comprensión lectora, tienen en común la noción de que la comprensión lectora implica la construcción de una representación mental coherente en la memoria del lector (Kintsch, 2013; Kendeou et ál., 2014). En otras palabras, en la comprensión lectora debe construirse una representación coherente en la memoria, a partir de las inferencias sobre las relaciones entre los elementos internos del texto, así como entre estos y los conocimientos que posea el lector sobre el tema; todo esto debe ejecutarse bajo las restricciones que imponen la capacidad de memoria de trabajo y los recursos atencionales de acuerdo con los estándares de coherencia del lector (Van den Broek, 2012).

Para conocer desde una perspectiva diacrónica los modelos existentes sobre la comprensión lectora, puede consultarse el trabajo de Alexander y Fox (2013). Desde un enfoque sincrónico, el texto de McNamara y Magliano (2009) da una visión general sobre los modelos más importantes en lo que respecta al estudio de este constructo. Finalmente, de obligada consulta es el texto editado por Alvermann, Unrau y Ruddell (2013), en

el que se expone con amplitud los aspectos psicológicos, educativos y sociales que conforman la investigación y enseñanza de la lectura. En la actualidad los modelos interactivos, en los que la comprensión es el producto de procesos ascendentes y descendentes son considerados los mejores para entender las diferencias individuales en la comprensión lectora (Leighton y Gierl, 2011).

LA MEDICIÓN DE LA COMPRENSIÓN LECTORA

Dado que la comprensión lectora es un constructo polifacético y los test siempre constituyen una muestra finita de todos los posibles ítems que podrían medir este constructo, ningún test de comprensión de lectura por sí mismo permitirá medir la comprensión lectora en su totalidad, sino solamente aspectos parciales de esta (Van den Broek, 2012).

Los cambios en las técnicas empleadas para la medición de los constructos asociados al lenguaje han sido moldeados por las contribuciones de la psicolingüística, la sociolingüística, el análisis del discurso y las investigaciones sobre enseñanza y adquisición de una segunda lengua, entre otros. Así pues, a lo largo de los años la medición de las habilidades relacionadas con la comprensión y producción de lenguaje ha incluido desde las estrategias más “indirectas” (i.e., ítems de selección o de completar textos incompletos) hasta las más “directas” (i.e., situaciones realistas de uso y comprensión de la lengua), lo cual ha enriquecido el campo de la medición del lenguaje con una gran variedad de tipos de indicadores diseñados con base en objetivos y enfoques teóricos a veces incompatibles entre sí (Bachman, 2007) y en muchas ocasiones no explicitados (Leslie y Caldwell, 2009; Perfetti y Adlof, 2012). Asimismo, en la elaboración de estos instrumentos la comprensión lectora se ha conceptualizado como procesamiento de información, como activación o como habilidad (Svetina, Gorin y Tatsuoka, 2011).

En diversos países se han venido utilizando pruebas para medir lo que, desde varios enfoques teóricos, se considera importante que un lector pueda desarrollar a la hora de enfrentarse a un texto (Reyes et ál., 2009). De igual forma, como se evidencia en uno de los primeros estudios sobre las pruebas de comprensión lectora (Davis, 1944), ya desde esta época existía el interés por determinar cuáles habilidades se requerían para contestar las preguntas típicas de este tipo de instrumentos de medición.

Es importante resaltar que no todas las pruebas denominadas “de comprensión lectora” miden este constructo de la misma manera (Svetina et ál., 2011), pues en cada una de ellas se enfatizan distintas facetas de este. Debido a que la comprensión lectora es un constructo complejo que implica la coordinación de varios procesos cognitivos (O'Reilly, Weeks, Sabatini, Halderman y Steinberg, 2014), existe una gran diversidad de pruebas de comprensión lectora, cada una de las cuales mide mejor ciertos aspectos de este constructo que otros (Keenan, Betjemann y Olson, 2008). Aunado a la complejidad de este constructo, también hay un conjunto de características que diferencian a esta gran cantidad de pruebas, tales como tiempo disponible para contestarlas, edad de la población meta, formato de respuesta (de elección o de respuesta construida), propósito (investigación, evaluación docente, selección o diagnóstico clínico), etc.

Magliano, Millis, Ozuru y McNamara (2007) afirman que los procesos de decisión implicados en responder preguntas de elección introducen procesos cognitivos irrelevantes a la capacidad para procesar textos escritos en contextos cotidianos. Así pues, este tipo de pruebas introducen en la medición habilidades que no están directamente relacionadas con la comprensión lectora (Farr, Pritchard y Smitten, 1990; Fletcher, 2006; Rupp, Ferne y Choi, 2006; Ozuru, Rowe, O'Reilly y McNamara, 2008; Cerdán, Vidal, Martínez, Gilabert y Gil, 2009). No obstante, debe señalarse que no existe un método infalible para medir la comprensión lectora, ya que cada uno permite medir aspectos parciales de la comprensión (Pérez, 2005), por lo que siempre será necesario elegir aquella técnica que mejor se ajuste a las restricciones particulares del caso (Alderson, 2000).

Así las cosas, es de vital importancia identificar específicamente cuáles procesos cognitivos mide una prueba porque de ello depende su utilidad en un determinado contexto. Por ejemplo, en el ámbito de la selección de

estudiantes que aspiran a ingresar a la educación superior (típicamente adultos jóvenes con un nivel alto de escolaridad) se requieren de pruebas que midan habilidades de alto nivel asociadas a las tareas usuales de un estudiante universitario (Jackson, 2005). En este sentido, y a la luz del estado de la cuestión con respecto a la medición de la comprensión lectora, la presente investigación se efectuó con el objetivo de desarrollar una prueba para medir las habilidades de comprensión lectora de estudiantes de universidad.

MÉTODO

La ejecución de este estudio fue autorizada por el Comité Ético Científico de la Universidad de Costa Rica. El diseño del estudio es observacional y transversal, por cuanto solamente se hizo una medición y no se efectuó una manipulación experimental.

Participantes

Los participantes del estudio conforman una muestra a conveniencia de 244 estudiantes (46% hombres) que ingresaron por primera vez a la carrera de Derecho de la Universidad de Costa Rica en 2018 y que estaban matriculados en el curso Introducción al estudio del Derecho I.

Instrumento

La Prueba de comprensión lectora fue diseñada con el objetivo de medir la capacidad de estudiantes universitarios de primer ingreso para (1) comprender las relaciones semánticas que existen entre las palabras, las frases y las oraciones de un texto y para (2) identificar la idea principal o el tema de este. Para ello, el instrumento está conformado por 60 ítems de selección única, en los cuales el estudiantado debe elegir una respuesta correcta de entre cuatro opciones, para lo cual cuentan con una hora y treinta minutos. Los 60 ítems se subdividen en grupos de 10, cada uno de los cuales se debe contestar con base en la lectura de seis textos. Asimismo, los ítems fueron elaborados para medir cuatro posibles niveles textuales: palabras, enunciados, párrafos y texto. En la Tabla 1 se muestra la distribución de los ítems para cada uno de los niveles textuales que se miden con la prueba, así como también por texto. Para cada uno de los seis textos se elaboraron 10 preguntas de modo que todos los niveles textuales estén representados en los textos.

TABLA 1
Distribución de ítems por texto y por nivel textual

Nivel textual	Texto 1	Texto 2	Texto 3	Texto 4	Texto 5	Texto 6
Palabras	2	2	2	2	2	2
Enunciados	2	2	2	2	2	2
Párrafos	4	4	4	4	4	4
Texto	2	2	2	2	2	2

Fuente: Elaboración propia.

Los textos fueron seleccionados de diversos libros introductorios universitarios de las áreas de ciencias sociales y de ciencias naturales. Por razones de espacio, los extractos de estos libros fueron modificados para que tuvieran un sentido completo en una página y posteriormente se presentan en una hoja aparte los respectivos ítems. Los ítems de la prueba son similares a los siguientes ejemplos: ¿Que# significa la palabra “impresiones” en el contexto en el que aparece?, exprese en sus propias palabras una síntesis del cuarto párrafo, ¿que# función cumple el último párrafo?, seleccione la opción que sintetiza el texto, etc.

En lo que respecta a los niveles textuales, estos fueron definidos antes de la elaboración de los ítems con el objetivo de que sirvieran de base para la elaboración de los ítems. En la Tabla 2 se presentan las respectivas definiciones.

Procedimiento

Los ítems de la prueba se elaboraron con base en la Tabla 1, de modo que cada uno de los niveles textuales puedan ser medidos por varios ítems. Dichos ítems fueron elaborados por dos de los investigadores a cargo del presente artículo. Posteriormente, el test se aplicó durante los meses de marzo y abril de 2018 en los grupos de estudiantes. Para ello, se coordinó con los profesores y las profesoras a cargo de dichos grupos, pues la prueba se aplicó en cada grupo en una sesión de aproximadamente una hora y treinta minutos. Para contestar las preguntas, el estudiantado empleó una hoja de respuesta diseñada para ser posteriormente leída por una máquina lectora óptica.

TABLA 2
Niveles textuales de la prueba y definiciones

Nivel textual	Definición
Palabras	Capacidad para identificar el significado de una palabra a partir de su contexto más inmediato
Enunciados	Capacidad para identificar la relación entre dos enunciados incluidos en un mismo párrafo
Párrafos	Capacidad para identificar la relación entre dos o más enunciados incluidos en distintos párrafos e identificar la idea principal de un párrafo
Texto	Capacidad para identificar el tema y la mejor síntesis de un texto completo

Fuente: Elaboración propia.

Fuente: Elaboración propia.

Análisis de datos

En primer lugar, se elaboró un análisis factorial confirmatorio para determinar la dimensionalidad del test. Luego, se estimó el coeficiente de confiabilidad alfa de Cronbach para estimar la consistencia interna de las puntuaciones generadas a partir del instrumento. Posteriormente, los datos fueron analizados en términos descriptivos mediante la Teoría Clásica de los Test con el paquete *CTT* (Willse, 2018) implementado en R (R Core Team, 2018). Se calcularon el índice de dificultad y el índice de discriminación de los 60 ítems con el fin de identificar ítems de baja calidad de la precisión de la medición del constructo.

Asimismo, para obtener evidencias sobre la pertinencia de los ítems con respecto a la población meta de la prueba (estudiantes de primer ingreso a la educación superior), se empleó el modelo de Rasch implementado en el programa *Winsteps* (Linacre, 2018). Este modelo de medición de la Teoría de la Respuesta al Ítem permite estimar la dificultad de los ítems y el nivel de habilidad del estudiantado de manera que pueden compararse directamente a través de un mapa de ítems y personas (también conocido como mapa de Wright). De esta manera, es posible identificar si el nivel de dificultad de la prueba es el apropiado en relación con la población meta del test.

Finalmente, se computó el valor del coeficiente de correlación entre la puntuación obtenida en la prueba y la nota obtenida por cada estudiante en un curso de primer año de la carrera de Derecho. Esta evidencia es pertinente para determinar la utilidad de la prueba en relación con el desempeño académico del estudiantado.

Una correlación positiva y moderada sería evidencia de que la prueba tiene un mínimo de capacidad predictiva en relación con el rendimiento académico en el primer año de la carrera de Derecho.

RESULTADOS

En la Tabla 3 se muestran los resultados del análisis factorial confirmatorio, el cual se estimó mediante el método de Máxima verosimilitud y permitió a poner a prueba la bondad de ajuste del modelo unidimensional que sustentó teóricamente la elaboración de los ítems. Este análisis se ejecutó con parcelas de ítems, las cuales se conformaron de acuerdo con los niveles textuales que se pretendía medir con los ítems. De esta manera, cada una de las parcelas corresponde a la suma de las puntuaciones obtenidas en los ítems respectivos. Así pues, 8 ítems pertenecen a la parcela *Palabra*, 10 a *Enunciados*, 20 a *Párrafos* y 6 a la parcela *Texto*.

Como se puede observar, el ajuste global del modelo es adecuado. El valor del Chi-cuadrado no es estadísticamente significativo, lo cual indica que la matriz de varianzas-covarianzas predicha por el modelo no es significativamente diferente de la matriz de varianzas-covarianzas observada. Asimismo, el índice de ajuste comparativo (CFI) es superior a 0,95, el error cuadrático medio de aproximación (RMSEA) es inferior a 0,05 y el residuo cuadrático medio estandarizado (SRMR) es inferior 0,08. Además, se puede observar que las cargas factoriales (solución completamente estandarizada) son superiores a 0,30. En este sentido, es posible concluir que las puntuaciones generadas a partir del test son unidimensionales.

TABLA 3
Análisis factorial confirmatorio

Parcela	Carga factorial	Error estándar	Intervalo de confianza (95%)		Residuo
Palabra	0,61	0,05	0,50	0,71	0,63
Enunciados	0,67	0,05	0,57	0,77	0,55
Párrafos	0,78	0,05	0,69	0,88	0,39
Texto	0,46	0,06	0,35	0,58	0,79
Chi-cuadrado	2,203	CFI	0,99		
Grados de libertad	2	RMSEA	0,02		
Significancia	0,33	SRMR	0,02		
		TLI	0,99		

Fuente: Elaboración propia.

En lo que respecta a la confiabilidad de las puntuaciones, el coeficiente alfa de Cronbach para los 44 ítems fue de 0,74, lo cual corresponde a un nivel de consistencia moderado.

En la Tabla 4 se muestran los ítems que obtuvieron índices de discriminación superiores a 0,10. En este sentido, 16 de los 60 ítems (4, 5, 6, 9, 10, 11, 12, 18, 19, 20, 32, 35, 40, 50, 51 y 54) fueron eliminados en esta etapa y no se incluyen en los análisis posteriores. Asimismo, en lo que respecta a la puntuación total obtenida a partir de la sumatoria de las respuestas correctas en estos 44 ítems, esta osciló entre 0 y 39 respuestas correctas, con una media de 23,95, una desviación estándar de 6,05 y una mediana de 25.

En términos generales, se puede apreciar que los ítems tendieron a ser relativamente fáciles, pues los índices de dificultad (proporción de respuestas correctas) fueron bastante altos. Sin embargo, 44 de los 60 ítems mostraron índices de discriminación positivos, lo cual indica que la correlación entre estos y la puntuación

total en el test es positiva. En otras palabras, 44 de los 60 ítems permiten discriminar entre estudiantes con puntuaciones bajas y aquellos(as) con puntuaciones altas.

TABLA 4
Índices de dificultad y discriminación

Ítem	Dificultad	Discriminación	Ítem	Dificultad	Discriminación
i1	0,81	0,17	i34	0,46	0,22
i2	0,73	0,10	i36	0,35	0,18
i3	0,81	0,18	i37	0,70	0,35
i7	0,66	0,19	i38	0,42	0,19
i8	0,55	0,20	i39	0,23	0,13
i13	0,49	0,16	i41	0,55	0,25
i14	0,80	0,18	i42	0,25	0,18
i15	0,45	0,12	i43	0,48	0,18
i16	0,52	0,20	i44	0,57	0,22
i17	0,53	0,15	i45	0,53	0,18
i21	0,46	0,38	i46	0,58	0,38
i22	0,73	0,26	i47	0,39	0,20
i23	0,61	0,33	i48	0,28	0,11
i24	0,61	0,28	i49	0,52	0,22
i25	0,65	0,15	i52	0,45	0,35
i26	0,52	0,29	i53	0,49	0,19
i27	0,41	0,18	i55	0,67	0,31
i28	0,53	0,26	i56	0,55	0,16
i29	0,66	0,17	i57	0,72	0,24
i30	0,40	0,16	i58	0,71	0,24
i31	0,58	0,12	i59	0,68	0,23
i33	0,45	0,27	i60	0,39	0,17

Fuente: Elaboración propia.

En lo que respecta al análisis efectuado con el modelo de Rasch, la confiabilidad de las personas fue de 0,75 y la de los ítems fue de 0,96. En lo que respecta al ajuste de los ítems, los valores del Infit MNSQ oscilaron entre 1,07 y 0,89. Por su parte, el ajuste de las personas medido por el Infit MNSQ osciló entre 1,19 y 0,78. Estos valores se encuentran dentro de un rango aceptable para concluir que los datos se ajustan relativamente bien al modelo de Rasch (Wright y Linacre, 1994).

En la Figura 1, la parte izquierda corresponde a los niveles de comprensión lectora estimados para quienes ejecutaron la prueba (cada símbolo # representa a 2 estudiantes y cada punto, a 1 estudiante). Por otra parte, en la zona derecha de dicha figura se representan los niveles de dificultad de los 44 ítems. Como se puede observar, los niveles de dificultad de los ítems y los de habilidad del estudiantado corresponden prácticamente al mismo rango de la escala (desde -2 a 2). Asimismo, las medias de la habilidad del alumnado ($M = 0,0$) y de dificultad de los ítems ($M = 0,00$), las cuales se indican con las flechas negras en la Figura 1, están muy cerca entre sí. En este sentido, se puede observar que la prueba no resultó ser ni demasiado fácil ni demasiado difícil para el estudiantado, lo cual implica que los ítems son apropiados para esta población.

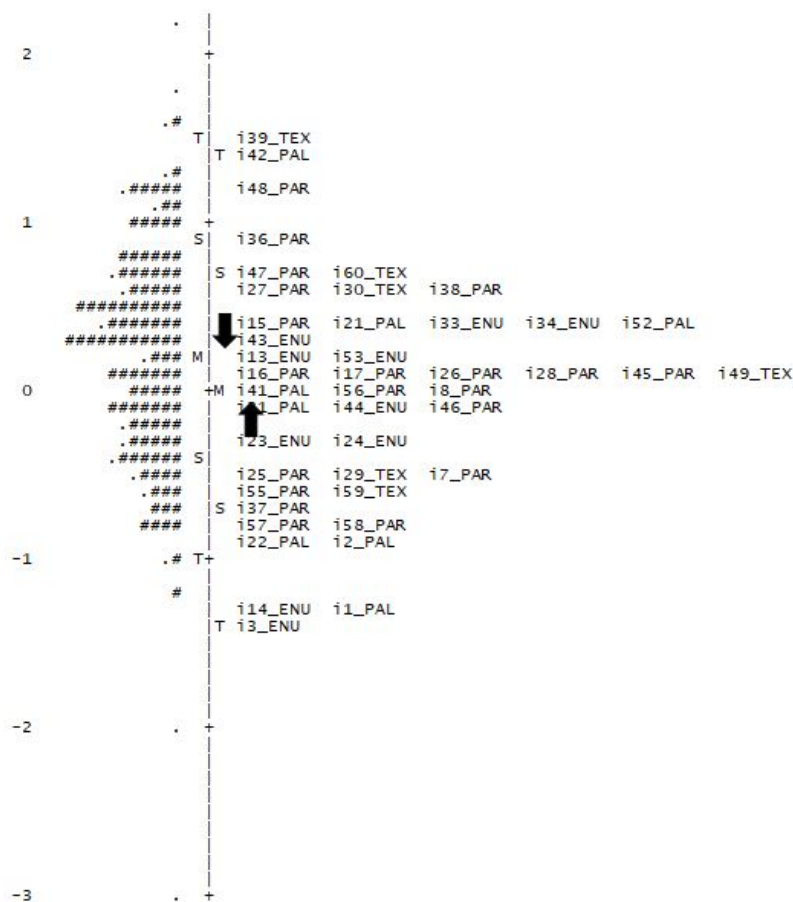


FIGURA 1
Mapa de ítems y personas
Fuente: Elaboración propia.

Finalmente, de los 244 que contestaron los ítems, fue posible obtener la nota final que 239 de ellos obtuvieron en un curso de primer año de la carrera de Derecho de la Universidad de Costa Rica. El coeficiente de correlación de Pearson observado entre la nota final en dicho curso y la puntuación total en la prueba (tomando en cuenta solamente los 44 ítems analizados previamente) fue de 0,21, estadísticamente significativo ($p < 0,000$) con un intervalo de confianza al 95% entre 0,09 y 0,33. Este valor positivo y moderado indica que el test podría ser de utilidad para anticipar el nivel de rendimiento académico de la población meta.

DISCUSIÓN

El objetivo del presente estudio fue proporcionar evidencias de validez y confiabilidad sobre un test nuevo para medir el nivel de comprensión lectora de estudiantes universitarios. Con base en los resultados obtenidos, esta prueba ha brindado resultados favorables los cuales señalan su utilidad para detectar estudiantes de primer ingreso a la educación superior que podrían presentar problemas de comprensión lectora y, por ende, ver disminuido su rendimiento académico.

En este sentido, este instrumento ayudaría a mejorar los niveles de comprensión lectora del estudiantado perteneciente a diferentes áreas del saber en el ámbito de la educación universitaria. Por ejemplo, podría

emplearse como una medición inicial que permita determinar si algunos estudiantes requieren ciertos refuerzos en el área de la lectura para afrontar con éxito los cursos propios de la carrera.

Asimismo, dado que fue construida con textos tanto de ciencias sociales como de ciencias naturales, el test podría emplearse como diagnóstico para ingresar a diversas carreras de la universidad. En este sentido, la exigencia de contar con un alto nivel de habilidad lectora es una exigencia de cualquier carrera universitaria, por lo que quienes obtengan bajas puntuaciones en esta prueba podrían presentar problemas de rendimiento académico en el futuro. Adicionalmente, el constructo subyacente a la prueba se basa en característica lingüísticas (palabras, oraciones, párrafos y texto) transversales a prácticamente la totalidad de los libros que usualmente se emplean en el ámbito académico (libros de texto, manuales, artículos científicos, ensayos, monografías, entre otros), por lo cual es posible incluir una gran variedad de textos con el propósito de adaptarla a diversos ámbitos disciplinares.

Otros usos posibles de esta prueba son aquellos relacionados con la investigación empírica relacionada con las habilidades verbales de estudiantes universitarios. En este sentido, las puntuaciones obtenidas en la prueba podrían combinarse con otros instrumentos de evaluación para generar modelos empíricos sobre la relación entre la lectura y otras variables de interés (oportunidades educativas, motivación, conocimientos especializados, entre otros).

CONCLUSIÓN

Como se pudo observar, los resultados dan evidencia empírica a favor de una estructura factorial unidimensional con un nivel de confiabilidad moderado. Dado que esta dimensión muestra un nivel de precisión relativamente moderado, la prueba puede emplearse como una herramienta de carácter diagnóstico en relación con las habilidades de comprensión lectora de estudiantes de primer ingreso a la educación superior. La prueba podría servir para propósitos de *screening*, es decir, para coadyuvar en la elaboración de planes remediales, implementación de intervenciones educativas y desarrollo de unidades didácticas orientadas al fortalecimiento de la comprensión lectora de los(as) estudiantes universitarios de primer ingreso.

Dadas sus características, esta prueba es relativamente corta en términos de la cantidad de ítems (44) y del tiempo requerido para su resolución (aproximadamente una hora y treinta minutos). Por ello, es un instrumento complementario que no pretende ser el único indicador requerido para evaluar una capacidad tan compleja como lo es la comprensión lectora. En este sentido, la prueba ha mostrado algunas propiedades deseables mínimas las cuales la convierten en una herramienta de *screening* para detectar a estudiantes que en principio podrían necesitar un acompañamiento más intensivo para enfrentar las demandas de lectura propias de la educación universitaria.

En lo que respecta a las recomendaciones que se derivan de este estudio, cabe señalar que este constituye un primer paso para posteriores mejoras en el proceso de validación de esta prueba. En este sentido, es necesario incluir a una mayor cantidad de estudiantes pertenecientes a carreras de otros ámbitos disciplinares para poder generalizar su utilidad. Si se toma en cuenta la fórmula de Spearman-Brown (Muñiz, 2018), sería necesario añadir aproximadamente 18 ítems para aumentar la confiabilidad de las puntuaciones al nivel generalmente aceptado para este tipo de pruebas (0,80). Finalmente, es importante incorporar en futuros estudios de validación evidencias de validez relacionadas con la relación de la prueba con otros instrumentos que evalúen comprensión lectora.

REFERENCIAS BIBLIOGRÁFICAS

Alderson, J. C. (2000). *Assessing Reading*. Reino Unido: Cambridge University Press.

- Alexander, P. A. y Fox, E. (2013). A Historical Perspective on Reading Research and Practice, Redux. En D. E. Alvermann, N. J. Unrau y R. B. Ruddell (Eds.), *Theoretical Models and Processes of Reading* (pp. 3-46). Estados Unidos: International Reading Association.
- Alvermann, D. E., Unrau, N. J. y Ruddell, R. B. (2013). *Theoretical Models and Processes of Reading*. Estados Unidos: International Reading Association.
- Bachman, L. F. (2007). What is the construct? The dialectic of abilities and contexts in defining constructs in language assessment. En J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E. Turner y C. Doe (Eds.), *Language Testing Reconsidered* (pp. 41-71). Canadá: University of Ottawa Press.
- Berko, J., Vellutino, F. y Wolf, M. (1999). Una explicación psicolingüística de la lectura. En J. Berko y B. Bernstein (Eds.), *Psicolingüística* (pp.433-475). España: MacGraw Hill.
- Camps, A. y Colomer, T. (1996). *Enseñar a leer, enseñar a comprender*. España: Celeste/M.E.C.
- Cassany, D. (2006). *Tras las líneas*. España: Anagrama.
- Cerdán, R., Vidal, E., Martínez, T., Gilabert, R. y Gil, L. (2009). Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction*, 19(1), 13-27.
- Cruz, J. (2018). Tests estandarizados de comprensión lectora [Mensaje en un blog]. *Comprensión lectora basada en evidencias*. Recuperado de <https://clbe.wordpress.com/tests-estandarizados-de-comprension-lectora/>
- Davis, F. B. (1944). Fundamental factors of comprehension in reading. *Psychometrika*, 9(3), 185-197.
- Day, R. y Bamford, J. (1998). *Extensive reading in the second language classroom*. Estados Unidos: Cambridge University Press.
- Farr, R., Pritchard, R. y Smitten, B. (1990). A description of what happens when an examinee takes a multiple - choice reading comprehension test. *Journal of Educational Measurement*, 27(3), 209-226.
- Fletcher, J. M. (2006). Measuring reading comprehension. *Scientific Studies of Reading*, 10(3), 323-330.
- Gutiérrez, M. (1999). Inferencias en la comprensión del lenguaje. En M. Vega y F. Cuetos (Eds.), *Psicolingüística del español* (pp. 231-270). España: Editorial Trotta.
- Jackson, N. E. (2005). Are university students' component reading skills related to their text comprehension and academic achievement? *Learning and Individual Differences*, 15(2), 113-139.
- Keenan, J. M., Betjemann, R. S. y Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, 12(3), 281-300.
- Kendeou, P., van den Broek, P., Helder, A. y Karlsson, J. (2014). A cognitive view of reading comprehension: implications for reading difficulties. *Learning Disabilities Research & Practice*, 29(1), 10-16.
- Kintsch, W. (2013). Revisiting the Construction-Integration Model of Text Comprehension and Its Implications for Instruction. En D. E. Alvermann, N. J. Unrau y R. B. Ruddell (Eds.), *Theoretical Models and Processes of Reading* (pp. 807 - 839). Estados Unidos: International Reading Association.
- Landi, N. (2010). An examination of the relationship between reading comprehension, higher level and lower-level reading sub-skills in adults. *Reading and writing*, 23(6), 701-717.
- Leighton, J. P. y Gierl, M. J. (2011). *The learning sciences in educational assessment: the role of cognitive models*. Estados Unidos: Cambridge University Press.
- Leslie, L. y Caldwell, J. (2009). Formal and Informal Measures of Reading Comprehension. En S. E. Israel y G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 403-427). Estados Unidos: Routledge.
- Linacre, J. M. (2018). *Winsteps® Rasch measurement computer program*. Beaverton, Oregon: Winsteps.com
- Lloréns, A. C., Gil, L., Abarca, E.V., Martínez, T., Mañá, A. y Gilabert, R. (2011). Prueba de Competencia Lectora para Educación Secundaria (CompLEC). *Psichotema*, 23(4), 808-817.
- López, M. (2010). *Estrategias cognitivas del aprendizaje y comprensión lectora en estudiantes de quinto grado de secundaria de una institución educativa del Callao* [Tesis de maestría]. Universidad de Costa Rica, San José. Recuperado de <https://bit.ly/30rlmdX>

- Magliano, J. P., Millis, K. K., Ozuru, Y. y McNamara, D. S. (2007). A multidimensional framework to evaluate reading assessment tools. En D. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 107-136). Estados Unidos: Lawrence Erlbaum Associates.
- McNamara, D. y Magliano, J. (2009). Towards a comprehensive model of comprehension. En B. Ross (Ed.), *The psychology of learning and motivation* (pp. 297-384). Estados Unidos: Academic Press.
- Montero, E., Rojas, S., Zamora, E. y Rodino, M. (2011). *Costa Rica en las pruebas PISA 2009 de Competencia Lectora y Alfabetización Matemática*. San José, Costa Rica: CONARE. Recuperado de <https://bit.ly/2JNLrOs>
- Muñiz, J. (2018). *Introducción a la psicometría / Teoría clásica y TRI*. España: Ediciones Pirámide.
- O'Reilly, T., Weeks, J., Sabatini, J., Halderman, L. y Steinberg, J. (2014). Designing reading comprehension assessments for reading interventions: how a theoretically motivated assessment can serve as an outcome measure. *Educational Psychology Review*, 26(3), 403-424.
- Organización para la Cooperación y el Desarrollo Económico. (2000). Programme for International student Assessment. Reading, mathematical and scientific literacy. París: OECD.
- Ozuru, Y., Rowe, M., O'Reilly, T. y McNamara, D. S. (2008). Where's the difficulty in standardized reading tests: The passage or the question? *Behavior Research Methods*, 40(4), 1001-1015.
- Pérez, M. J. (2005). Evaluación de la comprensión lectora: Dificultades y limitaciones. *Revista de Educación*, 1, 121-138.
- Perfetti, C. y Adlof, S.M. (2012). Reading Comprehension: A Conceptual Framework from Word Meaning to Text Meaning. En J. P. Sabatini, E. R. Albry y T. O'Reilly, *Measuring Up: advances in how we asses Reading ability* (pp. 3-20). Estados Unidos: Rowman & Littlefield Education.
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Recuperado de <http://www.R-project.org/>
- Reyes, S., Castillo, A., Zúñiga, A. y Llarena, R. (2009). *La evaluación objetiva de la habilidad lectora en la educación media superior*. México: CENEVAL. Recuperado de <https://bit.ly/2RH0dZw>
- Rupp, A. A., Ferne, T. y Choi, H. (2006). How assessing reading comprehension with multiple-choice questions shapes the construct: A cognitive processing perspective. *Language testing*, 23(4), 441-474.
- Solé, I. (2009). *Estrategias de lectura*. España: Editorial Graó.
- Svetina, D., Gorin, J. S. y Tatsuoaka, K. K. (2011). Defining and comparing the reading comprehension construct: A cognitive-psychometric modeling approach. *International Journal of Testing*, 11(1), 1-23.
- Van den Broek, P. (2012). Individual and Developmental Differences in Reading Comprehension: Assessing Cognitive Processes and Outcomes. En J. P. Sabatini, E. R. Albry y T. O'Reilly, *Measuring Up: advances in how we asses Reading ability* (pp. 39 - 57). Estados Unidos: Rowman & Littlefield Education.
- Willse, J. (2018). CTT / Classical Test Theory Functions. Recuperado de <https://CRAN.R-project.org/package=CTT>
- Wright, B. D. y Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3). Recuperado de <https://www.rasch.org/rmt/rmt83b.htm>