

DISEÑO Y VALIDACIÓN DE UNA PRUEBA DE COMPRENSIÓN LECTORA INFERENCIAL PARA ESTUDIANTES UNIVERSITARIOS

**ALEJANDRA PLATAS GARCÍA, VERÓNICA REYES MEZA Y
J. MARTÍN CASTRO MANZANO**

UPAEP

*alejandra.platas@upaep.edu.mx, veronica.reyesm@uatx.mx y
josemartin.castro@upaep.mx*

Resumen

Típicamente, los docentes universitarios se ven en la necesidad de medir la habilidad de comprensión lectora inferencial de los estudiantes a través de pruebas. Esto porque tal habilidad representa una tarea fundamental en la formación académica y porque las pruebas podrían servir como estrategias de enseñanza. En este trabajo presentamos el proceso de diseño de una prueba de comprensión lectora inferencial de opción múltiple por tipos de razonamiento (deductivo, inductivo y abductivo) en la que consideramos el efecto de la complejidad y la longitud del texto; asimismo, presentamos el proceso de validación de la prueba en tres etapas: la primera, en la que revisamos la validez del contenido a partir de un juicio de expertos; la segunda, en la que revisamos la viabilidad a través de la aplicación de la prueba a una muestra de 39 universitarios de áreas académicas diversas; y la tercera, en la que revisamos la confiabilidad de la prueba a través de un análisis de reactivos dados los resultados obtenidos en la aplicación. Concluimos describiendo el aporte de este tipo de pruebas en la formación de los universitarios y resaltando la necesidad de un trabajo futuro que considere las estrategias que utilizaron los estudiantes al resolverla.

Palabras clave: Evaluación, razonamiento, validez, confiabilidad.

Abstract

College teachers usually need to measure their undergraduate students' inferential comprehension skill by using tests. This is so because such a skill is fundamental for the academic training and because the used tests might serve as teaching strategies. In this paper, we present the design process of a multiple choice test for measuring reading inferential comprehension by taking into account three types of inference (deductive, inductive, and abductive). The effect of complexity and length of the questions within the test is also studied. Then, we present the validation process of the test in three stages: first, we check for validity by expert judgement; second, we check for feasibility by applying the test to a sample of 39 undergraduate students of various academic programs; third, we check for reliability through item analysis. Finally, we describe some potential features of this sort of tests in the academic training and we highlight future work regarding test taking strategies.

Keywords: Non-formal education, evaluation, social programs.

Introducción

La presente propuesta surge del contexto de la enseñanza de lenguas extranjeras (italiano) a estudiantes universitarios. En este ámbito se realizan pruebas de opción múltiple para medir la comprensión lectora con la finalidad de saber si los estudiantes son capaces de reconocer el enunciado que refleja el contenido de un texto, leído previamente, de entre un conjunto de afirmaciones. Manteniendo la misma finalidad y utilizando el mismo formato de opción múltiple, en este trabajo presentamos el proceso de diseño y validación de una prueba de comprensión lectora inferencial con características específicas. La primera característica tiene que ver con el idioma que elegimos para nuestra prueba, a saber, el Español. Hemos optado por el Español porque los estudiantes a los que va dirigida esta prueba son mexicanos y su lengua materna es el Español.

La segunda característica tiene que ver con la elección del nivel de comprensión lectora. Hemos decidido utilizar el inferencial porque es un nivel de mayor estructura, en comparación con un nivel literal. Gagné, citado por Jiménez (2004), distingue cuatro procesos implicados en la lectura, a saber: decodificación, comprensión literal, comprensión inferencial y control de la comprensión. De manera similar, García (1993) distingue cuatro procesos o niveles en la comprensión lectora, a saber: decodificación, la comprensión literal, la comprensión inferencial y la metacompreensión. En este trabajo, hemos seguido la clasificación de García (1993), para quien en su modelo de procesamiento y estrategias de mejora de la comprensión de textos, la elaboración de inferencias forma parte del procesamiento textual, el cual hace referencia a las representaciones y procesos que se dan a nivel textual y comprende la integración de las proposiciones del texto, el conocimiento del mundo por parte del lector, las inferencias según los esquemas cognitivos que el sujeto posee y la interpretación del texto.

La tercera característica tiene que ver con la decisión del nivel educativo al que va dirigida la prueba. Hemos elegido a estudiantes de nivel superior, es decir, universitarios, por dos razones: i) existen problemas de comprensión lectora a ese nivel, Guerra y Guevara (2013) afirman que tales problemas residen en la dificultad de identificar “la o las ideas principales y en el uso apropiado de la información para hacer inferencias y realizar un análisis crítico sobre la base de lo leído” (p. 278); y ii) porque “no se conocen pruebas estandarizadas que midan comprensión lectora de tipo inferencial a nivel universitario” (Velásquez, Cornejo y Roco, 2008, p. 124).

La cuarta característica tiene que ver con la elección de la tipología textual. Hemos optado por textos de tipo descriptivo, los cuales corresponden a descripciones internas en textos narrativos, expositivos y técnico-científicos, de acuerdo a la clasificación de Barni et al. (2009), porque esta clase de textos es común para los universitarios en cualquier

área académica a la que pertenezcan. Para este último criterio utilizamos como parámetro la *Certificazione di Italiano come Lingua Straniera: CILS* (Certificación de Italiano como lengua extranjera), la cual acredita el grado de dominio del idioma y es otorgada por la *Università per Stranieri di Siena*.

En otras palabras, seguimos el formato que se utiliza para evaluar la comprensión lectora en exámenes de certificación internacional de una lengua extranjera, donde los textos son de diferentes temáticas y de distintas tipologías, como la descriptiva. Buscamos, entonces, fragmentos de textos descriptivos como los que suelen estar presentes en artículos de investigación científica, artículos de enciclopedia o textos escolares, sabiendo que la descripción “rara vez constituye por sí sola un texto completo, sino que forma segmentos que gozan de cierta autonomía dentro de textos más amplios” (Álvarez, 2005, p. 67). Nuestro trabajo se articula en cuatro secciones: en la primera, exponemos los fundamentos teóricos de nuestra propuesta; en la segunda, describimos el proceso de diseño; en la tercera, describimos el proceso de validación; y por último, en la cuarta, presentamos nuestras conclusiones.

Antecedentes

A continuación exponemos antecedentes teóricos de nuestra propuesta: comprensión lectora, medición, pruebas de comprensión, diseño de pruebas. Entendemos por comprensión lectora, “un proceso a través del cual el lector elabora un significado en su interacción con el texto” (Pérez Zorrilla, 2005, p. 123). En esta interacción influyen tanto las características del lector, como las del texto. Sobre estas últimas características, afirma Zanutto (2007) que:

Un aspecto fundamental del enfoque de van Dijk y Kintsch corresponde al hecho de que, si bien se considera que el lector tiene un papel activo en el proceso de comprensión del texto, también el texto cuenta con una determinada estructura que influye en dicha comprensión, y que no permite la construcción de cualquier tipo de significado. (p. 42)

Los tipos de estructuras que se pueden encontrar en los textos, de acuerdo a van Dijk (Zanutto, 2007) son tres: a) La superestructura que determina el orden y la coordinación global de las partes, es un esquema al que se adapta el texto. b) La macroestructura que se forma a partir de las ideas principales del texto y corresponde a la representación abstracta de la estructura global del significado de un texto. c) La microestructura que corresponde a la coherencia semántica entre cada una de las proposiciones que componen al escrito.

García (1993) distingue cuatro niveles ordenados de comprensión lectora, que van desde un nivel básico a un nivel estructurado, a saber: decodificación (asociación de palabras

escritas con significados disponibles en la memoria o transformación de letras impresas en sílabas y en sonidos para activar significados); comprensión literal (combinación de significados de varias palabras de forma apropiada para formar proposiciones, dada la información explícitamente presentada en un texto); comprensión inferencial (elaboración de representaciones mentales más integradas y esquemáticas a partir de la información presentada y los conocimientos previos; que permite la extracción de información que no está explícita en el texto); y metacompreensión (conciencia y control que el lector tiene de su propio proceso de comprensión).

En la práctica educativa es necesario conocer el nivel de comprensión lectora porque, por un lado, para un estudiante universitario, la comprensión de textos es una tarea fundamental en su formación académica; y porque, por otro lado, para un docente universitario, la medición de la comprensión lectora ofrece información importante para decidir si hay necesidad de implementar cambios en los procesos de aprendizaje y enseñanza de la comprensión lectora. Brown y Glasner (2007) lo refieren de la siguiente manera:

Quando la evaluación se realiza correctamente, puede ser motivadora y productiva para los estudiantes, ya que les ayuda a saber si lo que están haciendo es correcto o si necesitan hacer algo más. Por otro lado, a los profesores les permite conocer mejor su tarea y les proporciona los indicadores de actuación que necesitan. (p. 24)

Asimismo, la evaluación a través de pruebas de comprensión lectora podría contribuir a resolver el problema que refieren Vidal-Moscoso y Manríquez-López (2016), de que la institución universitaria asume que el estudiante, al ingresar, conoce la cultura académica de la que formará parte y sabe analizar lo que lee, realizar inferencias, ofrecer argumentos coherentes cuando "lo cierto es que los estudiantes que ingresan a la universidad son inexpertos en la lectura que se les exige" (p. 98).

Ahora bien, un aspecto a considerar en esta medición corresponde al hecho de que la comprensión lectora, sobre todo a nivel inferencial, no se puede llevar a cabo con una observación inmediata; al respecto afirma Pérez Zorrilla (2005): "dado que no se puede observar directamente la comprensión lectora, hay que pedirle al alumno que realice algún tipo de tarea que indique el grado de comprensión alcanzado" (p. 128). Por lo anterior, la comprensión lectora se tiene que medir a través de pruebas y estas pruebas deben reunir ciertas características esenciales, a saber: validez que midan correctamente el conocimiento o las destrezas que pretende medir, confiabilidad que proporcionen consistencia en los resultados y viabilidad que sean de fácil aplicación, (Palencia del Burgo, 1990).

Existen distintos tipos de pruebas, pero si nos detenemos a considerar únicamente las pruebas de opción múltiple, podemos notar que estas reúnen las características anteriores: pueden alcanzar niveles adecuados de validez y confiabilidad, además, suelen ser

viabiles porque requieren de medios accesibles para su aplicación. Hay una escasa producción de instrumentos válidos y confiables para medir la comprensión lectora (Guerra y Guevara, 2013). Esto genera un área de oportunidad para diseñar pruebas, además de que estas pueden verse como estrategias de aprendizaje que podrían ayudar a mejorar el desempeño de los estudiantes en esta habilidad.

Entre la producción de pruebas de comprensión lectora dirigidas a universitarios, se pueden citar tanto el trabajo de Velásquez et al. (2008), quienes introdujeron una prueba de opción múltiple; como el de Guerra y Guevara (2013), quienes elaboraron una prueba que incluía reactivos de opción múltiple y de respuesta abierta; ambas pruebas se validaron bajo juicio de expertos. Está, además, el trabajo de Sánchez, Grajales y García (2011), quienes propusieron una prueba de opción múltiple validada estadísticamente.

Con respecto al diseño de pruebas de comprensión lectora inferencial dirigidas a universitarios; en un estudio nuestro (Platas-García, Castro-Manzano y Reyes-Meza, 2016), se diseñó una propuesta a partir del establecimiento de una analogía entre el concepto de inferencia y la estructura de las pruebas de opción múltiple que se explica a continuación.

En primer lugar, es conveniente aclarar lo que se entiende por inferencia. En términos lógicos, una inferencia es un proceso en el que un agente razonador parte de un conjunto finito de datos, premisas, para llegar a un nuevo dato, conclusión. (Cook, 2009).

García (1993), por su parte, caracteriza las inferencias como:

Actividades cognitivas mediante las cuales el lector adquiere informaciones nuevas a partir de informaciones disponibles. (...) Las inferencias permiten, por una parte, establecer relaciones entre los diversos elementos del texto y, por otra, integrar la información del texto con los conocimientos previos y esquema cognitivo del lector. (p. 98).

Con respecto a la estructura de una prueba de opción múltiple, esta se compone de un texto a partir del cual se genera una pregunta y se presentan varias opciones de respuesta en las que sólo una opción es la correcta. La analogía, entonces, es la siguiente: las premisas (o conjunto de datos), en una inferencia, correspondería al texto en una prueba de opción múltiple; el indicador de consecuencia (con palabras como: entonces, luego, por lo tanto), en una inferencia, correspondería a la pregunta en la prueba de opción múltiple; y finalmente, la conclusión (el dato nuevo) en una inferencia, correspondería a la respuesta correcta dentro del conjunto de opciones en la prueba de opción múltiple. A este diseño de pruebas lo llamaremos "Diseño por tipos de razonamiento" en el que el razonamiento es visto como un proceso que produce información nueva a partir de datos previamente dados, y la inferencia, podría considerarse como la unidad de medi-

da del razonamiento. Los razonamientos se pueden clasificar en tres tipos: deductivo, inductivo y abductivo.

La clasificación por “tipos de razonamiento” se realiza desde una perspectiva lógica; sin embargo, existen otras clasificaciones de las inferencias; por ejemplo, Otero, Graesser y León (como se citó en Zanotto, 2007) refieren la clasificación de “tipos cognitivos” que incluyen cinco categorías, a saber: inferencias anafóricas, inferencias puente, inferencias explicativas, inferencias elaborativas, e inferencias predictivas. Los tipos de razonamiento se explican de la siguiente manera.

Una deducción se define como una secuencia finita de enunciados dentro de un sistema formal donde cada enunciado en la secuencia es un axioma, una suposición o el resultado de la aplicación de una regla de inferencia a uno o varios enunciados precedentes; el enunciado final es la conclusión del argumento que ha sido derivado y cada suposición es una premisa del argumento derivado (Cook, 2009). La deducción supone, por tanto, que el conjunto de información está completo y que, por ende, la conclusión no agrega algo diferente a lo que ya está en las premisas (Rodríguez Rodríguez, 2005).

Una inducción es una inferencia que permite extraer conclusiones con cierto grado de soporte (Hawthorne, 2014); el grado de soporte que proveen las premisas ofrece la base conceptual que sostiene a la conclusión, pero este grado de soporte no es el deductivo porque no implica criterios de necesidad, sino de probabilidad.

Finalmente, una abducción es un proceso de razonamiento cuyo producto son explicaciones con cierta estructura inferencial que se dispara por un hecho que necesita explicación (Aliseda, 1998). Como la inducción, la abducción no tiene criterios de necesidad sino de probabilidad. Navarro (2012), por su parte, sostiene que la abducción “es una forma de razonamiento que parte de un conjunto de datos para concluir una hipótesis que explica plausiblemente ese conjunto de datos” (p. 237).

Proceso de diseño de la prueba

La primera actividad que realizamos, antes de comenzar la tarea de la redacción de los reactivos, fue acordar las características para nuestra prueba que se presentan a continuación.

- **Opción múltiple.** Con 4 posibilidades de respuesta (como el examen CILS) y una sola respuesta correcta (para garantizar fiabilidad); además, se redactan las justificaciones de las (in)correcciones de cada respuesta (para garantizar validez).
- **Tipos de razonamiento.** El diseño de la prueba de comprensión lectora por tipos de razonamiento se refleja en la pregunta de cada reactivo (el indicador de consecuencia de una inferencia); de tal manera que las preguntas para el tipo

de razonamiento deductivo se redactan de una forma similar a este ejemplo: “De acuerdo al texto anterior, qué se concluye necesariamente”; las preguntas para el tipo de razonamiento inductivo se redactan de una forma similar a este ejemplo: “De acuerdo al texto anterior, qué es lo más probable que suceda en el futuro”; y las preguntas para el tipo de razonamiento abductivo se redactan de una forma similar a este ejemplo: “De acuerdo al texto anterior, qué explicaría mejor la situación”.

- **Longitud del texto:** corto o largo. Un texto de 1 a 4 líneas es un texto corto y un texto de 5 a 8 líneas es un texto largo. Llegamos a estas medidas dividiendo entre dos el número de líneas del texto más largo de una prueba que habíamos elaborado previamente (Platas García et al., 2016).
- **Complejidad del texto: simple o complejo.** Después de haber elegido los diferentes textos descriptivos (de temáticas distintas), le solicitamos a siete lectores expertos (escritores, editores, docentes de español para extranjeros) que leyeron los textos que habíamos seleccionado, con la finalidad de establecer si, de acuerdo a su juicio, estos eran simples o complejos. Los criterios utilizados para tal clasificación fueron los siguientes: verbos, vocabulario, redacción, extensión y tema.

A partir de las características anteriores obtuvimos el número de reactivos que tendría nuestra prueba, esto es, 12 reactivos, correspondientes al producto de 3 tipos de razonamiento, por 2 tipos de longitud, por 2 tipos de complejidad. Al final, para cada reactivo, elaboramos una pregunta con cuatro opciones de respuesta (una sola correcta).

Proceso de validación de la prueba

El proceso de validación de la prueba se dividió en tres etapas: la primera, tuvo como propósito revisar la validez del contenido a partir de un juicio de expertos; la segunda, tuvo como propósito revisar la viabilidad a través de la aplicación de la prueba a una muestra de universitarios; y la tercera, tuvo como propósito revisar la confiabilidad de la prueba a través de un análisis de reactivos a partir de los resultados obtenidos en la aplicación.

Para la primera etapa, el procedimiento fue el siguiente: se informó a los jueces cuáles eran las características de la prueba y los aspectos que se pretendían medir. Posteriormente, se proporcionó una prueba a cinco lectores expertos que a partir de su formación académica y experiencia profesional están habituados a leer y analizar textos a nivel inferencial (estos expertos fueron distintos de los siete lectores referidos en el apartado sobre la característica de Complejidad del texto). Ellos leyeron la prueba y la respondieron de forma escrita. Posteriormente, en entrevista con tres de ellos, tuvimos retroalimenta-

ción de sus respuestas, observaciones y comentarios. Los jueces analizaron el formato de la prueba, el contenido de los textos, la pregunta de cada reactivo, la redacción de las opciones de respuesta, y su propia respuesta.

Después, discutimos sus respuestas; nuestro criterio consistió en observar si dos o más expertos estaban de acuerdo en que algún reactivo debería mejorarse y en ese caso, lo cambiaríamos. Esto sucedió con cinco reactivos de los doce que tiene la prueba. Finalmente, realizamos los cambios en la prueba antes de proceder a la aplicación. Las modificaciones se aplicaron a la redacción de las opciones de respuesta, con la intención de garantizar claridad y evitar más de una respuesta correcta en cada reactivo.

Para la segunda etapa, el procedimiento fue el siguiente: obtuvimos los permisos para aplicar la prueba en una universidad dentro del aula de clase en un horario de 8.00 a.m. a 12.30 p.m. Los estudiantes firmaron un documento de consentimiento para proporcionar sus datos y responder a la prueba; asimismo, se les informó de la confidencialidad con la que se conservarían sus datos personales. Los participantes fueron 39 estudiantes universitarios, 26 mujeres y 13 hombres, con una edad promedio de 20.2 años, pertenecientes a distintas áreas académicas. Esta muestra fue seleccionada al azar.

Después, observamos las condiciones de viabilidad de la prueba y registramos que su aplicación de forma grupal fue viable, las instrucciones fueron claras, el tiempo máximo empleado para concluir la prueba fue de 30 minutos.

Por último, analizamos el desempeño de los estudiantes en la prueba. Los resultados fueron los siguientes: el máximo de respuestas correctas por persona fue de 11 y el mínimo de respuestas correctas fue de 4. En total, las respuestas correctas de los estudiantes correspondieron al 61%; este porcentaje clasificado por tipo de razonamiento tuvo el siguiente orden: inducción (29.5%), abducción (20%) y deducción (11.5%); esto significa que los estudiantes están más habituados a realizar razonamientos inductivos que deductivos, ya que la inducción posibilita la solución de problemas en la vida cotidiana, esto es interesante porque la analogía, que es un tipo de argumento inductivo, "es la base de la mayoría de nuestros razonamientos ordinarios que van de la experiencia pasada a lo que sucederá en el futuro" (Copi y Cohen, 2007, p. 444).

Para la tercera etapa, el procedimiento fue el siguiente. Realizamos un análisis de reactivos de las respuestas obtenidas en la aplicación, esto siguiendo el procedimiento de la prueba CILS (Barni et al. 2009) y consideramos tres aspectos: a) revisión de las opciones elegidas (para conocer la frecuencia), b) revisión del índice de dificultad de los reactivos (para conocer cuáles fueron los reactivos fáciles o difíciles), c) revisión del índice de discriminación de los reactivos (para discriminar a los candidatos competentes de los menos competentes). Los resultados se presentan en la Tabla 1 a continuación.

TABLA 1. Resultados de los análisis de los reactivos

	FRECUENCIA DE ELECCIÓN DE LAS OPCIONES				ÍNDICE DE DIFICULTAD (Di)	ÍNDICE DE DISCRIMINACIÓN (Di)
	OPCIÓN A	OPCIÓN B	OPCIÓN C	OPCIÓN D		
Pregunta-1	5	*29	4	1	0.74	0.50
Pregunta-2	1	*14	11	13	0.35	0.30
Pregunta-3	2	1	35	1	0.89	0.40
Pregunta-4	16	*3	17	3	0.07	0.20
Pregunta-5	3	1	0	*35	0.89	0.30
Pregunta-6	*33	1	5	0	0.84	0.40
Pregunta-7	11	4	*20	4	0.51	0.30
Pregunta-8	*14	5	20	0	0.35	0.60
Pregunta-9	0	0	*38	1	0.97	0.10
Pregunta-10	6	1	3	*29	0.74	0.60
Pregunta-11	0	1	*34	4	0.84	0.10
Pregunta-12	21	*9	3	6	0.23	0.10

Nota: La respuesta correcta se indica con el asterisco. Si un Di es <0.40 se considera difícil; y si su valor está entre 0.91 y 1.0 se considera fácil. Si un Di es >0.39 su calidad se considera excelente y se recomienda conservar; si está entre 0.30 a 0.39, su calidad es buena y se puede mejorar; si está entre 0.20 y 0.29 su calidad es regular y se necesita revisar y, si el Di está entre 0.00 y 0.19 su calidad se considera pobre y se debe revisar a profundidad (Díaz y Leyva, 2013). Fuente: Elaboración propia.

- a. Con respecto a la revisión de las opciones elegidas, obtuvimos la frecuencia en la que cada opción de respuesta fue seleccionada. Los resultados muestran que en los reactivos 5, 6, 8, 9 y 11 una o dos opciones no fueron elegidas por ningún estudiante, por lo cual convendría revisarlas.
- b. Con respecto a la revisión del índice de dificultad de los reactivos (Di) dividimos el número de respuestas correctas de cada reactivo (A) entre el número total de estudiantes que contestaron el reactivo (N=39), la fórmula es: $Di=A/N$ (Díaz y Leyva, 2013). Los resultados muestran que tenemos cuatro reactivos difíciles (2, 4, 8 y 12) con un Di de <0.40 y un reactivo fácil (el 9) con un Di de 0.91; estos reactivos necesitarían una revisión, especialmente para reconsiderar la clasificación propuesta de simple-complejo.

- c. Con respecto al índice de discriminación de los reactivos (D_i), ordenamos de mayor a menor las respuestas correctas. Después, dividimos la muestra en dos grupos correspondientes al 27%; de esta forma tuvimos dos extremos de 10 estudiantes: un extremo superior y un extremo inferior. Le restamos al número de correctas en el reactivo i del 27% de personas con las puntuaciones más altas en la prueba (GA) el número de correctas en el reactivo i del 27% de personas con las puntuaciones más bajas en la prueba (GB) y lo dividimos entre el número de personas en el grupo más numeroso (GA o GB); siendo la fórmula: $D_i = (GA - GB) / N$ (Díaz y Leyva, 2013). Los resultados muestran que tenemos 3 reactivos de calidad pobre (9, 11 y 12) con valores entre 0.00 y 0.19 y por lo tanto, tendremos que revisarlos a profundidad.

En suma, hasta este momento, nuestra prueba tiene validez de contenido, gracias al juicio de expertos, como lo han hecho para validar sus pruebas autores como Velásquez et al. (2008) y Guerra y Guevara (2013); sin embargo, para la siguiente versión de nuestra prueba, en la que se realizarán las modificaciones necesarias, sería favorable aplicarla a una muestra más representativa de estudiantes universitarios. Asimismo, sería conveniente obtener un índice de confiabilidad del instrumento mediante el coeficiente de Kuder Richardson (KR-20) y una validez más precisa con el desarrollo de un análisis de factores.

Conclusiones

El uso de pruebas de comprensión lectora inferencial por tipos de razonamiento podría contribuir a la formación general de los universitarios, independientemente del área académica a la que pertenezcan, porque al realizar actividades guiadas que requieran el ejercicio de los distintos tipos de razonamiento (a través de los reactivos), los estudiantes llevarían a cabo un proceso de aprendizaje crítico y reflexivo. Esto es compatible con los resultados presentados en Platas-García et al. (2016).

Un aspecto esencial, que no consideramos en este trabajo, pero que hemos investigado y que influye en el diseño y validación de una prueba de comprensión lectora inferencial de opción múltiple, corresponde a las estrategias que eligen los universitarios al momento de resolver este tipo de pruebas (Platas García et al., 2016; Platas-García, Castro-Manzano y Reyes-Meza, 2017).

Esto último ha sido referido por Andrew Cohen (1984), quien explica por un lado, que es importante esforzarse por buscar una adecuación entre la intención que tienen los diseñadores de pruebas para que estas se resuelvan y el cómo, de hecho, los estudiantes

las resuelven (p. 79); y por otro lado, que la investigación empírica en las estrategias para la resolución de pruebas (test-taking strategies) puede proporcionar información valiosa sobre lo que las pruebas están realmente midiendo (Cohen, 2007).

Por lo tanto, una investigación futura podría analizar las justificaciones de las (in)correcciones de las respuestas en el diseño de pruebas de comprensión lectora de esta naturaleza y las justificaciones de las respuestas de los estudiantes, con la intención de verificar si la comprensión lectora inferencial que se quería medir a través de la prueba, realmente se midió.

Referencias

- Aliseda, A. (1998). La abducción como cambio epistémico: C. S. Peirce y las teorías epistémicas en inteligencia artificial. *Analogía*, 12, 125-144.
- Álvarez, A. (2005). *Escribir en español*. España: Ediciones Nobel, S. A.
- Barni, M., Bandini, A., Sprugnoli, L., Lucarelli, S., Scaglioso, A., Strambi, B., Fusi, C., Arruffoli, A. (2009). *Linee Guida CILS. Certificazione di Italiano come Lingua Straniera Università per Stranieri di Siena*. Perugia: Guerra Edizioni.
- Brown, S. y Glasner, A. (2007). *Evaluar en la Universidad. Problemas y nuevos enfoques*. (2a ed.). Madrid: Narcea.
- Cohen, A.D. (1984). On taking language tests: What the students report. *Language Testing*, 1(1), 70-81.
- Cohen, A.D. (2007). The coming of age for research on test-taking strategies. En J. Fox, M. Weshe, D. Bayliss, L. Cheng, C. Turner, y C. Doe (Eds.), *Language testing reconsidered*. Ottawa: Ottawa University Press, 89-111.
- Cook, R.T. (2009). *A Dictionary of Philosophical Logic*. Edinburgh: Edinburgh University Press.
- Copi, I. y Cohen, C. (2007). *Introducción a la Lógica*. México: Limusa.
- Díaz, P. A. y Leyva, E. (2013). Metodología para determinar los instrumentos de evaluación. En *Educación Médica Superior*, 27(2), pp. 269-286.
- García, E. (1993). La comprensión de textos. Modelo de procesamiento y estrategias de mejora. *Didáctica (Lengua y literatura)*, 5, 87-114.
- Guerra, J. y Guevara, Y. (2013). Validación de un instrumento para medir comprensión lectora en alumnos universitarios mexicanos. *Enseñanza e Investigación en Psicología*, 18(2), 277-291.
- Hawthorne, J. (2014). Inductive Logic. En E. N. Zalta (Ed.), *The Stanford Encyclopedia of philosophy* (Winter 2014 ed.). Recuperado de <http://plato.stanford.edu/archives/win2014/entries/logic-inductive>
- Jiménez, V. (2004). *Metacognición y comprensión de la lectura: Evaluación de los componentes estratégicos (procesos y variables) mediante la elaboración de una Escala de Conciencia Lectora (ESCOLA)*. (Tesis doctoral). Universidad Complutense de Madrid, España.

- Navarro, M. (2012). La interpretación como evento cognitivo expresado en razonamientos abductivos. *Trans/Form/Ação*, 35(3), Universidade Estadual Paulista, Departamento de Filosofia Marília, Brasil. doi:10.1590/S0101-31732012000300012
- Palencia del Burgo, R. (1990). La evaluación como diagnóstico y control. En P. Bello, A. Fera, J.M. Ferrán, T. García Hernández, P. Gómez, M.C. Guerrini, J. López Hernández, D. Martín Catalán, M.M. Martín Viaño, M.R. Martos, C. Mata Barreiro, A. Navarro, R. Palencia, M. Ravera, S. Salaberri, J. M. Sierra y M. Verdú. *Didáctica de las segundas lenguas. Estrategias y recursos básicos*. Madrid: Aula XXI Santillana.
- Pérez Zorrilla, M.J. (2005). Evaluación de la comprensión lectora: dificultades y limitaciones. *Revista de Educación*, num. Extraordinario, 121-138.
- Platas-García, A., Castro-Manzano, J.M., y Reyes-Meza, V., (2016). Razonamiento y heurísticas en pruebas de comprensión lectora. *Revista de Educación*, 371, 157-187.
- Platas-García, A., Castro-Manzano, J.M. y Reyes-Meza, V. (2017). Estrategias para la resolución de pruebas de comprensión lectora inferencial en lengua extranjera. *Revista A & H. Revista de Artes y Humanidades*, 5, 106-117.
- Rodríguez Rodríguez, R. (2005). Abducción en el contexto del descubrimiento científico. *Revista de Filosofía de la Universidad de Costa Rica*, 43(109/110), 87-97.
- Sánchez, S., Grajales, I. y García, E. (2011). Un instrumento para el diagnóstico de habilidades de lectura de los estudiantes de la Licenciatura en Enfermería, UNSIS. *Perfiles Educativos*, 33(132), 110-126.
- Velásquez, M., Cornejo, C. y Roco, A. (2008). Evaluación de la competencia lectora en estudiantes de primer año de carreras del área humanista y carreras del área de la salud en tres universidades del consejo de rectores. *Revista de Estudios pedagógicos [online]*. Universidad Austral de Chile. 34(1), pp. 123-138.
- Vidal-Moscoso, D. y Manríquez-López, L. (2016). El docente como mediador de la comprensión lectora en universitarios. *Revista de la Educación Superior*, 1(177), 95-118.
- Zanotto, M. (2007). *Estrategias de lectura en lectores expertos para la producción de textos académicos*. (Tesis doctoral). Universidad Autónoma de Barcelona, Barcelona, España.