

## Confiabilidad del test de Lectura y Escritura en Español para su aplicación en Costa Rica

*María de los Ángeles Carpio Brenes<sup>1</sup>*

*Ericka Méndez Chacón<sup>2</sup>*

*Universidad de Costa Rica, Costa Rica*

Recepción: 14-10-2017 / Aceptación: 03-01-2018

### Resumen

Este artículo presenta los resultados del estudio realizado entre los años 2015 y 2016 sobre la confiabilidad del Test de Lectura y Escritura en Español (LEE), batería argentino-española para analizar su aplicabilidad en estudiantes costarricenses. Su enfoque es instrumental con metodología cuantitativa, donde el cálculo de confiabilidad y la consistencia interna del test se determinaron con el estadístico Alfa de Cronbach. El instrumento se aplicó en una muestra de 285 estudiantes de la región educativa de Cartago, zona donde se realizó la validez de contenido del mismo test en un estudio realizado entre los años 2013 y 2014 (Carpio y Méndez, 2016). Los resultados globales muestran que de esta batería conformada por siete subpruebas, solo se pueden utilizar, con algunas modificaciones, las subpruebas de Lectura de palabras, Lectura de pseudopalabras, Escritura de palabras y Escritura de pseudopalabras. Por esta razón, se recomienda elaborar una prueba específica para Costa Rica.

**Palabras clave:** aprendizaje, educación, evaluación, lectura, escritura, comprensión.

### Abstract

This article presents the results of the study made in the years 2015 and 2016 about the reliability of the Reading and Writing in Spanish (LEE by its initials in Spanish), Argentinian-Spanish battery test for analyzing its applicability on Costa Rican students. Its perspective is instrumental with quantitative methodology, where the calculation of reliability and the internal consistency of the test were determined with the Cronbach's Alpha statistic. The instrument was applied on a sample of 285 students of the educative region of Cartago, area where the content validity of the same test was carried out in a study conducted between 2013 and 2014 (Carpio and Méndez, 2016). The global results show that this battery test is composed by seven sub-tests, they can only be used with some modifications, the Word-reading sub-tests, pseudo-word reading, word-writing and pseudo-word writing. For this reason, it is recommended to elaborate a specific test for Costa Rica.

**Key Words:** learning, education, evaluation, reading, writing, comprehension.

- 
1. Doctora en Educación, investigadora docente, Universidad de Costa Rica, Costa Rica; email: marycarpiob@gmail.com
  2. Máster en Estadística, docente, Universidad de Costa Rica, Costa Rica; email: ericka.mendez@ucr.ac.cr

## 1. Introducción

Para determinar la competencia lectora del estudiantado que cursa primaria es indispensable contar con instrumentos de medición formales que permitan identificar el nivel de logro en las tareas que esta conlleva, a saber, la segmentación fonética, la decodificación, la fluidez y la comprensión.

En Costa Rica no hay ninguna prueba de medición que evalúe formalmente estos aspectos en Educación Primaria, razón por la cual se consideró la opción de adaptar el Test de Lectura y Escritura en Español (LEE) al contexto costarricense, ya que había sido utilizado en una investigación de tesis doctoral, con 216 estudiantes del país quienes mostraron una adecuada comprensión léxica de las siete subpruebas que la componen (Carpio, 2012).

En el año 2014 se realizó un estudio enfocado a validar el contenido léxico de los textos para la comprensión lectora de este test, cuyo principal resultado fue la identificación de muy pocas palabras que podrían generar confusión en la comprensión de los mismos por un asunto léxico, de las cuales se recomendaron sinónimos como ajuste en caso de que usara con estudiantes costarricenses (Carpio y Méndez, 2016). Este resultado motivó a realizar la investigación que se presenta en este artículo, con el objetivo de determinar la confiabilidad del Test LEE para su aplicación en Costa Rica, ya que además de ser pertinente por las áreas que evalúa (conciencia fonológica, lectura y comprensión), es oportuna; porque coincide con el enfoque fonológico que

promueve el nuevo Programa de Español del Ministerio de Educación Pública (Costa Rica, Ministerio de Educación Pública, 2014) para la enseñanza de la lectoescritura inicial.

### Test LEE

El Test de Lectura y Escritura en Español (LEE) (Defior, Fonseca, Gottheil, Aldrey, Jiménez, Pujals, Rosa, y Dolores, 2006) es el producto del trabajo de un equipo de investigación de la Universidad de Granada, España y la Universidad General San Martín, Argentina. Su finalidad es evaluar los principales procesos implicados en la lectura y escritura, obtener un perfil en el que se observan fortalezas y debilidades, haciendo referencia a los logros medios esperados por año escolar, según los baremos establecidos para cada país.

La batería de evaluación consta de dos pruebas complementarias que se utilizan como evaluación adicional cuando se sospecha la presencia de problemas de lectura; y siete pruebas que pueden administrarse en conjunto o seleccionarse de acuerdo a los aspectos que se necesiten indagar, pues cada una cuenta con sus propios baremos. Las pruebas son las siguientes:

### Componentes

1. **Segmentación fonética (prueba complementaria).** Evalúa la conciencia fonémica al solicitar al estudiantado que aísle los sonidos que componen las palabras que son presentadas de forma oral, pudiendo decir el sonido de la letra o el nombre.



Consta de tres ítems de ensayo y catorce palabras organizadas en orden creciente de dificultad. El rango de puntuación en esta tarea es de 0 a 14 puntos.

2. **Lectura de letras (prueba complementaria).** El objetivo es comprobar si el estudiante conoce todas las letras o manifiesta dificultades en el reconocimiento de alguna de ellas, diciendo el sonido o el nombre de cada letra. Se consigna si realiza rotaciones (b por d, p por qu, etc). El rango de puntuación es de 0 a 29.

### Pruebas

1. **Lectura de palabras.** Valora la utilización de los procesos léxicos y subléxicos que intervienen en la lectura de las palabras, además de la fluidez y la velocidad lectora. Cada estudiante debe leer una lista de 42 palabras seleccionadas, teniendo en cuenta los criterios de frecuencia, longitud y tipo de complejidad ortográfica. Se registra el tipo de lectura (silabeante, vacilante o fluida) y el tiempo de lectura en segundos para obtener una medida de velocidad lectora. La puntuación de los ítems depende del tipo de lectura realizada. Se adjudican dos puntos a la palabra leída en forma fluida; un punto a la palabra leída en forma vacilante o silabeante y cero puntos a la palabra leída en forma incorrecta. El rango de puntuación en esta prueba es de 0 a 84.
2. **Lectura de pseudopalabras.** Su objetivo es evaluar, mediante la lectura de 42 pseudopalabras, los procesos subléxicos de forma estricta, sin posibilidad de apoyo del conocimiento léxico, así como la fluidez y la velocidad lectora. Es importante comparar las diferencias de ejecución que presenta cada estudiante entre la lectura de palabras y de pseudopalabras para identificar si maneja con destreza las reglas de correspondencia grafema fonema o si se apoya en su conocimiento léxico para leer. El rango de puntuación es de 0 a 84 puntos.
3. **Comprensión de palabras y frases.** Evalúa los procesos morfosintácticos que intervienen en la comprensión del texto. Consta de cuatro subpruebas: 3.A. Frases, donde se presentan estímulos simples que puedan ser resueltos por la mayoría del estudiantado; 3.B. Familia de palabras, que mide los conocimientos morfológicos y la comprensión de palabras; 3.C. Preguntas, que evalúa los procesos sintácticos que colaboran con los procesos de comprensión; y 3.D. Completar, con el que se indaga acerca de los procesos morfosintácticos. El rango total de la prueba Comprensión de palabras y frases es de 0 a 43 puntos.
4. **Prosodia.** Mide la capacidad de comprender el significado de una frase teniendo en cuenta la utilización de los signos de puntuación. Cada ítem se compone de dos oraciones que contienen las mismas palabras, pero distintos signos de puntuación, lo que hace que cambie significado.

El estudiante une cada frase con su significado. Consta de cinco ítems y uno de ensayo. A los estudiantes de primero y segundo grado solo se les presentan los tres primeros ítems; a los de tercer y cuarto grado se les aplica todos. El rango de puntuación es de 0 a 6 en primer y segundo grado; y de 0 a 10 en tercer y cuarto grado.

5. **Comprensión de textos.** Valora la comprensión de palabras y de oraciones, y la relación e interpretación de unas oraciones con otras. Se le va presentando a cada estudiante tres textos que debe leer y comprender, luego se le formulan las preguntas acerca del contenido de esos textos. La prueba incluye un ítem de selección del título y otro de selección del resumen en cada texto. La puntuación parcial de cada texto oscila entre 0 y 16 puntos. El rango de puntuación total de la prueba es de 0 a 48.
6. **Escritura de palabras.** Evalúa los conocimientos fonológicos y ortográficos que utiliza el estudiantado al escribir palabras. Consiste en un dictado de 44 palabras. El rango de puntuación de esta prueba oscila entre 0 y 44.
7. **Escritura de pseudopalabras.** Mide el conocimiento de las reglas de correspondencia grafema fonema de forma estricta. Consiste en un dictado de 32 pseudopalabras. El rango de puntuación es de 0 a 32 puntos.

La puntuación en cada año escolar corresponde a la finalización del mismo. De modo

que si una persona es evaluada al comienzo del año escolar se deben aplicar las pruebas y puntuaciones correspondientes al año anterior. Aunque en algunas pruebas se pueden obtener puntuaciones parciales para realizar el perfil se tendrán en cuenta solo las puntuaciones totales.

Al completar el Registro de Respuestas (RegR) y analizar el perfil, se puede observar que los percentiles se encuentran divididos en tres áreas, lo que ubica al estudiante en alguno de los siguientes niveles de ejecución:

- Con dificultad, del percentil 0 al percentil 30
- Competente, del percentil 30 al percentil 70 (en este rango se ubicaría el mayor porcentaje de la población evaluada)
- Muy competente, del percentil 70 al percentil 100

Como se reportó en el artículo producto de la primera parte de esta investigación correspondiente a la validez de contenido del test LEE (Carpio y Méndez, 2016), Defior y otras (2006) indican que los estudios de confiabilidad arrojaron indicadores satisfactorios tanto acerca de la consistencia interna de los ítems que conforman todas las pruebas de la batería como de la confiabilidad test-retest con un mes de intervalo entre las dos administraciones, y con la confiabilidad por división por mitades de las pruebas de lectura y escritura. Las propiedades estudiadas del test, según las autoras, indican que este instrumento constituye una herramienta adecuada para evaluar procesos vinculados



a la lectoescritura, en estudiantes de primero a cuarto grado de primaria.

Para el análisis de confiabilidad de los resultados obtenidos en la aplicación con el estudiantado costarricense se utilizó la Teoría clásica de los test (TCT) (Nunnally y Bernstein, 1995), debido a que este fue el enfoque utilizado por las autoras. Este modelo consiste en asumir que la puntuación que una persona obtiene en un test, que se denomina puntuación empírica y que suele asignarse con la letra  $X$ , está formada por dos componentes: por un lado, la puntuación teórica ( $T$ ) o verdadera ( $V$ ) de esa persona en ese test; y por otro, un error ( $E$ ), que puede ser debido a muchas causas que se escapan y que no se pueden controlar. Esto se expresa formalmente así:  $X=T + E$  (Muñiz, 2010).

Este modelo parte de los siguientes supuestos:

1. La puntuación observada de una persona en un test se compone de dos variables latentes: la puntuación teórica o verdadera ( $T$ ) y el error de medida ( $E$ ).
2. La variable latente  $E$  es una variable aleatoria, que sigue una distribución normal con media cero y varianza  $\sigma^2$ .
3. Las variables latentes  $T$  y  $E$  no están correlacionadas.
4. Cuando se realizan dos mediciones, los errores  $E$  son independientes entre sí.
5. El error que se comete con una medida es independiente de la puntuación teórica de una segunda medición.

A partir de los supuestos anteriores, se define la confiabilidad de la medida como una correlación entre dos conjuntos de puntuaciones obtenidas con el mismo test o con formas paralelas de test. Como complemento y en busca de mejoras a la teoría clásica, han surgido nuevas teorías para evaluar datos en el campo de la psicometría; sin embargo, dado que las autoras del test LEE utilizaron el enfoque clásico, se estimó el mismo cálculo para Costa Rica, a fin de mantener la comparabilidad.

En el caso de evaluar la capacidad de lectura y escritura (constructo), esta se podría definir como una cualidad del estudiantado no observable directamente. Para ello se miden variables (ítems o preguntas) que sí son observables en cada estudiante. Por ejemplo, lectura de la palabra *Péndulo*. Se supone que las variables que se miden están relacionadas con la magnitud del constructo de interés (capacidad lectora y de escritura) que es inobservable. Entre ellas, las mediciones son estables, consistentes y correlacionadas; es decir, esperamos que todos los componentes de cada prueba estén altamente correlacionados entre sí, y además que los resultados sean similares entre estudiantes con un nivel o capacidad lectora parecidos.

El alfa de Cronbach es un estadístico desarrollado por J. L. Cronbach. Según González y Pasmíño (2015), este estadístico “es una forma sencilla y confiable para la validación del constructo de una escala y como una medida que cuantifica la correlación existente entre los ítems que componen esta” (p. 63). Es una

correlación entre la escala medida y el constructo subyacente.

Para su cálculo se requiere de una sola administración del instrumento de medición y produce valores que varían entre 0 y 1 (pues es una correlación). También permite evaluar cuanto mejora la confiabilidad de una prueba si se excluye una determinada variable o ítem. De acuerdo con González y Pasmíño (2015), “un valor del alfa de Cronbach entre 0.70 y 0.90 indica una buena consistencia interna para una escala unidimensional” (p. 63).

## 2. Metodología

El propósito principal de esta investigación fue determinar la confiabilidad del Test de Lectura y Escritura en Español (LEE), es decir, el grado en que este instrumento produce resultados consistentes y coherentes para su aplicación en el contexto escolar costarricense. Se enmarca en un estudio instrumental con metodología cuantitativa porque, de acuerdo con León y Montero (2003), a esta categoría pertenecen aquellos estudios orientados al “desarrollo de pruebas y aparatos, incluyendo tanto el diseño (o adaptación) como el estudio de las propiedades psicométricas de los mismos” (p. 799).

El instrumento se aplicó en una muestra de estudiantes de segundo a quinto grado del Circuito 06 de la Región Educativa de Cartago, Costa Rica donde se realizó la validez de contenido del Test LEE (Carpio y Méndez, 2016) como primera etapa de esta investigación. Cabe aclarar que en ese primer estudio se hizo referencia al Circuito 09 porque esa

era su numeración antes del 2016, año en que se hizo una re-organización y pasó a ser en la actualidad Circuito 06.

El estudiantado de primer grado fue excluido del estudio por no haber alcanzado un nivel de decodificación lectora aceptable en el momento de la aplicación de las pruebas. También se descartó a los estudiantes que el profesorado reportó con necesidades educativas especiales. La muestra se seleccionó estratificada por nivel. Cada una se obtuvo fijando un nivel de confianza del 95% (precisión [ $Z\alpha/2$ ]), un margen de error del 10% ( $\epsilon$ ) y asumiendo una variable de interés de tipo categórica, por ejemplo: ¿Cómo es el nivel general de lectura?, con categorías de respuesta: Excelente o No excelente, cuya proporción de ocurrencia ( $p$ ) es de 0.8. La fórmula utilizada para el cálculo muestral por estrato fue la siguiente:

$$n_0 = \frac{z^2_{\alpha/2} * p * (1 - p)}{\epsilon^2}$$

El tamaño muestral final se ajustó por finitud, tomando en cuenta la matrícula provisional del 2015  $n=n_0 / (1 + [n_0/N])$ . En aquellas escuelas donde se seleccionó un grupo por nivel y había dos o más; se seleccionó al azar, ya fuera en el sitio o en el grupo que se quedó en la muestra. Por diferentes motivos se debió sustituir algunos de los grupos seleccionados durante la aplicación de los instrumentos. En la Tabla 1 se detalla las diferentes circunstancias. Todas las sustituciones se realizaron en forma aleatoria.



Tabla 1. Resultados de la ejecución de la muestra

Escuela	Grupos seleccionados	Estudiantes en la muestra	Observaciones	Muestra efectiva
<b>Segundo grado</b>				
San Francisco	1	20	No se realizó en la escuela de San Francisco porque la maestra del grupo reusó participar debido a que la mayoría de los estudiantes tenía necesidades educativas especiales. Se sustituyó por un segundo grado en la escuela Villas de Ayarco.	22
Calle Naranjo	2	41		38
San Vicente	1	21		21
	<b>Total</b>	<b>82</b>	<b>Total</b>	<b>81</b>
<b>Tercer grado</b>				
Villas de Ayarco	1	24		22
Carolina Bellelli	1	23	Escuela en construcción y estaban reubicados en las aulas de catecismo de la iglesia. Espacio no adecuado para aplicación de las pruebas. Se sustituyó por la escuela San Vicente.  Se tuvo que sustituir por el tercer grado de la escuela Quebrada del Fierro por incompatibilidad con los horarios.	27
Central	1	27		17
	<b>Total</b>	<b>75</b>	<b>Total</b>	<b>66</b>
<b>Cuarto grado</b>				
San Vicente	1	26	No se evaluó porque el grupo estaba en capacitación. Se sustituyó por el cuarto grado de la escuela Santiago del Monte.	27
Fernando Terán	1	23		24
Central	2	48		41
	<b>Total</b>	<b>75</b>	<b>Total</b>	<b>92</b>
<b>Quinto grado</b>				
Unidad	1	25		15
Pedagógica Central	1	28		16
Quebrada del Fierro	1	22		15
	<b>Total</b>	<b>75</b>	<b>Total</b>	<b>46</b>
	<b>Muestra Total</b>	<b>328</b>	<b>Muestra total ejecutada</b>	<b>285</b>

Al final se contó con una muestra efectiva de 285 estudiantes, distribuidos de la siguiente manera: 28% segundo grado, 23% tercer grado, 32% cuarto grado y 16% quinto grado.

Como se describió en el referente teórico, el Test de Lectura y Escritura en Español (LEE), está conformado por: dos pruebas complementarias (Segmentación fonética y Lectura de letras); cuatro pruebas que evalúan el cumplimiento de las reglas de correspondencia grafema fonema RCGF (Lectura de palabras, Lectura de pseudopalabras, Escritura de palabras, Escritura de pseudopalabras), y tres pruebas para determinar la comprensión lectora (Comprensión de palabras y frases, Prosodia, Comprensión de textos).

Con las pruebas para la evaluación de las RCGF, el estudiantado debe cumplir las normas idiomáticas del español, razón por la cual no corresponde realizar ningún cambio o aclaración del léxico en estas. No sucede lo mismo con las pruebas de comprensión, pues al ser un test extranjero, el vocabulario utilizado en los textos podía generar sesgo en los resultados del estudiantado costarricense. Por esta razón, se sustituyeron las palabras difíciles (ovillos, estufa, asintió, era cuaternaria, rudimentarias, anzuelo) por los sinónimos (bolas-rollos, cocina-chimenea, aceptó, era de hielo, simples, trampa) que resultaron de la primera parte del estudio que se reporta aquí (Carpio y Méndez, 2016).

### **Procedimiento**

Durante el año 2015 se realizó el trabajo de campo en coordinación con el Supervisor de

Centros Educativos del Circuito 06, la investigadora principal y los directores de las instituciones educativas que fueron seleccionadas para la muestra. Estas últimas fueron las personas encargadas de escoger las secciones de los niveles que debían participar, y de solicitar la colaboración al profesorado en la ejecución del proyecto.

Se coordinó con cada docente las dos fechas de aplicación de la prueba al estudiantado de su grupo correspondiente, y se les hizo entrega de las cartas de consentimiento informado para que las familias de sus estudiantes aprobaran su participación en el estudio.

Las siete pruebas fueron aplicadas a inicios del periodo lectivo del año 2016; esto aseguró que los 285 estudiantes participantes ya habían concluido el nivel requerido, pues en el manual técnico del test se establece que, si el estudiantado es evaluado al comienzo del año escolar se le debe aplicar las pruebas y puntuaciones correspondientes al año anterior. De esta manera: los de segundo grado ya cuentan con el nivel de primer grado completo, los de tercero con el segundo nivel, los de cuarto con tercer año y los de quinto con cuarto grado.

Las pruebas 1, 2 y 5 se aplicaron en forma individual, con grabación de voz, pues el instrumento solicita medir el tiempo de lectura en las tres tareas. Para las pruebas 3 y 4 se le entregó a cada estudiante el material impreso que establece el Test LEE, y se dio la instrucción general al grupo para cada una de las tareas que solicitan ambas pruebas. Esta parte



estuvo dirigida por la investigadora principal que tiene formación en Educación Especial, con el apoyo de la estudiante asistente, también estudiante de Educación Especial, quien colaboró en la supervisión del estudiantado. La prueba 6, que corresponde a un dictado de palabras, se realizó al grupo en general, siguiendo la consigna de repetirla solo dos veces. La prueba 7 fue aplicada por la investigadora principal y la estudiante asistente, a cada estudiante por separado, ya que al ser un dictado de pseudopalabras, podría generar confusión si se aplicaba al grupo completo.

Cada una de las pruebas del estudiantado fue organizada en folders, por institución y sección, y las grabaciones de voz guardadas en carpetas, también por institución y sección para su posterior análisis. La revisión de las pruebas y el volcado de datos en el registro impreso de resultados fue realizado por la investigadora principal y la estudiante asistente de la investigación, según el grupo de estudiantes que le correspondió realizar las evaluaciones individuales.

Estos registros fueron entregados en el mes de enero de 2016 a la investigadora colaboradora, quien es especialista en estadística, para los análisis correspondientes. Durante el año 2016, esta investigadora colaboradora se encargó de ingresar los datos al programa estadístico y realizar los análisis e interpretaciones correspondientes a la luz de los objetivos propuestos en el estudio.

El procedimiento que se utilizó para calcular la confiabilidad del Test LEE fue la medida

de consistencia interna mediante el Alfa de Cronbach, estadístico que utiliza resultados arrojados por el mismo instrumento que es administrado una sola vez a la población en estudio. Se utilizó el análisis de confiabilidad de la teoría clásica de los tests para evaluar la pertinencia de los ítems de la prueba en esta población, explicada en el referente teórico de este artículo.

Hay que recordar que en el alfa de Cronbach, para que la escala sea confiable, entre más se aproxime a 1 el valor el coeficiente de alfa, mayor es la confiabilidad; los valores iguales o superiores a 0.7 son los adecuados para garantizar la confiabilidad (Frías-Navarro, 2014).

### 3. Resultados

El Test de Lectura y Escritura en Español: LEE está organizado en los dos grandes componentes de la lectura: el reconocimiento de palabras y la comprensión lectora, subdivididos en las siguientes siete pruebas: Lectura de palabras, Lectura de pseudopalabras, Comprensión de palabras y frases, Prosodia, Comprensión de textos, Escritura de palabras y Escritura de pseudopalabras. A continuación se muestran los resultados obtenidos para cada una de esas pruebas, en el respectivo orden en que se presentan en el Test LEE:

#### Lectura de palabras

El objetivo de esta prueba es evaluar la utilización de los procesos léxicos y subléxicos que intervienen en la lectura de las palabras, además de la fluidez y la velocidad lectora.

Al estudiantado participante se le presentó, en forma individual, una lista de 42 palabras que debían leer en voz alta, para obtener su registro de voz y hacer la evaluación posterior según los siguientes criterios: 0 puntos si la respuesta es incorrecta; 1 punto si la palabra es leída correctamente pero en forma no fluida (vacilante o silabeante); y 2 puntos si la palabra es leída en forma correcta y fluida.

**Tabla 2.** Alfa de Cronbach para la prueba de lectura de palabras según grado

Grado	Alfa de Cronbach	Alfa de Cronbach basada en elementos estandarizados
Segundo	0.947	0.951
Tercero	0.917	0.922
Cuarto	0.836	0.862
Quinto	0.897	0.907

Como se observa en la tabla 2, la evaluación de la confiabilidad de esta prueba, a partir de la medida alfa de Cronbach, es superior a 0.90 para segundo y tercer grado; y superior a 0.80 para cuarto y quinto grado, lo que corresponde a los niveles de Excelente y Bueno, respectivamente. Sin embargo, se identificó que las palabras: *chiste*, *rima*, *fiel*, *flan*, *faro*, *duquesa* y *fijó*, no funcionan como parte de la prueba porque presentan poca variabilidad.

El promedio general de palabras leídas en forma fluida fue de 30 con una desviación de 10. Para los niveles de segundo y tercer grado fue de 25, y para cuarto y quinto grado fue de 37 y 34, respectivamente. Las palabras que resultaron más fáciles de leer en todos los niveles

fueron: *pavo*, *honda*, *cruel*, *faro*, *fiel*, *flan*, *chiste* y *rima*. Estas palabras tienen un promedio de nota de lectura superior o igual a 1.7.

Las palabras más difíciles para todo el estudiantado fueron: *fijó*, *péndulo*, *gitano*, *anguila*, *entretenimiento*, *animó*, *asfalto* y *fachada*. En el caso de la palabra *fijó*, el promedio en todos los niveles fue inferior a la unidad; esto quiere decir que la lectura fue incorrecta y en algunos casos vacilante o silabeante, pero que no se logró una adecuada lectura. En igual condición están *péndulo* y *gitano* en el caso del estudiantado de segundo y tercer grado. Las otras palabras mencionadas -*anguila*, *entretenimiento*, *animó*, *asfalto* y *fachada*-, muestran un promedio de nota de lectura menor a 1.5.

### Lectura de pseudopalabras

Esta prueba evalúa los procesos subléxicos de forma estricta, sin posibilidad de apoyo del conocimiento léxico, así como la fluidez y la velocidad lectora. Se califica asignando la puntuación de la misma forma que en la lectura de palabras.

**Tabla 3.** Alfa de Cronbach para la prueba de lectura de pseudopalabras según grado

Grado	Alfa de Cronbach	Alfa de Cronbach basada en elementos estandarizados
Segundo	0.93	0.93
Tercero	0.82	0.82
Cuarto	0.85	0.84
Quinto	0.90	0.90

La tabla 3 muestra que la confiabilidad de esta prueba resultó superior a 0.80, lo que indica que tiene una buena consistencia interna. Se identificaron las pseudopalabras: *balma, dujó, pénfano, padunó, cetilno, agenso, soque, chisño, yesa, pénfano y guite* con correlaciones menores a 0.30 o negativas, lo que se consideran inaceptables, por lo que corresponde ser eliminadas de la prueba.

La lectura de pseudopalabras fue un poco más difícil para el estudiantado, pues el promedio general obtenido en la lectura fue de 1.4, en comparación con el 1.6 que se obtuvo en la lectura de palabras. La cantidad media de pseudopalabras leídas en forma fluida fue de 24 con una desviación de 10; mientras que la cantidad media de palabras leída fluidamente fue de 30 con una desviación de 10.

Las pseudopalabras que resultaron más fáciles de leer en todos los niveles fueron: *ripa, pioz, yesa, soque, maro y us*, las cuales tienen un promedio de nota de lectura superior o igual a 1.7. Las pseudopalabras más difíciles de leer para todo el estudiantado fueron: *padunó, dujó, pénfano e ipitanami*, pues ni siquiera alcanzaron el promedio de 1 en los cuatro niveles, lo que indica que ninguna fue leída fluidamente. En condición similar están las pseudopalabras: *astridamiensa, mastán, funtimiente, duguizo, peyata, disnutible, astenda, cetilno, genies, farroda, chisño, guite, ingrónico, pansata, sasgla, halade, agenso, hemildo y detel*, las cuales muestran un promedio de nota de lectura menor a 1.5.

## Comprensión de palabras y frases

El objetivo de esta prueba es evaluar los procesos morfosintácticos que intervienen en la comprensión del texto. Se subdivide en cuatro subpruebas: 3A. Frases, 3B. Familia de palabras, 3C. Preguntas y 3D. Completar.

### 3A. Frases

Esta subprueba está compuesta por partes inferenciales que sirven de introducción a las otras subpruebas. La puntuación obtenida varía entre 0 y 5, se obtiene 1 punto si la respuesta es correcta y 0 si es incorrecta. En general hubo un buen desempeño del estudiantado dando respuesta a estas preguntas, a excepción del ítem 2, que corresponde a una pregunta de un texto redactado en voz pasiva: “El ratón es atrapado por el perro”/ ¿Qué hace el perro?

### 3B. Familia de palabras

Con esta prueba se evalúa conocimientos morfológicos y comprensión de palabras. La puntuación obtenida varía entre 0 y 28. En esta subprueba se mostraba un grupo de 28 palabras al estudiantado, de las cuales debían indicar si eran o no familia de otra palabra que se encontraba encerrada en un círculo. Eran cuatro grupos de familias. La media global de acierto fue de 80%. Los ítems donde se presentó mayor dificultad fueron el dos, relacionado con la familia de la palabra *domador*; y el cuatro, relacionado con la familia de la palabra *calificado*.

### 3C. Preguntas

Esta subprueba evalúa procesos sintácticos que colaboran con los de comprensión lectora, por medio de la presentación de dos frases con tres preguntas por cada una de ellas. La puntuación obtenida varía entre 0 y 6. El porcentaje global de aciertos en esta prueba fue de apenas 58%, es decir, las respuestas correctas las lograron casi sólo la mitad del estudiantado. El mayor porcentaje de acierto, en todos los niveles, se dio con el primer ítem de cada una de las propuestas, el cual corresponde a ¿Quién leyó el cuento? (Comprensión 1 y 4). Las demás preguntas derivadas del texto no tuvieron alto porcentaje de acierto.

### 3D. Completar

Indaga sobre procesos morfosintácticos. Cada frase debe ser completada siguiendo procesos de coherencia y cohesión. La puntuación obtenida varía entre 0 y 4. En esta prueba, el porcentaje de acierto medio al completar las dos primeras frases fue de 86% y 31% para las dos últimas frases.

### Confiabilidad

Según el manual técnico del Test LEE, para el cálculo de la confiabilidad de esta prueba no se toma en cuenta la sección 3B.

Como se puede observar en la Tabla 4, el alfa de Cronbach para esta prueba no alcanza el 0.50, por lo que se considera inaceptable.

**Tabla 4.** Alfa de Cronbach para la prueba de comprensión de palabras y frases, según grado

Grado	Alfa de Cronbach	Alfa de Cronbach basada en elementos estandarizados
Segundo	0.93	0.93
Tercero	0.82	0.82
Cuarto	0.85	0.84
Quinto	0.90	0.90

### Prosodia

Con esta prueba se pretende evaluar la capacidad de comprender el significado de una frase teniendo en cuenta los signos de puntuación. Cada ítem se compone de dos oraciones que contienen las mismas palabras, pero signos de puntuación diferentes, lo que hace que cambie el significado. El estudiantado debe unir cada frase con el significado correcto. La puntuación es de 0-6 para el estudiantado de primer y segundo grado completo, es decir tres ítems y de 0-10 para el estudiantado de tercer y cuarto grado completo, es decir cinco ítems.

Para el cálculo de la confiabilidad se proponen cinco indicadores que conforman los ítems de la prueba. Es decir, se toma como correcto el ítem si contesta en forma correcta las dos frases de ese ítem. En este caso el ítem 1 estaría correcto si contestan adecuadamente Prosodia 11 y Prosodia 22.



**Tabla 5.** Alfa de Cronbach para la subprueba prosodia, según grado.

Grado	Alfa de Cronbach	Alfa de Cronbach basada en elementos estandarizados
Segundo	0.24	0.21
Tercero	0.26	0.32
Cuarto	0.51	0.52
Quinto	0.49	0.49

De acuerdo con los datos que se presentan en la Tabla 5, se puede decir que esta prueba no es confiable y que carece de consistencia interna, esto por lo bajo de los indicadores de Cronbach y porque son muy pocos los ítems que la componen.

### Comprensión de textos

El objetivo de esta prueba es evaluar la comprensión, relación e integración de las oraciones. Para primero y segundo grado se presentan tres textos: Dos narrativos (“Pupi” y “Los Dinosaurios”) y uno expositivo (“El Gallo Pepe”). Para tercero y cuarto grado, se emplean un texto narrativo directo (“Los Delfines”) y dos expositivos (“Lucas” y “La Era Cuaternaria”).

En la evaluación de la comprensión de los textos, se realizan tres preguntas literales y tres inferenciales; además se pregunta por la selección del título y el resumen de cada texto. Las preguntas literales e inferenciales se califican con 0 si la respuesta es incorrecta, con 1 si es incompleta y con 2 puntos si es correcta.

En la primera etapa de este proyecto se ajustaron las palabras de los textos al contexto

de nuestro país (Carpio y Méndez, 2016), las cuales se consideraron en la aplicación de esta prueba para que el componente léxico no sesgara los resultados.

### Sobre los textos de primero y segundo grado

La media global de comprensión de los textos: “Pupi”, “Los Dinosaurios” y “El Gallo Pepe”, correspondientes al estudiantado de segundo y tercer grado, fue de 1.5. El ítem más difícil de responder en los tres textos fue el relacionado con las preguntas tipo inferencial puente, que son las que se responden mediante la adición de elementos semánticos no explícitos en el texto.

Con respecto a la selección del título, el 58% del estudiantado de ambos grupos respondieron correctamente. El texto en el que salieron con mejores resultados fue el denominado “Pupi”, pues un 60% del estudiantado de ambos grados lo hizo bien; mientras que el texto que más les costó fue el de “Los Dinosaurios”. Por otra parte, el promedio general en la selección correcta de la frase que resume el texto fue de sólo el 38% del estudiantado de segundo y tercer grado. En el texto: “El Gallo Pepe” fue donde presentaron mejor resultado con poco más del 52% de respuestas correctas.

### Sobre los textos de tercero y cuarto grados

La media global de comprensión de textos del estudiantado de cuarto y quinto grado a quienes se les aplicó esta prueba fue de 1.3, ligeramente menor a la media observada por los estudiantes de grados inferiores. El ítem que fue más difícil de responder en el texto

de “Los Delfines” fue el correspondiente a la pregunta de tipo literal, mientras que en los textos “Lucas” y “La Era Cuaternaria” fueron las preguntas inferenciales puente e inferenciales elaborativas (se responden con información que no se encuentra de forma explícita en el texto).

En la selección correcta del título, el resultado más alto obtenido fue en el texto denominado “Lucas”, con un porcentaje superior a 75% en ambos grados. El texto más difícil fue el de la “Era Cuaternaria”. Por otra parte, la respuesta correcta de la frase que resume el texto, también fue mayor en el texto “Lucas”, con más del 60% del estudiantado. En general, el 66% del estudiantado que cursa cuarto y quinto grados seleccionó correctamente el título y sólo el 56% realizó la selección correcta de la frase que resume el texto.

### Escritura de palabras

Evalúa los conocimientos fonológicos y ortográficos. Se les realiza un dictado de 44 palabras.

**Tabla 6.** Alfa de Cronbach para la subprueba escritura de palabras según grado.

Grado	Alfa de Cronbach	Alfa de Cronbach basada en elementos estandarizados
Segundo	0.83	0.84
Tercero	0.87	0.86
Cuarto	0.69	0.71
Quinto	0.81	0.81

En la Tabla 6 se muestra que la confiabilidad de esta prueba es mayor a 0.80 para los niveles segundo, tercero y quinto grado, por lo que se

considera adecuada para esos tres niveles. No sucede lo mismo con la de cuarto grado, que da un puntaje de 0.69, quedando con un criterio de cuestionable. Algunas de las palabras no son adecuadas para esta prueba debido a que presentan correlaciones negativas o muy bajas. Según el año escolar, las palabras con esas características son las que se muestran en la Tabla 7:

**Tabla 7.** Palabras que presentan correlaciones negativas por año escolar.

Grado	Alfa de Cronbach	Alfa de Cronbach basada en elementos estandarizados
Segundo	0.83	0.84
Tercero	0.87	0.86
Cuarto	0.69	0.71
Quinto	0.81	0.81

Las palabras que resultaron más difíciles de escribir en forma correcta fueron: *péndulo, sultán, hamaca, fijó, humilde, animó, yema, guiño y pompa*; puesto que el porcentaje de escritura correcto fue menor a 40%. Cabe resaltar que las palabras: *fijó, péndulo, animó*, también fueron de las más difíciles de leer. Por otra parte, las palabras: *dedal, pantano, fachada, rima, tableta, mantel, chiste y fiel*, fueron escritas correctamente en promedio el 85% de las veces.

### Escritura de pseudopalabras

Se evalúa el conocimiento de las reglas de correspondencia grafema-fonema. Se les realiza un dictado de 32 pseudopalabras (Defior y otras, 2006).



**Tabla 8.** Alfa de Cronbach para la subprueba escritura de pseudopalabras según grado

Grado	Alfa de Cronbach	Alfa de Cronbach basada en elementos estandarizados
Segundo	0.83	0.81
Tercero	0.82	0.81
Cuarto	0.60	0.58
Quinto	0.69	0.67

Como se muestra en la Tabla 8, esta prueba tiene una confiabilidad ligeramente superior a 0.80 para primero y segundo grado, y superior a 0.60 para tercero y cuarto grado, lo que sugiere que es mejor para los años inferiores que para los superiores. Algunas pseudopalabras que tienen correlaciones negativas o muy bajas, son las que se muestran en la Tabla 9, mismas que se sugieren ser eliminadas de la prueba:

**Tabla 9.** Pseudopalabras que tienen correlaciones negativas por año escolar

Segundo	Tercero	Cuarto	Quinto
Pénfano	Dujó	Celtino	Maro
Mastán		Croiz	Anqueta
Padunó		Detel	Fual
		Anqueta	Soblesa
		Mandal	
		Mastán	
		Farroda	
		Pansata	

Las palabras *deplo*, *denotre* y *ripa* fueron las pseudopalabras más fáciles de escribir, ya que el porcentaje que las escribió correctamente es superior al 80%. Por otro lado, *chisño*, *em-*

*pasa*, *maro* y *entrinaumo* fueron las más difíciles de escribir, ya que menos de un 10% las escribió correctamente. En general, se mantiene la tendencia de que los niveles superiores escriben mejor las pseudopalabras.

#### 4. Discusión

De acuerdo con el objetivo general de la investigación, en cuanto a la confiabilidad global del Test de Lectura y Escritura en Español (LEE), se puede decir que por la carencia de consistencia interna, al presentar alfas muy bajos en la mayoría de los ítems, no es conveniente la aplicación de la batería completa de pruebas del test en el contexto educativo costarricense.

Sin embargo, se pudo determinar que la confiabilidad de las pruebas de Lectura de Palabras, Lectura de Pseudopalabras, Escritura de Palabras y Escritura de Pseudopalabras es adecuada; por lo que se podrían utilizar en Costa Rica, siempre y cuando los ítem que presentaron correlaciones bajas o negativas en estas pruebas, sean sustituidas por otros que busquen el mismo fin; o bien, eliminarlos y ajustar el puntaje al número de palabras que sí obtuvieron valores de alfas adecuados.

Con respecto a las otras pruebas, la de Comprensión de palabras y frases, en términos globales, resultó difícil para el estudiantado de los cuatro niveles (de 2° a 5° grado). En la sección de Frases, el ítem problemático corresponde a una pregunta de un texto redactado en voz pasiva. En la sección de Familia de palabras tuvieron problemas con dos de

las familias (mitad de los ítems) relacionadas con las palabras *domador* y *calificado*. En la sección de Preguntas, la dificultad estuvo en la relación entre los personajes de las dos oraciones planteadas y sus nombres, lo que generó respuestas incorrectas a las tres preguntas literales que se plantearon en cada oración. Finalmente, en la sección de Completar, la dificultad se dio en los ítems C y D, correspondientes a tiempos más complejos.

En la prueba de Prosodia, el desempeño tampoco fue bueno; la media global de apro-

bación fue de 54%, esto demuestra que el estudiantado tiene problemas con la interpretación de los signos de puntuación. La prueba denominada Comprensión de textos, a nivel global, también obtuvo resultados muy bajos. La dificultad mayor se dio en la comprensión de preguntas inferenciales y al solicitarles que resumieran la idea central de los textos. Por los resultados globales obtenidos en esta investigación, se recomienda diseñar una prueba específica para Costa Rica.

## 5. Referencias

- Carpio, M. y Maldonado, A. (2012). *Eficacia de las estrategias pictofónicas en la enseñanza de la lectura inicial en Costa Rica: Un enfoque longitudinal* (Tesis de doctorado). Universidad Autónoma de Madrid, España. Recuperado de [https://repositorio.uam.es/bitstream/handle/10486/11756/58112\\_carpio\\_brenes\\_maria\\_de\\_los\\_angeles.pdf?sequence=1](https://repositorio.uam.es/bitstream/handle/10486/11756/58112_carpio_brenes_maria_de_los_angeles.pdf?sequence=1)
- Carpio, M. y Méndez, E. (2016). Validación de contenido léxico de los textos para la comprensión lectora del Test de Lectura y Escritura en Español (LEE) para su aplicación en Costa Rica. *Revista Actualidades Investigativas en Educación*, 18(2), pp. 1-28
- Defior, S., Fonseca, L., Gottheil, B., Aldrey, A., Jiménez, G., Pujals, M., Graciela, R. y Serrano, F. (2006). *LEE Test de lectura y escritura en español*. Buenos Aires: Paidós
- Frías-Navarro, D. (2014). *Apuntes de SPSS*. Recuperado de: <http://www.uv.es/friasnav/ApuntesSPSS.pdf>
- González, J. y Pazmiño, M. (2015). Cálculo e interpretación del alfa de Cron Bach para el caso de validación de la consistencia interna de un cuestionario, con dos poblaciones escalas tipo Likert. *Revista Publicando*, 2(1), pp. 62-77
- León, I. y Montero, O. (2003). *Métodos de Investigación en Psicología y Educación*. Madrid: McGraw-Hill



Costa Rica, Ministerio de Educación Pública. (2014). Programa de Estudios de Español. I Ciclo Educación General Básica. Recuperado de <http://www.mep.go.cr/programa-estudio/espano>

Muñiz, J. (2010). Las teorías de los tests: Teoría Clásica y Teoría de Respuesta a los Ítems. *Papeles del Psicólogo*, 31(1), pp. 57-66

Nunnally, J y Bernstein, I. (1995). Teoría Psicométrica. Editorial McGrawHill. Tercera edición