

Propiedades psicométricas de una prueba de lectura de palabras (LEO-1-min)

**Edurne Goikoetxea¹, Wim Van Bon², Gorka Fraga³ y
Naroa Martínez¹**

¹ Departamento de Psicología y Educación, Universidad de Deusto, Bilbao

² Facultad de Ciencias Sociales, Universidad de Radboud, Nijmegen

³ Departamento de Psicología del Desarrollo, Universidad de Zurich, Zurich

España, Los Países Bajos, Suiza

Correspondence: Edurne Goikoetxea. Av. de las Universidades, 24. 48007 – Bilbao, Vizcaya, SPAIN. E-mail: egoikoetxea@deusto.es

© Universidad de Almería and Ilustre Colegio Oficial de la Psicología de Andalucía Oriental (Spain)

Resumen

Introducción. Maestros e investigadores necesitan con frecuencia evaluar la lectura de palabras en grupo y en poco tiempo. El test LEO-1-min fue creado para medir la lectura de palabras a través de una tarea de decisión léxica donde el examinado debe identificar pseudopalabras dentro de una lista de palabras frecuentes.

Método. Participaron 284 niños de 1º to 6º de una escuela pública concertada. Se crearon cuatro formas alternativas del LEO-1-min, cada una con 180 estímulos (132 palabras y 48 pseudopalabras).

Resultados. Los resultados mostraron que una adecuada fiabilidad para formas alternas (rangos = de .57 a .81). Se encontraron altas correlaciones entre las puntuaciones en el LEO-1-min y las de un test estandarizado de lectura en voz alta. El análisis discriminante de las puntuaciones del LEO-1-min mostró un alto nivel de éxito en la predicción del rendimiento en lectura oral.

Discusión y Conclusion: La fiabilidad de las puntuaciones del LEO-1-min es aceptable a buena. La tarea de decisión léxica del LEO-1-min mostró una alta correlación con la lectura oral, lo que apoya el uso de la decisión léxica como test grupal para identificar rápidamente lectores con pobre lectura de palabras. Se ofrece la Forma A del test y baremos provisionales por curso.

Palabras clave: Lectura de palabras, desarrollo lector, test de lectura, tareas de decisión léxica.

Abstract

Introduction. Teachers and researchers often need to evaluate word decoding skill in group-wise and in a short time. The LEO-1-min test is created to measure word reading through a lexical decision procedure where the examinee identifies pseudowords in a list of frequent words.

Objective. To examine the reliability and validity of LEO-1-min, a silent word reading test, suitable for quick assess of reading abilities in a wide age range of students.

Method. Participants were 284 children from 1st to 6th grade of a subsidized Primary School. We created four alternate forms of the LEO-1-min, each with 180 stimuli (132 words and 48 pseudowords).

Results. The results show an adequate parallel forms reliability of the scores (range r_s = from .57 to .81). High correlations were found between the scores on the LEO-1-min and the scores on a standardized reading aloud test. The discriminant analysis of the scores on the LEO-1-min shows a high level of success in predicting the oral word decoding performance.

Discussion and Conclusion. LEO-1-min reliability is acceptable to good. Lexical decision in LEO-1-min and oral reading are highly correlated, which support using lexical decision as a groupwise test to screen for poor word readers. Form A of the test and provisional scales are presented for each primary grade.

Keywords: Word decoding, reading development, reading test, pencil-and-paper test, lexical decision task.

Introducción

Los profesores de primaria utilizan una amplia variedad de estrategias de enseñanza y de evaluación de la lectura (Lacina & Block, 2011). El uso de pruebas o exámenes breves e individuales de lectura en voz alta es parte frecuente de la rutina de clase para supervisar el progreso de la lectura a lo largo del curso escolar. Si estas pruebas están organizadas adecuadamente dentro de la rutina de enseñanza, ofrecen un beneficio máximo a un coste mínimo para el niño y la escuela porque un profesor bien formado suele ser más eficiente en la detección de problemas que muchos tests (Wilson & Jungner, 1968). Además, los profesores pueden establecer objetivos y adaptar su enseñanza sin demora si perciben que el progreso del niño es insuficiente (Förster & Souvignier, 2015; Förster, Kawohl, & Souvignier, 2018). Sin embargo, la aplicación rutinaria de pruebas de lectura en voz alta es difícil cuando el número de niños es muy grande, o cuando se necesitan varias mediciones para seguir el progreso de los estudiantes e identificar rápidamente a aquellos que están quedándose atrás, con riesgo de fracasar.

Para evaluar de lectura de forma frecuente y rápida y en grandes grupos de niños, es necesario tener tests que cumplan estándares científicos, idealmente con formas paralelas. En España, existen tests bien fundamentados para evaluar la lectura (incluyendo la lectura de palabras) en la escuela primaria (Cuetos, Rodríguez, Ruano y Arribas, 2007; Defior et al., 2006; Jiménez, Gove, Crouch y Rodríguez, 2014) y tests breves que miden aspectos específicos de la lectura, como la fluidez lectora (González-Trujillo, Calet, Defior y Gutiérrez-Palma, 2014), pero su administración individual los hace laboriosos. Hasta donde sabemos, el único test breve para grupos evalúa la comprensión lectora (Marín y Carrillo, 1999).

No conocemos ninguna prueba breve grupal para medir el reconocimiento visual de palabras o la lectura de palabras, es decir, la recuperación de la fonología y el significado de una palabra, a pesar de que estas habilidades determinan en gran medida la comprensión lectora (Jenkins, Fuchs, Van den Broek, Espin, & Deno, 2003; Kim, Petscher, Schatschneider, & Fooman, 2010; Klauza & Guthrie, 2008; Perfetti & Stafura, 2014). Si la identificación de palabras funciona mal se desencadena un círculo vicioso que limita la progresión de la comprensión lectora, es decir, el proceso lector que permite aprender. Los niños con problemas para lograr el reconocimiento visual de palabras tienen dificultad para alcanzar la fluidez lectora. Su lectura es laboriosa y su motivación para leer, su práctica lectora y su comprensión lectora

suelen ser bajas (Rasinski, Reutzel, Chard y Linan-Thompson, 2011). De hecho, la investigación sobre lo que realmente miden los tests de comprensión lectora muestra que la variabilidad en esta medida se debe principalmente a la habilidad para leer palabras (Keenan & Meenan, 2014).

La evaluación de la lectura de palabras, ya sea en papel y lápiz o computarizada, se realiza principalmente con dos tareas: nombrado o denominación (lectura en alto o lectura en voz alta) y decisión léxica (decidir si una cadena de letras es una palabra existente o no). La evidencia sobre la validez de estas tareas proviene de estudios que las han comparado con la lectura silenciosa, específicamente con los tiempos de fijación ocular. Los tiempos de fijación ocular podrían considerarse el estándar de oro porque se registran durante la lectura natural, y se ha demostrado que reflejan los procesos cognitivos de la lectura (por ejemplo, Foster, Ardoin y Binder, 2018; Rayner y Reichle, 2010; Rayner, Sereno, Morris, Schmauder y Clifton, 1989). Los estudios con participantes de habla inglesa revelan que, aunque los tiempos de respuesta de la tarea de nombrado de la palabra son más semejantes a los tiempos de fijación ocular, la tarea de decisión léxica también tiene una fuerte correlación con los tiempos de fijación ocular, lo que demuestra que ambas tareas, nombrado y decisión léxica, ofrecen datos válidos para estimar la lectura silenciosa (Schilling, Rayner, & Chumbley, 1998; véase también Forster, 1976; Forster & Chambers, 1973). Otra pregunta es si los resultados anteriores se replican en diferentes ortografías. Faltan datos sobre los movimientos oculares durante la lectura en español, una ortografía poco profunda y con una estructura silábica simple comparada con el inglés, pero Kuperman, Drieghe, Keuleers y Brysbaert (2013) muestran que los datos en holandés (relativamente poco profundo) e inglés (muy profundo) presentan el mismo patrón de correlaciones entre los movimientos oculares, la decisión léxica y las latencias de denominación. Nótese además que estos autores encuentran un patrón diferente al de la investigación anterior, este es, existe una correlación más fuerte entre las latencias de los movimientos oculares y la decisión léxica que entre las primeras y las latencias de nombrado.

El test que se presenta en este estudio se basa en trabajos anteriores. En primer lugar, nuestra experiencia con las tareas de decisión léxica en ordenador (Goikoetxea & Ferrero, 2019). Aunque el ordenador permite registrar el tiempo de respuesta y parece más escalable que la evaluación en papel y lápiz, en realidad impone una limitación en el número de niños evaluados simultáneamente debido al gran número de errores cometidos por los niños mientras trabajan con ordenadores (Moret-Tatay & Perea, 2011). Por lo tanto, decidimos crear una

prueba de papel y lápiz. En segundo lugar, el LEO-1-min se basa en una prueba que mide la lectura de palabras en holandés: el Doorstreepleestoets o la tarea de decisión léxica de papel y lápiz de Van Bon (2007; véase también Van Bon, Hoevenaars, & Jongeneelen, 2004; van Bon & Libert, 1997; Van Bon, Tooren, & Van Eekelen, 2000). Esta prueba fue desarrollada para evaluar a lectores pobres de 2º y 3er grado de la escuela primaria que necesitan medidas de apoyo o refuerzo. La tarea de decisión léxica, realizada en 1 minuto, demostró una buena correlación entre las formas paralelas (media de las r s: 0,82 para 2º grado y 0,66 para 3er grado), lo que indica una buena fiabilidad de las puntuaciones. Una baja correlación con una prueba de símbolos (rango de las r s: de 0,12 a 0,25 para 2º grado y de 0,07 a 0,20 para 3er grado) indica que las habilidades motoras no fueron decisivas en el rendimiento. También mostró la validez del criterio debido a las altas correlaciones con las puntuaciones en una prueba estandarizada de lectura en voz alta (media de las r s: 0,79 para el 2º grado y 0,67 para el 3er grado). Estos resultados muestran que esta prueba fue apropiada para medir la lectura de palabras en niños de 2º y 3er grado. Por lo tanto, hemos basado la construcción de nuestra nueva prueba en Doorstreepleestoets.

Objetivos e hipótesis

Los objetivos del presente estudio fueron, en primer lugar, crear una prueba de lectura de palabras llamada LEO-1-min, utilizando la tarea de decisión léxica, con cuatro formas alternativas. Según la Asociación Americana de Investigación Educativa, la Asociación Americana de Psicología y el Consejo Nacional de Medición en la Educación (2014), dos o más versiones de un examen que se consideran intercambiables se denominan formas alternativas (o equivalentes). Por consiguiente, diseñamos las cuatro formas del LEO-1-min para tener la misma distribución general de contenido, formato de los ítems y procedimiento de administración. Aunque las formas alternativas no eliminan completamente los efectos de la práctica o la memoria, sí reducen estos efectos porque los elementos no son idénticos. Basados en estas consideraciones, esperamos que las cuatro formas del LEO-1-min tengan aproximadamente la misma puntuación media y las mismas desviaciones estándar en cada curso escolar. El segundo objetivo fue investigar la fiabilidad y la validez convergen-discriminante de LEO-1-min. Basados en el proceso de construcción, esperamos que la fiabilidad paralela sea satisfactoria, el LEO-1-min se correlacione más con una prueba de lectura en voz alta que con una prueba de aritmética oral, y que un análisis discriminante revele que las puntuaciones en el LEO-1-min predicen los juicios de los profesores acerca del nivel de lectura de sus estudiantes. El LEO-1-min está diseñado para evaluar la lectura en niños de educación primaria, se-

cundaria y bachiller, y en personas adultas, pero en este trabajo sólo se muestran las propiedades psicométricas de las formas A y B en una muestra de 1er a 6º curso de educación primaria, y de las formas C y D en una muestra de tres cursos de educación primaria.

Método

Participantes

En total, 284 niños de 1º a 6º curso de una escuela primaria concertada (parcialmente subsidiada) de Guecho, Vizcaya, participaron en el experimento: 48 de 1er curso (31,3% niñas; 35,4% datos no disponibles; edad media = 5,94 años, $DT = 0,24$), 47 de 2º curso (51% niñas, 21% datos perdidos; edad media = 7,0; $DT = 0,21$), 51 de 3er curso (14% niñas; 72% datos perdidos; edad media = 7,96; $DT = 0,20$), 42 de 4º curso (31% niñas, 33% datos perdidos; edad media = 8,98; $DT = 0,27$), 50 de 5º curso (18% niñas; 70% datos perdidos; edad media = 10,08; $DT = 0,40$), 46 de 6º curso (33% niñas, 33% datos perdidos; edad media = 11,00; $DT = 0,30$). La escuela atiende a población de clase media. Además, diecisiete participantes fueron excluidos de los análisis debido a errores en el uso de la prueba. Un niño mostró tener necesidades educativas especiales. Sus datos también se excluyeron de los cálculos que figuran a continuación. El muestreo fue incidental. La lengua materna y primera lengua de enseñanza en la escuela era el castellano, que representaba el 65% de las horas de clase, mientras que el 35% restante eran en euskera y en inglés. La escuela utilizó un método alfabético para enseñar a leer y escribir, donde los nombres de las letras y sus sonidos se aprendían en combinación con cada una de las 5 vocales.

Instrumentos

LEO-1-min es una prueba de velocidad diseñada para determinar la fluidez del reconocimiento visual de palabras en lectores de una amplia gama de edades y competencias de lectura. Se pide a los niños que identifiquen las pseudopalabras entre las palabras reales de una lista de ítems, durante un minuto.

Las palabras de las cuatro formas alternativas del LEO-1-min, cada una con 180 estímulos, fueron seleccionadas al azar de entre los 528 sustantivos más frecuentes, excluyendo nombres y coloquialismos (32 palabras de una sílaba, 279 de dos sílabas, 166 de tres sílabas, 43 de cuatro sílabas, y 8 de cinco sílabas) seleccionados del vocabulario de LEXIN (Corral, Ferrero, & Goikoetxea, 2009). La lista de 528 palabras se utilizó para crear las pseudopalabras sustituyendo las letras centrales según la longitud de las palabras. En palabras monosilábicas, la última letra de cada palabra fue cambiada (por ejemplo, gas a gar), en palabras bisilábicas, la primera consonante de la segunda sílaba fue sustituida (por ejemplo, mesa a meda).

En palabras con tres o más sílabas, la primera consonante en la segunda y tercera sílabas fue sustituida (p. ej., mañana a mavaga). Se mantuvo la estructura silábica de las palabras.

A continuación, se examinó si las formas alternativas eran similares en las variables psicolingüísticas con efectos conocidos sobre el reconocimiento visual de palabras en español: frecuencia léxica (Defior, Justicia y Martos, 1996), también controlada en estudios anteriores (Van Bon, Hoevenaars y Jongeneelen, 2004), longitud (Acha y Perea, 2008), vecindad (Perea y Rosa, 2000) y frecuencia silábica (Carreiras, Álvarez y de Vega, 1993). Cada lista incluía 132 sustantivos. La frecuencia media de estos sustantivos es 66,77 (rango: 37,13-101,4) en LEXIN y 119,56 (rango: 30,89-850,89) en LEXESP (Sebastián-Gallés, Martí, Carreiras, & Cuetos, 2000). La longitud media es de 5,19 letras (rango: 3-12), y la vecindad media es de 3,39 (rango: 0-23). La frecuencia silábica media de los sustantivos es 876,97 (rango: 73,45- 6737,32) en la primera sílaba, 1289,43 (rango: 34,46- 9665,36) en la segunda, 1682,04 (rango: 48,57- 9600,36) en la tercera, y 2008,47 (rango: 51,43- 6210,18) en la cuarta, en Buscapalabras (Davis & Perea, 2005). Cada lista incluía una proporción de tres palabras por cada pseudopalabra, como en la prueba anterior de primaria (Van Bon et al., 2000, 2004), con listas de 180 estímulos: 132 sustantivos y 48 pseudopalabras en orden aleatorio.

En cada lista, los estímulos se dispusieron en cuatro columnas. La fuente es Arial 13, a un solo espacio. Las pseudopalabras se distribuyeron entre las cuatro columnas, y al menos una pseudopalabra aparece en la mitad inferior de la última columna, y nunca como el primer estímulo en la primera columna. Los estudiantes deben ir columna por columna para subrayar las pseudopalabras. Después de haber hecho esto durante un minuto, escriben una X al lado de la última palabra vista. Las instrucciones y estímulos de la forma A del LEO-1-min se encuentran en el Apéndice A. La prueba completa con las cuatro formas alternativas se puede obtener gratuitamente solicitando una copia directamente al primer autor o en PsycTESTS (Base de datos de la Asociación Americana de Psicología). La prueba de Símbolos fue creada para estimar la influencia de los componentes perceptivo-motores de la tarea de subrayado en las puntuaciones del LEO-1-min. La lista incluía 180 cadenas de letras compuestas por la misma letra del alfabeto (excepto la X), repetidas un mínimo de 3 veces y un máximo de 12. En 48 de las 180 cadenas, una o más letras fueron reemplazadas por una X. La tarea requería subrayar las cadenas que contenían una o más X, hasta que el experimentador dijera 'Parad'. La puntuación fue el número de ítems correctamente completados en un minuto.

Lectura de palabras y lectura de pseudopalabras de la Batería para la Evaluación de Procesos de Lectura en Primaria revisada (PROLEC-R; Cuetos et al., 2007), una de las pruebas más utilizadas en España para medir la lectura de palabras. Los subtests de Lectura de Palabras y Lectura de Pseudopalabras requieren que el niño lea en alto una lista de palabras y pseudopalabras, respectivamente, y que se registre la precisión y el tiempo. La fiabilidad de la consistencia interna de las puntuaciones de Lectura de Palabras y Lectura de Pseudopalabras en esta muestra fue de 0,81 y 0,83, respectivamente.

Aritmética de la Escala de Inteligencia de Wechsler para Niños IV (WISC-IV; Wechsler, 2005). El subtest de Aritmética consiste en problemas aritméticos que el niño debe resolver mentalmente en un tiempo limitado. Este subtest tiene diferentes puntos de partida según la edad. La fiabilidad de la consistencia interna de las puntuaciones de Aritmética en esta muestra fue de 0,86.

Procedimiento

La dirección de la escuela y los adultos responsables de los participantes dieron su consentimiento informado para llevar a cabo el experimento. Los niños fueron evaluados en la escuela por dos de los autores y tres colaboradores más previamente capacitados para realizar la tarea. Cada curso de 1º a 6º incluyó dos aulas. Todos los participantes completaron dos de las cuatro formas del LEO-1-min y la tarea de símbolos. La forma A y la tarea de símbolos se administraron a todos los participantes de 1º a 6º de las dos aulas, la forma B se administró a la mitad de la muestra, es decir, a los estudiantes de un aula por curso, la forma C a los estudiantes de un aula de 1er, 3er y 5º curso, y la forma D a los estudiantes de un aula de 2º, 4º y 6º curso. La administración de ambas formas se realizó en la misma sesión con un intervalo de unos pocos minutos. Debido a limitaciones prácticas, el orden de aplicación de las dos formas no se balanceó; la forma A fue la primera en administrarse cuando la forma alternativa era la B, y lo contrario cuando las formas alternativas eran las C o D, siempre terminando con la tarea de los símbolos. Las instrucciones para los símbolos eran las mismas que para el LEO-1-min, pero considerando los estímulos objetivo aquellos que incluían X.

Para estudiar la validez, se solicitó a los profesores de las aulas que eligieran y clasificaran a tres niños de cada aula, en tres categorías: "muy buen lector" o categoría 1, "lector normal" o categoría 2 y "mal lector" o categoría 3. Los profesores clasificaron así a 108 niños de 1er a 6º curso (6 cursos x 2 aulas x 3 categorías x 3 niños). Esta submuestra completó la prueba PROLEC-R y la prueba Aritmética en un orden contrabalanceado y siguiendo un pro-

cedimiento de doble ciego, es decir, sin que los evaluadores ni los niños conocieran los datos del LEO-1-min ni la clasificación realizada por los profesores.

Análisis estadístico

Para conocer las principales comparaciones entre las formas de la prueba, curso y sexo, se utilizó un análisis multivariado de varianza (MANOVA) de los diferentes tipos de respuesta, es decir, omisiones correctas, aciertos, falsas alarmas y omisiones. Las comparaciones post-hoc de Tukey se utilizaron para examinar los efectos de interés entre los grupos cuando eran más de dos. Para estimar la fiabilidad, se empleó la correlación de Pearson entre las dos formas del LEO-1-min y, posteriormente, se utilizó un análisis de varianza (ANOVA) de medidas repetidas con la forma como factor intrasujeto para examinar la diferencia entre los cursos. La correlación de Pearson también se utilizó en el análisis convergente-discriminante utilizando las puntuaciones LEO-1-min, el PROLEC-R, y Aritmética del WISC, y en el análisis de la equivalencia de las formas al correlacionar las puntuaciones LEO-1-min y variables demográficas de la muestra. Finalmente, se realizó un análisis discriminante predictivo utilizando las puntuaciones de LEO-1-min como predictor y el nivel de rendimiento lector según el juicio de los profesores. La normalidad e igualdad de la matriz de varianza-covarianza en estos análisis fueron probadas con las pruebas de Komogorov y Box, respectivamente. Se utilizó el método dejando-uno-fuera para estimar la tasa de clasificación y la Z de Huberty's para evaluar estadísticamente nuestra clasificación. Todos los análisis estadísticos se realizaron con una probabilidad del error de tipo I fijada en 0,05. Los análisis se realizaron utilizando el software SPSS, versión 21.0.

Resultados

Descriptivo

Las respuestas se etiquetaron como "aciertos" para pseudopalabras subrayadas correctamente, "omisiones correctas" para palabras reales que no estaban subrayadas, "falsas alarmas" si las palabras reales estaban subrayadas, y "omisiones" para pseudopalabras que fueron omitidas y no estaban subrayadas (Figura 1). Luego, el número de respuestas correctas (efectividad total) se calculó como el número total de omisiones correctas y aciertos (esto es lo mismo que el número de estímulos intentados o el número total de elementos leídos por cada niño menos las omisiones y las falsas alarmas). Los estadísticos descriptivos para este tipo de respuestas en el LEO-1-min y las puntuaciones totales en la tarea de símbolos se muestran en la Tabla 1.

Se realizó un MANOVA con los tipos de respuesta - omisiones correctas, aciertos, falsas alarmas y omisiones - en cada forma de LEO-1-min con el curso y el sexo como factores independientes. Para las Formas A y B, el sexo no tuvo ningún efecto y no interactuó con el curso. El curso, sin embargo, tuvo un efecto estadísticamente significativo (ver Tabla 1). Las comparaciones post-hoc de Tukey revelaron diferencias estadísticamente significativas ($p < .05$) entre 1° y todos los demás cursos; 2° y todos los demás cursos; 3° y 4° y todos los demás cursos, pero no entre estos dos cursos. Este resultado puede explicarse por la típica disminución del crecimiento de la lectura de los niños después de los primeros años de aprendizaje, pero también podría estar relacionado con la baja variabilidad de las puntuaciones del 4° curso de esta muestra. Los cursos 5^a y 6^a no mostraron ninguna diferencia entre ellos.

Tabla 1. *Puntuación total en las formas A y B de LEO-1-min y en la tarea de Símbolos por curso*

	Grade						$F(5,283)$	p	η^2
	1	2	3	4	5	6			
Forma A	$n = 46$	$n = 47$	$n = 51$	$n = 42$	$n = 50$	$n = 46$			
Puntuación total (0-180)	15.17 (10.75)	28.74 (11.59)	41.53 (13.09)	48.17 (11.71)	61.62 (17.55)	63.24 (15.71)	85.38	<.001	.61
Omisiones correctas (0-132)	9.43 (8.27)	19.26 (9.32)	29.53 (10.31)	34.40 (9.34)	45.82 (14.55)	46.35 (12.56)	81.40	<.001	.60
Aciertos (0-48)	5.74 (2.99)	9.49 (2.59)	12.00 (3.01)	13.76 (2.77)	15.80 (3.55)	16.89 (3.38)	83.23	<.001	.60
Falsas alarmas (0-132)	1.33 (2.03)	0.49 (0.72)	0.25 (0.56)	0.12 (0.33)	0.16 (0.37)	0.35 (0.60)	10.16	<.001	.16
Omisiones (0-48)	1.41 (1.57)	1.36 (1.11)	1.55 (1.12)	1.48 (1.38)	2.26 (2.25)	1.35 (1.30)	2.65	.023	.05
	Grade								
Forma B	$n = 26$	$n = 24$	$n = 27$	$n = 20$	$n = 26$	$n = 23$	$F(5,145)$	p	η^2
Puntuaciones totales (0-180)	14.54 (11.62)	25.92 (11.77)	41.15 (11.93)	43.20 (10.63)	53.12 (16.55)	57.04 (16.75)	36.12	<.001	.56
Omisiones correctas (0-132)	10.15 (8.15)	17.17 (8.32)	29.22 (9.24)	30.75 (8.30)	37.23 (11.61)	39.74 (11.20)	35.70	<.001	.56

Aciertos (0-48)	4.38 (3.79)	8.75 (8.33)	11.93 (3.10)	12.45 (2.70)	15.88 (5.04)	16.43 (4.91)	33.36	<.001	.54
Falsas alarmas (0-132)	1.38 (1.24)	0.88 (0.90)	0.52 (0.75)	0.05 (0.22)	0.31 (0.84)	0.22 (0.52)	8.64	<.001	.24
Omisiones (0-48)	1.15 (1.64)	0.42 (0.58)	1.11 (1.53)	1.10 (1.17)	0.54 (0.95)	1.04 (1.22)	1.70	.14	.06
Forma C	<i>n</i> = 20		<i>n</i> = 24		<i>n</i> = 23		<i>F</i> (5, 67)	<i>p</i>	η^2
Puntuación total (0-180)	16.05 (8.64)	42.63 (15.04)		70.09 (21.53)		59.58	<.001	.65	

Tabla 1 (continúa)

	Grade						<i>F</i> (5,283)	<i>p</i>	η^2
	1 <i>n</i> = 46	2 <i>n</i> = 47	3 <i>n</i> = 51	4 <i>n</i> = 42	5 <i>n</i> = 50	6 <i>n</i> = 46			
Forma A							<i>F</i> (5,283)	<i>p</i>	η^2
Omisiones Correctas (0-132)	10.55 (5.85)	29.83 (10.77)		49.83 (15.87)		59.77	<.001	.65	
Aciertos (0-48)	5.50 (2.89)	12.79 (4.35)		20.39 (6.49)		50.00	<.001	.61	
Falsas alarmas (0-132)	0.90 (0.97)	0.63 (0.65)		1.17 (1.23)		1.87	.163	.06	
Omisiones (0-48)	0.45 (0.69)	0.33 (0.48)		0.65 (0.93)		1.61	.163	.04	
Forma D	<i>n</i> = 23		<i>n</i> = 22		<i>n</i> = 24		<i>F</i> (2, 68)	<i>p</i>	η^2
Puntuación total (0-180)	32.26 (9.55)		50.95 (15.48)		65.88 (15.48)		36.45	<.001	.53
Omisiones correctas (0-132)	22.65 (7.25)		36.36 (10.01)		46.33 (11.01)		36.13	<.001	.52
Aciertos (0-48)	9.61 (2.41)		14.59 (4.91)		19.54 (4.63)		33.94	<.001	.51
Falsas alarmas (0-132)	0.78 (1.04)		0.55 (0.72)		1.08 (1.18)		1.65	.201	.05
Omisiones (0-48)	1.26 (0.75)		1.00 (0.93)		0.67 (0.76)		3.15	.049	.09
Símbolos	<i>n</i> = 46	<i>n</i> = 47	<i>n</i> = 51	<i>n</i> = 42	<i>n</i> = 50	<i>n</i> = 47	<i>F</i> (5,282)	<i>p</i>	η^2
Puntuación total (0-180)	48.33 (12.46)	63.34 (14.42)	78.04 (13.66)	83.07 (14.49)	102.06 (16.51)	108.40 (23.46)	92.29	<.001	.63

Este patrón fue idéntico para las puntuaciones totales, omisiones correctas y aciertos. Sin embargo, el 1er curso fue el único que hizo más falsas alarmas y omisiones que todos los demás, probablemente porque para los lectores principiantes unas pocas letras son suficientes para reconocer una palabra (falsas alarmas) pero tienen un léxico limitado (omisiones). De manera similar, para las Formas C y D, el curso tuvo un efecto significativo (ver Tabla 1), pero el sexo no. Las comparaciones posthoc de Tukey mostraron diferencias significativas entre los cursos 1, 3 y 5 en el Forma C, y diferencias entre los cursos 2, 4 y 6 en el Forma D.

La respuesta de subrayar las pseudopalabras tras detectarlas, involucra innegablemente habilidades perceptivo-motoras. Ahora bien, mientras más estímulos pueda procesar un niño en la prueba de símbolos, en comparación con los que procesa en LEO-1-min, más participan en el rendimiento en el LEO-1-min otras habilidades que no son las perceptivo-motoras. Por lo tanto, eran de esperar las diferencias que se muestran en la Tabla 1.

Se calcularon las correlaciones entre las puntuaciones de las palabras y de los símbolos; una alta correlación indicaría una importante contribución de las habilidades motoras en ambas tareas. Estas correlaciones se muestran en la Tabla 2. Los resultados muestran correlaciones moderadas y positivas entre el desempeño en el LEO-1-min en la mayoría de los cursos, excepto en los cursos 3° y 4°, donde la correlación disminuye.

Fiabilidad de las formas paralelas

Primero, se calculó la correlación de Pearson entre las puntuaciones de las dos formas del LEO-1-min. Los resultados de este análisis, que se muestran en la Tabla 2, muestran correlaciones positivas y de moderadas a altas a lo largo de la educación primaria.

Para examinar más a fondo la fiabilidad de las puntuaciones en las formas alternativas del LEO-1-min, se compararon las medias de las formas de la prueba, buscando diferencias que cuestionaran su equivalencia. Un ANOVA de medidas repetidas con la forma como factor intrasujeto (dos niveles: A y B), y la puntuación como factor entre-sujetos, comparó las formas del LEO-1-min. No hubo diferencias estadísticamente significativas en el rendimiento entre las formas A y B en ningún curso: $F(1,25) = 1,97, p = 0,173, \eta^2 = 0,073$ en 1er curso, $F(1,23) = 1,97, p = 0,173, \eta^2 = 0,079$ en 2° curso, $F(1,26) = 1,06, p = 0,313, \eta^2 = 0,039$ en 3er curso, $F(1,19) = 2,93, p = 0,103, \eta^2 = 0,134$ en 4° curso, $F(1,25) = 1,14, p = 0,296, \eta^2 = 0,044$ en 5° curso, $F(1,22) = 1,71, p = 0,205, \eta^2 = 0,072$ en 6° curso. Un segundo ANOVA con las formas A y D no mostró diferencias en ningún curso, $F(1,22) = 2,65, p = 0,118, \eta^2 = 0,107$ en 2° curso, $F(1,21) < 1$ en 4° curso, y F

(1,28) < 1 en 6º curso. Un tercer ANOVA con las formas A y C, sin embargo, mostró diferencias significativas en dos cursos: en 1er curso, $F(1,19) = 4,84$, $p = 0,040$, $\eta^2 = 0,203$ y en 3er curso, $F(1,23) = 6,72$, $p = 0,016$, $\eta^2 = 0,226$, pero no hubo diferencias en el 5º curso, $F(1,22) < 1$.

Table 2

Correlaciones de Pearson entre las Formas A y Otras Formas de LEO-1-min y la Tarea de Símbolos

Grade	Forma A			Forma B			Forma C			Forma D		
	<i>n</i>	<i>r</i>	Todas las demás	<i>n</i>	<i>r</i>	Forma A	<i>n</i>	<i>r</i>	Form A	<i>n</i>	<i>r</i>	Símbolos
1	46	.81**	.40**	26	.84**	.32	20	.85**	.60**			
2	47	.76**	.51**	24	.84**	.69**				23	.67**	.52*
3	51	.81**	.19	27	.65**	.40*	24	.94**	.20			
4	42	.57**	.27	20	.56**	.26				22	.57**	.05
5	50	.72**	.49**	26	.63**	.07	23	.83	.63**			
6	47	.69**	.46**	23	.68**	.49**				24	.68**	.48*

* $p < .05$. ** $p < .01$. *** $p < .001$.

Validez

Se llevaron a cabo dos tipos de análisis para examinar las relaciones con otras variables, específicamente, pruebas de convergencia y discriminación y pruebas concurrentes con las puntuaciones del test LEO-1-min.

La evidencia convergente y discriminante acerca del LEO-1-min como una tarea para medir el desempeño en la lectura de palabras fue examinada calculando las correlaciones de Pearson entre las puntuaciones en la forma A de LEO-1-min y las puntuaciones en los subtests de Lectura de Palabras y Lectura de Pseudopalabras del PROLEC-R y el subtest Aritmética del WISC-IV. La Tabla 3 muestra las altas correlaciones logradas en todos los cursos entre el rendimiento en la forma A del LEO-1-min y el PROLEC-R, especialmente, para la lectura de palabras. Por el contrario, las correlaciones de la forma A de LEO-1-min con las puntuaciones en Aritmética fueron más bajas, y la mayoría de ellas no fueron significativas, en ningún curso, excepto en 1º y 2º. Cabe señalar que el 4º curso mostró menores correlaciones con el PROLEC-R y con Aritmética. Este resultado se debe posiblemente a la menor variabilidad de las puntuaciones de LEO-1-min en este curso.

Tabla 3. *Correlaciones de Pearson entre las Respuestas Correctas de LEO-1-min Forma A, el PROLEC-R y Aritmética por Curso.*

	PROLEC-R		Arithmetic
	Words	Pseudowords	
LEO-1-min Form A			
1 (n= 17)	.60*	.51*	.56*
2 (n= 18)	.74**	.60**	.59**
3 (n= 16)	.85**	.82**	.44
4 (n= 17)	.55*	.38	.24
5 (n= 14)	.72**	.70**	.38
6 (n= 17)	.58*	.58*	.22

* $p < .05$. ** $p < .01$. *** $p < .001$.

La equivalencia es otro índice de validez. Si varias formas de un test miden lo mismo, estas formas deberían mostrar una correlación de magnitud y signo semejante con una variable demográfica. La Tabla 4 muestra las magnitudes con las que las formas del LEO-1-min correlacionan con la edad y el género. Las cuatro formas correlacionan moderadamente con la edad; por el contrario, las correlaciones con el género son modestas.

Se utilizó el análisis discriminante predictivo para examinar hasta qué punto la puntuación del test predijo el nivel de lectura asignado por los profesores. La variable predictiva fue la puntuación del LEO-1-min en la forma A, y la variable de agrupamiento fue el nivel de lectura. Para maximizar la diferencia en los niveles de lectura, agrupamos juntos a los niños de los niveles normal y alto. De esta manera, el grupo de rendimiento normal-alto estaba formado por quienes habían recibido puntuaciones de 1 y 2 de sus profesores, y el grupo de rendimiento bajo estuvo formado por aquellos que habían recibido una puntuación de 3 de sus profesores. La normalidad de la variable predictiva se cumplió según la prueba de Kolmogorov ($p = 0,064$), y la prueba de Box para la igualdad de la matriz de varianza-covarianza no fue significativa. $F < 1$. El análisis discriminante reveló una función discriminante significativa, $\chi^2(1, N = 102) = 18,747, p < ,000$ (autovalor = ,207, correlación canónica = ,414, Wilks's $\Lambda = ,828$). La media de los grupos centroides en la función discriminante reveló que el grupo de lectura normal-alta alcanzó una media positiva (.347), y el grupo de bajo nivel alcanzó una media negativa (-,585). Estos resultados muestran que el grupo de rendimiento normal-alto

lee un número mayor de palabras e identifica un número mayor de pseudopalabras en el LEO-1-min que el grupo de bajo rendimiento. Aplicando el método dejando-uno-fuera para controlar la tendencia a sobreestimar la exactitud de las tasas de clasificación, la tasa de clasificación correcta para toda la muestra fue del 72%. La Z de Huberty indicó que la tasa de clasificaciones correctas es estadísticamente mejor de lo que cabría esperar por azar, $z = 2,02$, $p = 0,002$.

Normas. Para completar la creación del LEO-1-min, se elaboraron baremos con la muestra de estudio (1° a 6° curso) que permiten transformar las puntuaciones directas en centiles (ver Tabla 5). Esto permite situar a los niños dentro de su grupo normativo de referencia. Obsérvese que un resultado de interés no previsto es que los cuartiles del LEO-1-min coinciden con las puntuaciones medias que describen a los niños de los tres niveles de lectura según el juicio de los profesores.

Tabla 4. *Correlaciones de Pearson entre Edad, Sexo y Formas de LEO-1-min.*

	LEO-1-min			
	Form A n = 283	Form B n = 146	Form C n = 67	Form D n = 69
Edad	.765**	.729**	.805**	.725**
Sexo ^a	-.017	.002	.038	-.063

^aDebido a la falta de datos de sexo, n para las Formas A, B, C y D son 155, 64, 32 y 59, respectivamente.

** $p < .01$.

Discusión y conclusión

El estudio ilustra la potencial utilidad del LEO-1-min para medir la lectura de palabras en un amplio rango de edades, aunque aquí sólo discutimos la evidencia obtenida en una muestra de niños de primaria. El desempeño en la prueba mostró el progresivo desarrollo de la lectura a lo largo de la educación primaria, ya observado en estudios anteriores en español (Goikoetxea & Ferrero, 2019). Los niños de 1er y 2º curso progresan aceleradamente, con un crecimiento más lento a partir del 3er curso, lo que hace que los cursos de 3º y 4º y también los cursos de 5º y 6º sean muy similares entre sí. Es interesante observar de nuevo el bajo número de errores incluso desde el 1er curso, como se observa en español (Valle-Arroyo, 1989) y otras ortografías transparentes (Landerl, & Wimmer, 2008) en comparación con las opacas

(Seymour, Aro, Erskine, & COST Action A8, 2003; Simões & Alves, 2018). A diferencia de Van Bon (2007), que encontró una diferencia significativa, pero pequeña (y probablemente también socialmente irrelevante), no encontramos ningún efecto del sexo en el desempeño del LEO-1-min. Sin embargo, la pérdida de datos sobre el sexo de los niños de nuestro estudio hace que este resultado sea tentativo. También en contraste con Van Bon (2007), el rendimiento en el LEO-1-min se correlaciona con el rendimiento en la tarea de los símbolos. Una explicación tentativa de estos resultados es que Van Bon (2007) usó letras griegas en la tarea de los símbolos que no tienen sentido para los lectores de nuestros alfabetos. Pero aquí, usamos cadenas de letras, y gracias a tareas como la de Stroop sabemos que incluso cadenas de letras sin sentido desvían la atención y retrasan el nombrado de colores e imágenes porque los lectores competentes no pueden evitar leer.

Con respecto a la equivalencia de las formas LEO-1-min, las formas A y B muestran puntuaciones medias similares, y la distribución de las puntuaciones es también bastante similar en ambas formas. Debido al alto nivel de comparabilidad entre las formas A y B, estas dos formas pueden considerarse equivalentes en dificultad. Las formas C y D también revelan medias y desviaciones típicas similares a las de la forma A, al menos en los tres cursos comparados. Además, la magnitud de los coeficientes de fiabilidad del retest es de moderada a alta en todos los cursos, alcanzando los requisitos mínimos para una prueba que se utilizará en la toma de decisiones en muestras similares a la de este estudio. Los coeficientes de fiabilidad son similares a los obtenidos por los niños holandeses de 2º y 3er grado (Van Bon et al., 2004). Estos resultados sugieren que las formas A y B pueden considerarse equivalentes, lo que las hace útiles para exámenes repetidos a lo largo de la educación primaria. Aunque no examinamos el efecto del orden y, en consecuencia, no conocemos el efecto de la práctica en el LEO-1-min, la similitud de las medias sugiere que no hay tal efecto de orden. Sin embargo, las investigaciones futuras deben abordar esta cuestión.

En cuanto a la evidencia de la validez convergente-discriminante, se puede observar que las puntuaciones en el LEO-1-min correlacionan fuertemente con la lectura en alto en el PROLEC-R, y en menor medida con el cálculo matemático en el subtest de Aritmética. La correlación entre la decisión léxica y la lectura en voz alta apoya la validez de la tarea de decisión léxica como medida de los procesos cognitivos en la lectura de palabras. La menor correlación con el subtest de Aritmética muestra que LEO-1-min no mide un factor general de rendimiento escolar general, sino más bien habilidades específicas de lectura. Además, las puntuaciones del LEO-1-min discriminan entre los grupos de lectores buenos y pobres identi-

ficados por los profesores. Esta evidencia de la validez concurrente de la prueba es un indicador de su posible utilidad en la realización de actividades que aumenten la habilidad para leer palabras y la fluidez en la lectura.

A pesar de la pequeña y reducida muestra de este estudio, se presentan baremos provisionales basados en los datos presentados. Estos baremos se pueden emplear en el contexto escolar y en la investigación sobre el rendimiento lector, pero con cautela porque nuestra muestra procede de una sola escuela de la provincia de Vizcaya.

Las limitaciones obvias de este estudio, que justifican una mayor investigación, son el hecho de que no hemos proporcionado datos de rendimiento a partir de la evaluación de un muestreo representativo de la población de niños de primaria, o datos de estabilidad de las puntuaciones (test-retest), ni tampoco pruebas de validez de las puntuaciones a lo largo del tiempo. Asimismo, es necesario conocer las propiedades psicométricas del LEO -1-min en muestras representativas para evaluar su uso en poblaciones juveniles y adultas de lengua española. Los datos aquí ofrecidos no pueden generalizarse más allá de las características de los participantes en este estudio.

En conclusión, las puntuaciones del LEO-1-min muestran fiabilidad y validez como medida de la habilidad de lectura de palabras en niños hispanohablantes con las características de la muestra que participó en este estudio. Una buena enseñanza de la lectura requiere evaluaciones repetidas de los mismos individuos a lo largo del tiempo. LEO-1-min permitirá a los profesores e investigadores una evaluación rápida, grupal y repetida de las habilidades de lectura de palabras de los niños. Hoy se sabe que la detección temprana del bajo rendimiento en lectura en los cursos de primaria es la mejor intervención de remedio porque permite a los profesores reaccionar inmediatamente, adaptando su enseñanza de la lectura a las necesidades individuales.

Referencias

Acha, J., & Perea, M. (2008). The effects of length and transposed-letter similarity in lexical decision: Evidence with beginning, intermediate, and adult readers. *British Journal of Psychology*, 99 (2), 245-264. doi:10.1348/000712607x224478

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: Authors.
- Carreiras, M., Alvarez, C.J., & de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language*, 32, 766-780. doi:10.1006/jmla.1993.1038
- Corral, S., Ferrero, M., & Goikoetxea, E. (2009). LEXIN: A lexical database from Spanish kindergarten and first-grade readers. *Behaviour Research Methods*, 41(4), 1009-1017. doi:10.3758/brm.41.4.1009
- Cuetos, F., Rodríguez, B., Ruano, E., & Arribas, D. (2007). *PROLEC-R: Battery of evaluation of reading processes for children from primary education reviewed*. Madrid: TEA.
- Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indexes in Spanish. *Behavior Research Methods*, 37, 665-671. doi:10.3758/bf03192738
- Defior, S., Fonseca, L., Gottheil, B., Aldrey, A., Rosa, G., Pujals, M., ... Serrano, F. D. (2006). *LEE. Test de lectura y escritura en español*. Buenos Aires: Paidós.
- Defior, S., Justicia, F., & Martos, F. (1996). The influence of lexical and sub lexical variables in normal and poor Spanish readers, *Reading and Writing*, 8, 487-497. doi:10.1007/bf00577024
- Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 257-287). Amsterdam: North-Holland.
- Forster, K. I., & Chambers, I. M. (1973). Lexical access and naming time. *Journal of Verbal Learning & Verbal Behavior*, 12, 627-635. doi:10.1016/s0022-5371(73)80042-8
- Foster, T. E., Ardoin, S. P., & Binder, K. S. (2018). Reliability and validity of eye movement measures of children's reading. *Reading Research Quarterly*, 53 (1), 71-89. doi:10.1002/rrq.182
- Förster, N., Kawohl, E., & Souvignier, E. (2018). Short-and long-term effects of assessment-based differentiated reading instruction in general education on reading fluency and reading comprehension. *Learning and Instruction*, 56, 98-109. doi.org/10.1016/j.learninstruc.2018.04.009
- Försters, N., & Souvignier, E. (2015). Effects of providing teachers with information about their students' reading progress. *School Psychology Review*, 44 (1), 60-75. doi:10.17105/spr44-1.60-75

- Goikoetxea, E., & Ferrero, M. (2019). Word recognition growth in Spanish children. Manuscript submitted for publication.
- González-Trujillo, M. C., Calet, N., Defior, S., & Gutiérrez-Palma, N. (2014) Escala de fluidez lectora en español: midiendo los componentes de la fluidez, *Estudios de Psicología*, 35 (1), 104-136. doi:10.1080/02109395.2014.893651
- Jenkins, J. R, Fuchs, L. S., van den Broek P, Espin C, & Deno S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, 95, 719–729. doi:10.1037/0022-0663.95.4.719
- Jiménez, J. E., Gove, A., Crouch, L., & Rodríguez, C. (2014). Internal structure and standardized scores of the Spanish adaptation of the EGRA (Early Grade Reading Assessment) for early reading assessment. *Psicothema*, 26(4), 531-537.
- Keenan, J. M., & Meenan, C. E. (2014). Test differences in diagnosing reading comprehension deficits. *Journal of Learning Disabilities*, 47 (2), 125-135. doi:10.1177/0022219412439326
- Kim, Y.S., Petscher, Y., Schatschneider, C., & Foorman, B. (2010). Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? *Journal of Educational Psychology*, 102, 652–667. doi:10.1037/a0019643
- Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, 100 (2), 310-321. doi:10.1037/0022-0663.100.2.310
- Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *The Quarterly Journal of Experimental Psychology*, 66 (3), 563–580. doi.org/10.1080/17470218.2012.658820
- Lacina, J., & Block, C. (2011). What matters most in distinguished literacy teacher education programs? *Journal of Literacy Research*, 43(4), 319-351. doi:10.1177/1086296x11422033
- Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology*, 100(1), 150-161. doi:10.1037/0022-0663.100.1.150
- Marín, J., & Carrillo, M. S. (1999). *Test Colectivo de Eficacia Lectora (TECLE)*. Unpublished manuscript, Universidad de Murcia.

- Moret-Tatay, C., & Perea, M. (2011). Is the go/no-go lexical decision task preferable to the yes/no task with developing readers. *Journal of Experimental Child Psychology, 110* (1), 125-132. doi:10.1016/j.jecp.2011.04.005
- Perea, M., & Rosa, E. (2000). The effects of orthographic neighborhood in reading and laboratory word identification tasks: A review. *Psicológica, 21*, 327-340.
- Perfetti, C. & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*(1), 22-37. doi:10.1080/10888438.2013.827687
- Rasinski, T. V., Reutzel, C. R., Chard, D. & Linan-Thompson, S. (2011). Reading fluency. In M. L. Kamil, P. D. Pearson, B. Moje, & P. Afflerbach E. (Eds), *Handbook of Reading Research, Volume IV* (pp. 286-319). New York: Routledge.
- Rayner, K., & Reichle, E. D. (2010). Models of the reading process. *Wiley Interdisciplinary Reviews: Cognitive Science, 1* (6), 787-799. doi:10.1002/wcs.68
- Rayner, K., Sereno, S., Morris, R., Schmauder, R., & Clifton, C., Jr. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4*, 21-49. doi:10.1080/01690968908406362
- Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: word frequency effects and individual differences. *Memory & Cognition, 26* (6), 1270-1281. doi:10.3758/bf03201199
- Sebastián-Gallés, N., Martí, M. A., Carreiras, M. F., & Cuetos, F. (2000). *LEXESP: Léxico informatizado del español* [LEXESP: A computerized word-pool in Spanish]. Barcelona: Universitat de Barcelona.
- Seymour, P. H., Aro, M., Erskine, J. M., & COST Action A8 (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143-174. doi:10.1348/000712603321661859
- Simões, E., & Alves-Martins, M. (2018). Reading acquisition in beginner readers: Typical errors in European Portuguese. *Educação e Pesquisa, 44*, e165734. Retrieved from <http://dx.doi.org/10.1590/S1678-4634201844165734>
- Valle-Arroyo, F. (1989). Reading errors in Spanish. In P. G. Aaron & R. M. Joshi (Eds.), *Reading and writing disorders in different orthographic systems* (pp. 163-175). The Netherlands: Kluwer Academic.
- Van Bon, W. H. J. (2007). *De Doorstreepleestoets* [Paper-and-pen lexical decision task]. Leiden, The Netherlands: PITS.
- Van Bon, W. H. J., Hoevenaars, L. T. M., & Jongeneelen, J. J. (2004). Using paper-and-pencil lexical-decision tests to assess word decoding skills: Aspects of validity and re-

liability. *Journal of Research in Reading*, 27 (1), 58–68. doi:10.1111/j.1467-9817.2004.00214.x

Van Bon, W. H. J., & Libert, J. E. A. (1997). Oral reading and silent reading compared: Evidence for a subtype of poor readers. *Polish Psychological Bulletin*, 28, 59-70.

Van Bon, W. H., Tooren, P. H., & van Eekelen, K. W. (2000). Lexical decision and oral reading by poor and normal readers. *European Journal of Psychology of Education*, 15(3), 259-270. doi:10.1007/bf03173178

Wechsler, D. (2005). *Escala de inteligencia de Wechsler para niños IV*. Madrid: TEA.

Wilson, J. M. G., & Jungner, G. (1968). *Principles and practice of screening for diseases*. Geneva: World Health Organization.

Apéndice A

LEO-1-min

TEST LEO-1-min

Nombre:

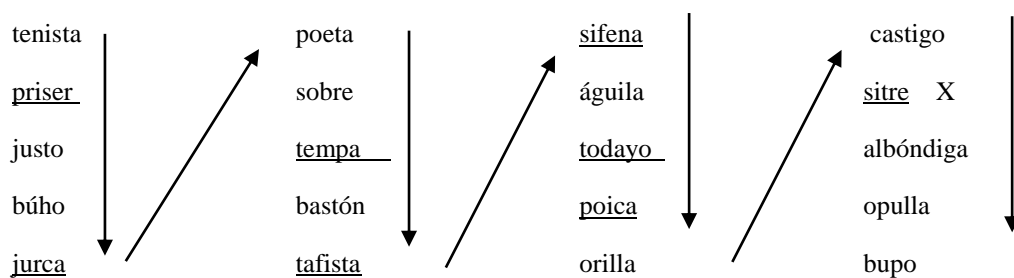
Edad:

Curso:

INSTRUCCIONES

1. Vas a encontrar algunas palabras escritas en español en varias columnas.
2. Algunas de esas palabras realmente existen, y otras son inventadas.
3. Subraya o haz una línea debajo de las palabras que no existen.
4. Comienza con la columna de la izquierda de arriba a abajo.
5. Si terminas la primera columna, pasa a la siguiente de la derecha. Sigue siempre este orden.
6. Cuando se acabe el tiempo, oirás: "PARAR". En ese momento, deja de leer y pon una X a la derecha de la última palabra, real o inventada, que hayas leído.

Example:



Apéndice B

LEO-1-min Form A

amigo	tarea	cabeza	sopefad
pretio	sur	papel	doctor
fuego	blanco	atención	piso
ayuta	literatura	gato	daño
pol	piano	par	barfe
imagen	río	encuentro	valor
morenso	presente	contenido	objeto
conjunto	carne	reamión	codi
lista	vicina	capital	loño
flus	patio	relisma	ayer
media	resultado	vino	francés
sala	cine	título	motor
rul	kilómetros	signo	alcohol
payamo	semana	borde	tipo
vía	secunto	unidad	prejitente
oportunidad	díe	letra	plan
red	abuelo	teléfono	caja
cazano	preufa	cola	ciendia
resto	cuarenta	caso	sombra
radio	sueño	imborcancia	lado
trel	época	hermano	gloria
casirad	serie	humor	rojo
habitación	catanidad	color	estado

Apéndice B (Continuación)

hambre	madera	curso	cuarlo
pena	rato	aspecto	dueño
espajo	esbrimor	pelo	cocina
esfuerzo	inteligencia	baru	café
dama	seguridad	ripo	centro
frase	nada	pie	intención
televisión	cantidad	policía	plaro
planeta	programa	nombre	pedra
cultura	peligro	otode	sitio
bope	impresión	sentimiento	jardín
ojo	bar	tierra	nitel
gente	mado	canal	guena
escena	ciepa	luba	alión
vaso	producto	pleno	favor
hujo	revolución	vida	fatibia
deseo	poatía	trabajo	palabra
quinle	energía	máquina	medida
estación	mundo	llegada	fiesca
renueldo	decisión	lunes	problema
frecuencia	puerta	carta	
cona	comida	forsagión	
cura	cacilán	mayoría	
espectáculo	salud	térjilo	

Respuesta

Item	Pseudopalabra	Item subrayado	No acción
		Palabra	Acierto
		Falsa alarma	Omisión correcta

Figure 1. *Matriz estímulo-respuesta*

Apéndice C

Tabla 5. Normas Provisionales de la Forma A del LEO-1-min en Centiles por Curso

Centiles	Curso					
	1 <i>n</i> = 48	2 <i>n</i> = 47	3 <i>n</i> = 51	4 <i>n</i> = 42	5 <i>n</i> = 50	6 <i>n</i> = 46
95	41	48	62	76	91	97
90	31	41	55	63	78	85
80	21	36	52	54	72	77
70	16	34	51	53	67	69
60	14	30	45	50	64	66
50	13	26	44	48	60	65
40	11	24	36	45	56	55
30	10	21	32	42	53	53
20	6	19	31	39	48	49
10	3	16	21	34	40	40
5	2	15	20	26	33	36
<i>Media</i>	15.17	28.74	41.53	48.17	61.62	63.24
<i>(DT)</i>	(10.75)	(11.59)	(13.09)	(11.71)	(17.55)	(15.71)

Recibido: 08-10-2018
Aceptado: 04-04-2019