

PROPIEDADES PSICOMÉTRICAS DE UNA PRUEBA EXPERIMENTAL PARA LA EVALUACIÓN DEL NIVEL ORTOGRÁFICO ¹

PSYCHOMETRIC PROPERTIES OF AN EXPERIMENTAL TEST FOR THE ASSESSMENT OF THE ORTHOGRAPHICAL LEVEL

Daniel Costa Ball

Universidad Católica del Uruguay, Uruguay

Ana Laura Palombo

Universidad Católica del Uruguay, Uruguay

Ariel Cuadro

Universidad Católica del Uruguay, Uruguay

Resumen: Se presenta el desarrollo de una prueba para la evaluación del nivel ortográfico de escolares (Test de Eficacia Ortográfica: TEO), con el objetivo de contribuir a las prácticas educativas, así como al diagnóstico de una competencia que impacta fuertemente en la escritura. A partir de una muestra de 693 sujetos de 2° a 6° de primaria de cuatro colegios de Montevideo se analizaron las propiedades psicométricas de los ítems que componen la prueba. Con una segunda muestra de 123 sujetos de un único colegio de Montevideo se estudió la fiabilidad de la prueba con el método de test-retest. Los estudios de validez de contenido, análisis de ítem, dimensionalidad y fiabilidad, han aportado evidencia de la validez de una estructura unidimensional; acorde con los planteamientos teóricos. En conclusión, se puede afirmar que esta prueba de nivel ortográfico (TEO) constituye un instrumento de medida válido y fiable que permite explorar a partir de un único factor el nivel ortográfico.

Palabras clave: *Psicometría, ortografía, escritura.*

Abstract: This article presents the development of the “Test de Eficacia Ortográfica (TEO)” for the assessment of orthographical level of elementary school students, and aims to contribute to educational practices and diagnosis of a competence strongly impacting writing skills. The psychometric properties of the measure’s items were analyzed based on a sample of 693 participants attending second to sixth grade at four private elementary schools in Montevideo. Test-retest Reliability of the instrument was measured based on 123 of the participants. The study of content validity, item analysis, dimensionality and reliability provides evidence of a unidimensional structure, in line with theoretical framework. In conclusion, our results suggest the appropriateness of the TEO as a valid and reliable single-factor measure for the assessment of orthographical level.

Keywords: *Psychometrics, orthography, writing.*

¹ (N. de la Ed.). A continuación se transcriben algunas consideraciones realizadas por la coautora de los estudios de nivel ortográfico en escolares uruguayos -publicado en 1978 - y que se toma como referencia en esta investigación para la construcción del TEO:

“Este trabajo denota un uso de recursos que enriquecen los análisis y contiene un planteamiento actualizado de los requerimientos básicos para este tipo de desarrollos instrumentales, que atienden un área compleja y relevante.

Al mismo tiempo y digno de ser resaltado, supone un reconocimiento a la obra de una investigadora y docente como la Prof. M. A. Carbonell de Grompone, cuyos aportes enriquecieron los albores de la Psicología Educativa en el Uruguay”. Prof. Emérita Elida Tuana, 2011.

INTRODUCCIÓN

El tema de la ortografía, en particular por las dificultades que presentan los niños en su adquisición, es un asunto recurrente en la agenda educativa y constituye un tema de preocupación creciente.

La ortografía es una actividad que tiene características propias, que impacta fuertemente en la realización de la escritura, tanto desde el punto de vista conceptual como formal y que el adecuado manejo que los sujetos posean de esta actividad tiene impactos positivos desde el punto de vista personal, social y profesional.

La adquisición de la ortografía se da con el desarrollo de la escritura y constituye el conjunto de normas que la regula; siendo un componente central de ésta, en tanto permite asegurar la correcta comunicación escrita entre los hablantes de una misma lengua (Real Academia Española, 2011).

La ortografía se introduce como una variable que es posible determinar, en la medida que el sujeto se ha podido apropiarse de las reglas de conversión grafema - fonema, y trascendido este momento, es capaz de considerar las peculiaridades ortográficas de las palabras almacenándolas en su léxico ortográfico.

En la lectura, la investigación cognitiva, ha señalado a lo largo de las últimas décadas la importancia de reconocer las palabras escritas en forma exacta y automática, para acceder así a la activación de un número importante de procesos mentales superiores que le hagan posible la construcción de significados.

El acceso a la palabra escrita, en forma automática, requiere el dominio y automatización de las reglas de conversión grafema-fonema; mecanismo por el cual se atribuyen secuencias fonológicas a secuencias ortográficas que se ensamblan en una palabra (ensamblaje fonológico). En el aprendizaje de la lectura, este último mecanismo va permitiendo desarrollar progresivamente un léxico ortográfico que almacena representaciones de palabras escritas, las que luego podrán ser identificadas en forma directa y a un bajo costo cognitivo (Cuadro, 2010). Precisamente el modelo de autoaprendizaje de Share (1995, 1999) ha

defendido la idea de que la decodificación permite la adquisición de las representaciones ortográficas necesarias para un reconocimiento visual de la palabra, rápido y autónomo.

Algo similar ocurre en la escritura, ya que estos mecanismos de acceso se reconocen también en este proceso. En lo que refiere al procedimiento fonológico se plantea que la escritura consiste en analizar y segmentar la representación fonológica de la palabra que se quiere escribir, y convertir los segmentos fonológicos que se derivan de esa representación en su equivalente ortográfico. Para el procedimiento léxico, se plantea que es posible escribir una palabra partiendo directamente de la representación ortográfica sin realizar ningún análisis intermedio. Esto supone disponer de una representación definida de la ortografía de la palabra que se impone en la escritura (Sanchez & Cuetos, 1998). El léxico ortográfico implica entonces, que cada palabra almacenada o más precisamente cada representación de la misma, tenga un nivel alto de activación, para que su recuperación, al momento de escribirla, sea automática (García & González, 2000).

Jiménez-Fenández, Defior & Serrano (2010) agregan que la influencia del conocimiento léxico en la escritura se ha puesto de manifiesto, fundamentalmente, a través de dos efectos: frecuencia y lexicalidad. El efecto de frecuencia según estas autoras, está vinculado a que la escritura de una palabra de alta frecuencia se escribe mejor que otra de baja frecuencia. El efecto de lexicalidad estaría sustentando que la escritura de palabras es mejor que la escritura de pseudopalabras. Ambos efectos estarían sustentados por la intervención del conocimiento léxico ortográfico que tenga el sujeto.

En este sentido Jiménez et al (2008) estudiaron la evolución de la adquisición de palabras con ortografía arbitraria en alumnos de educación primaria en España. Encontraron que hasta 3er año, los alumnos utilizan en forma predominante el procedimiento fonológico para escribir y que es recién a partir de 4to año que comienzan a usar el conocimiento ortográfico.

Sánchez, Diuk, Borzone & Ferroni (2009) estudiaron la interacción entre el conocimiento fonológico y el ortográfico, analizando las es-

trategias que utilizan los niños de 1er y 2do año escolar en la escritura de palabras. Las conclusiones que obtienen es que los mecanismos léxicos son relativamente más tardíos y se adquieren en forma gradual. Agregan que ni el conocimiento de las correspondencias ni los encuentros casuales con palabras de ortografía irregular, parecen suficientes para garantizar un desarrollo del conocimiento léxico-ortográfico entendido en sentido estricto.

Perfetti (2010) por su parte, en su modelo de lectura DVC (Decodificación, Vocabulario, Comprensión) hace mención a estos aspectos en el apartado referido a la relación vocabulario – comprensión, donde plantea que la decodificación afecta el vocabulario directamente, porque la decodificación exitosa (1) recupera significados de palabras conocidas y por ello refuerza las conexiones forma-significado y (2) establece conexiones dependientes del contexto entre palabras desconocidas y contextos que contengan significado. El vocabulario (conocimiento del significado de una palabra) afecta la decodificación, porque el decodificar una palabra cuyo significado es conocido refuerza la conexión entre la forma ortográfica de esa palabra (la ortografía) y su significado. El proceso ayuda a establecer una representación específica de esa palabra, que especialmente ayuda en el caso de palabras con mapeos de ortografía y pronunciación excepcionales o irregulares y teóricamente ayuda para todas las palabras. Cuando los niños decodifican palabras, refuerzan el conocimiento del vocabulario, y cuando los niños recuperan el conocimiento que tienen del significado de una palabra mientras la decodifican, refuerzan la capacidad de identificar esa palabra.

Dada la inexistencia, en nuestro medio, de instrumentos válidos y confiables que tengan como objetivo central evaluar el desarrollo del léxico ortográfico en los niños de edad escolar, nos propusimos elaborar una prueba, que llamamos Test de Eficacia Ortográfica (TEO); que respondiera a las premisas actuales de la psicolingüística cognitiva y que permitiera ir construyendo un perfil ortográfico de los alumnos, así como, analizar su evolución a lo largo de la escolaridad.

El objetivo de este trabajo es presentar el desarrollo y las características psicométricas

de la prueba (TEO), previo a iniciar el proceso de baremación de la misma. La prueba fue diseñada para ser un instrumento sencillo y de fácil aplicación. Puede ser aplicada en forma individual o colectiva en escuelas primarias o contextos clínicos.

1. MÉTODO

1.1. Participantes

El estudio de dimensionalidad y fiabilidad se realizó con dos muestras. La primera quedó compuesta por 693 alumnos pertenecientes a cuatro colegios privados de Montevideo, emplazados en estratos socioeconómicos medio y medio-alto (ver Tabla 1). La versión experimental del test de evaluación del nivel ortográfico (TEO) fue aplicada a alumnos de segundo a sexto año de primaria y fueron evaluados durante el mes de septiembre de 2010.

Para el estudio de fiabilidad test-retest, se trabajó con una segunda muestra a conveniencia de un colegio de la ciudad de Montevideo y se eligió al azar un grupo por año, evaluando en total a 122 alumnos de segundo a sexto año de primaria.

Tabla 1.
Descripción de la muestra por curso

Curso	Primera fase	Segunda fase
	Estudio de dimensionalidad	Estudio de fiabilidad
2º	126	22
3º	138	27
4º	142	24
5º	137	23
6º	150	26
Total	693	122

La elección de este bloque poblacional se basa en que los sujetos ya se iniciaron en la lectura y escritura, y al mismo tiempo, transcurrida la escolaridad van aumentando su dominio de la ortografía, fundamentalmente en el manejo de las palabras de uso frecuente, por la reiterada exposición y uso de las mismas.

1.2. Instrumento

Se desarrolló el *Test de Evaluación del nivel Ortográfico (TEO)*, siendo el mismo un test de velocidad; con el fin de evaluar el léxico ortográfico de los alumnos de 2do a 6to de nivel primario (ver Tabla 2). La prueba consta de 99 ítems formados por una palabra y su homófono (escrito con el grafema incorrecto), donde el alumno debe elegir la palabra bien escrita, disponiendo para ello de un tiempo de tres minutos (ver Tabla 3).

Para la elaboración de esta prueba se realizó una revisión de la literatura y se tomaron como base referencial los estudios de Carbonell & Tuana (1978), quienes seleccionaron una lista de palabras de uso frecuente para los escolares de nuestro país a partir de una lista elaborada por García-Hoz (citado en Carbonell & Tuana, 1978, p. 245) para el español.

La lista de Carbonell & Tuana (1978) está compuesta por vocablos distribuidos en centenas, de tal forma que las palabras de la centena 1 son las más frecuentes y continúan en orden decreciente de frecuencia. De esta lista general, estas mismas autoras, elaboran otra lista de palabras que contienen compromiso ortográfico. De todas las palabras que contiene la lista planteada, se toman en cuenta para la realización de esta prueba, solamente aquellas palabras que contienen la alternancia del grafema *b* y *v* para representar el fonema /b/, los grafemas *j* y *g* para la representación del sonido /j/, las palabras que pueden escribirse con *ll* o *y*, las palabras que contienen el grafema para el que no hay fonema: *h*, y los grafemas que representan el sonido /s/.

La lista elaborada por Carbonell & Tuana (1978) es más abarcativa. Para la elaboración de la prueba, se tomaron en cuenta estrictamente las palabras con alternancia de grafemas para el mismo fonema, de tal forma que no se consideraron las palabras con el grafema *x*, dado que a este grafema no le corresponde ningún sonido de la lengua, sino que principalmente representa la secuencia /k+s/ (Real Academia Española, 2011, pp.156). Se han eliminado también las palabras con el grafema *z*, dado que en nuestra lengua la diferenciación entre el sonido /z/ y /s/ no se realiza. En este sentido, en la mayoría de los hablantes hispanos, se ha

reducido al fonema /s/ dando lugar al seseo. (Real Academia Española, 2011, p.56) Dada la frecuencia de aparición se seleccionaron las palabras que contienen *ce-ci*, y no se agrega otra dificultad ortográfica para el mismo sonido. Carbonell & Tuana (1978, p.249) plantean que una de las dificultades ortográficas que se visualiza con mayor frecuencia son las sílabas "ce" y "ci" ("c" o "s"), las sílabas con "b" o "v".

De esta forma la lista de palabras queda determinada por 425 palabras, que se distribuyen de la siguiente manera según el porcentaje que presenta cada dificultad:

Tabla 2.
Distribución de las palabras según el porcentaje de dificultad en la lista original

Dominio	Porcentaje
V	20,5
B	26,8
J	0,9
G	4,0
C	33,2
H	7,8
Y	4,0
LL	2,8
Total	100

La primera versión de la prueba fue sometida a un proceso de depuración a través del *juicio de expertos*, donde cuatro docentes por curso (20 en total) analizaron la adecuación de los ítems con la dimensión que evaluaban mediante una escala Likert de cuatro puntos. Se utilizó el acuerdo inter-jueces para eliminar los ítems más problemáticos.

Esta decisión se sustentó en el hecho de que la lista tenía aproximadamente 30 años, por lo que era necesario confirmar que las palabras continuaban presentándose como frecuentes. A esta lista de palabras de Carbonell & Tuana (1978) se le agregaron también las palabras que componen la prueba de Carrillo (2010), de más reciente elaboración y que incluyen grafemas *b/v*, *g/j* seguidos de "e-i", tanto en posición inicial como intermedia y palabras con "h" o vocal en posición inicial. En esta última prueba se le pide

al sujeto que decida entre dos opciones (palabra y pseudohomófono) la palabra correctamente escrita. Costa de 61 palabras de alta frecuencia para los escolares españoles. Esta lista no posee todos los grafonemas inconsistentes que establecimos como necesarios para nuestro medio, como los grafemas “c” para el sonido /s/, que resulta significativos para población uruguaya.

Los criterios para la elección de los jueces fueron: tener un mínimo de dos años de permanencia en el grado; tener como mínimo cinco años de haber conseguido su titulación de maestro y tener experiencia de trabajar con niños de nivel socio económico cultural medio de Montevideo. Cada juez, recibió el listado de las palabras seleccionadas según los criterios antes mencionados, el cual estaba compuesto por 231 palabras. Cada juez debía ubicar las palabras en una de las siguientes categorías: Nada frecuente, Poco frecuente, Frecuente y Muy frecuente. Por último, de las 231 palabras estudiadas se seleccionaron 99 para conformar el TEO (ver Tabla 3).

Prueba de Eficacia Lectora (TECLE) de Cuadro, Costa, Trias & Ponce de León (2009): Esta prueba evalúa la habilidad lectora y está conformada por 64 ítems (frases incompletas). Es un test de velocidad con un tiempo límite de 5 minutos. El alumno debe responder a cada ítem leyendo cada frase incompleta presentada y completar la misma entre cuatro opciones que se le presenta (tres incorrectas y una correcta). Los estudios psicométricos evidencian que es una prueba unidimensional, fiable (fiabilidad test-retest de .88) y válida. Los estudios psicométricos evidencian que es una prueba unidimensional, fiable (fiabilidad test-retest de .88) y válida. Los

estudios de análisis de funcionamiento diferencial de los ítems (FID) arrojó resultados de al menos un 6% de los ítems con sospecha leve de FID (Costa Ball, Grundel & Cuadro, 2011).

1.3. Procedimiento

Primero se recogió el consentimiento de los cuatros centros para participar en la investigación, la escala en su versión inicial (experimental) fue administrada en horario escolar por una psicopedagoga y en coordinación con las maestras de las mismas instituciones.

A los alumnos se les señaló que la participación era voluntaria y que el cuestionario debía completarse de forma individual. A su vez se les informó que en todo momento se mantendría la confidencialidad y el anonimato de los datos.

Una vez repartidos los protocolos se procedió a una lectura conjunta de la consigna, la cual está especificada en la primera hoja del cuadernillo de trabajo. al mismo tiempo se realizan 3 ítems de ejemplo en forma colectiva, explicando en cada caso cuál es la respuesta correcta. Una vez finalizado este proceso se les indicó que comiencen a trabajar. A partir de ese momento se contabilizaron 3 minutos, una vez llegados a este tiempo se detiene el cronómetro. Se les solicitó que detengan el trabajo y tracen una línea en el último ítem marcado. Una vez que todos realizaron esta línea se les pidió que continuaran realizando el test y se volvió a iniciar el cronómetro. A los 5 minutos (es decir 2 minutos después del corte anterior) se les volvió a pedir que tracen una línea y a partir de allí no se contabilizó más el tiempo. Se les indicó a los alumnos que finalicen la prueba.

Tabla 3.
Palabras que conforman la prueba TEO

Natación	Vestido	Cenar	Conocer	Sombra	Varios	Bosque	Imaginar	Mover
Invitar	Bañar	Resolver	Abrazar	Noticia	Precioso	Amanecer	Humo	Favorito
Colegio	Decir	Verdura	Voz	Bien	Emoción	Cambiar	Humano	Hallar
Trabajo	Llegar	Edificio	Yo	Crecer	helado	Abril	Blanco	También
Selva	Situación	Educación	Levantar	Gracias	Bizcocho	Nueve	Agradecer	Energía
Entonces	Vacaciones	Primavera	Ventana	Visitar	Mayor	Árbol	Ejercicio	Llover
Problema	Felicitar	Haber	Fábrica	Volver	Quince	Abeja	Cine	Amable
Sábado	Conocimiento	Arriba	Playa	Hombre	Aparecer	Bueno	Canción	vida
Fácil	Página	Ejemplo	nube	Libre	Solución	Vaso	Ciudad	Abierto
Verano	Libro	Silencio	Ahora	nombre	Hojas	Once	Atención	Ayer
Febrero	Barrio	Dirección	Aburrido	Vecino	Almacén	Receta	Cocina	llorar

De esta forma, al terminar la prueba, el protocolo presenta 2 marcas: a los 3 minutos y a los 5 minutos de iniciada. También se tiene el trabajo completo en su totalidad, más allá del tiempo que le insuma al estudiante.

1.4. Análisis de datos

Con el objetivo de reportar la unidimensionalidad de la prueba y el estudio de fiabilidad, se trabajó con dos muestras. Con la primera (693 alumnos) se realizaron los análisis dirigidos al estudio métrico de los ítems, validez convergente y la evaluación dimensional de la prueba sin tiempo límite. Para la consecución de este objetivo se utilizó el programa estadístico NOHARM (Fraser & McDonald, 1988) junto con el paquete estadístico SPSS 18.0. Con la segunda muestra (122 alumnos) los análisis están dirigidos a la obtención de los estudios de fiabilidad con el método test-retest utilizando el paquete estadístico SPSS 18.0.

2. RESULTADOS

2.1. Análisis estadístico de los ítems en la aplicación sin tiempo

El análisis de la calidad métrica del test es una fase importante en el desarrollo de instrumentos de medida. Este estudio, se realiza luego de los análisis cualitativos de los ítems,

análisis de validez de contenido, de constructo, evaluación de jueces, etc. (Carretero-Dios & Pérez, 2005). Los resultados muestran (ver Figura 1) que todos los índices de dificultad en los 99 ítems son altos o próximos a uno, resultado de que la mayoría de los alumnos resolvieron bien los ítems en situación sin tiempo límite. Cuando los valores son bajos representan ítems muy difíciles (la mayoría de los sujetos han fracasado en su resolución).

Al tratarse el TEO de un test de velocidad, se entiende que estos test deben estar compuestos por ítems fáciles de tal forma que las diferencias de las puntuaciones en los sujetos se deba a la velocidad de ejecución utilizada en cada ítem. Al poner un tiempo límite, no todos los sujetos llegan a contestar la misma cantidad de ítems en el test y en el mismo plazo de tiempo (Martínez, Hernández & Hernández, 2006). En definitiva, para que el test sea de velocidad, todos los ítems deben ser fáciles por lo tanto el índice de dificultad podríamos denominarlo índice de facilidad. En nuestro estudio, el índice se calculó sin la consideración del tiempo, dejando que los alumnos completen todos los ítems. A los ítems no respondidos se los considera como ítems contestados incorrectamente.

Como se puede apreciar en la Figura 1, únicamente tres ítems presentan un índice de dificultad muy bajo, se ubican dos desvíos estándar (*DE*) por debajo de la media, evidenciando que

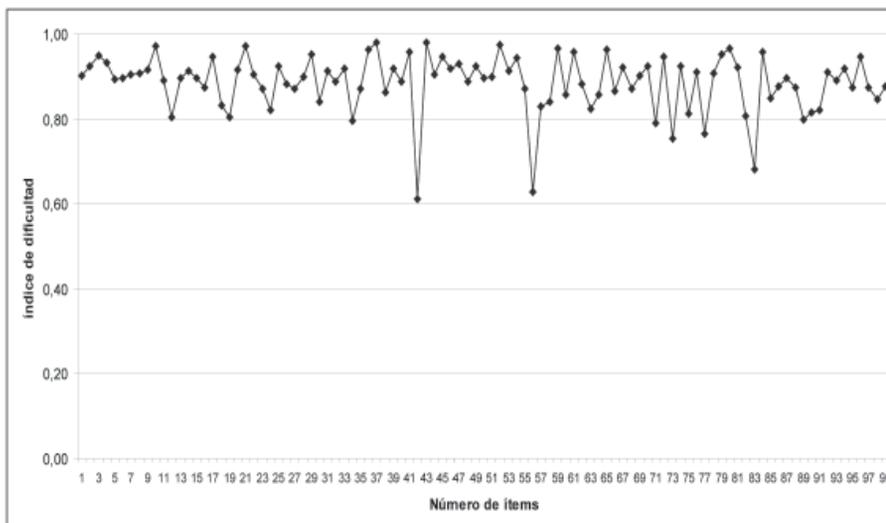


Figura 1. Índice de dificultad por ítem.

en estos ítems los sujetos fracasan más que en el resto y esto no colabora para ser utilizados en una prueba de velocidad, por lo tanto son candidatos a ser eliminados o ponerlos últimos en el protocolo. El promedio de los 99 ítems fue de $M = 0.88$ con un $DE = 0.067$.

En la Tabla 4, se muestran la cantidad de sujetos, medias, mediana, la desviación típica, asimetría, y la prueba de Kolmogorov-Smirnov para una muestra. Esta última es una prueba de bondad de ajuste que sirve para contrastar la hipótesis nula de que la distribución del puntaje total en el TEO sin tiempo se ajusta a una distribución normal. En consonancia con los análisis de ítems, la asimetría es negativa en todos los cursos, junto a los resultados del estudio de normalidad, se debe rechazar la hipótesis de normalidad en todos los cursos y concluir que las puntuaciones en el TEO sin tiempo, no se ajustan a una distribución normal. Esto último, se alinea con la definición que dimos previamente sobre las características de la prueba

de tipo de velocidad. Por este motivo sería adecuado usar pruebas no paramétricas para comprobar la existencia de diferencias entre los grupos (*Anova de Kruskal-Wallis*).

Para comprobar si existen diferencias significativas en el puntaje total sin tiempo en el TEO entre la variable curso, se empleó el *Anova de Kruskal-Wallis*. Se encontraron diferencias significativas ($p < ,000$) en el puntaje total de la prueba. El estadístico de *Kruskal-Wallis (Chi-cuadrado)* tiene un valor de 184,66 con 4 grados de libertad y con un nivel crítico de 0,000. Por lo tanto, se puede rechazar la hipótesis de igualdad de promedio poblacional y concluir que los grupos comparados difieren en el puntaje total en la prueba TEO sin tiempo.

Una vez detectado que existen diferencias significativas entre la variable curso, utilizamos la prueba de *U de Mann-Whitney* (ver Tabla 5) para dos muestras independientes, acompañada de la corrección de Bonferroni para controlar la tasa de error de tipo I (dividimos el

Tabla 4.
Estadísticos descriptivos de la prueba TEO sin tiempo

	N	Media	Mediana	DE	Asimetría	Normalidad
2º	115	73,63	77,0	15,57	-1,26	1,42*
3º	133	83,62	90,00	14,71	-1,69	2,06*
4º	140	90,33	93,0	9,49	-2,07	2,37*
5º	134	92,75	95,0	5,68	-1,75	1,98*
6º	136	94,10	96,0	5,68	-2,10	2,26*
Total	658	87,33	92,0	13,58	-2,15	5,00*

* $p < .05$ (Estudio de normalidad, Prueba de Kolmogorov-Smirnov para una muestra)

Tabla 5.
Prueba de *U de Mann-Whitney* de comparación del Puntaje TEO sin tiempo entre la variable curso

	TEO sin tiempo			
	2º - 3º	3º - 4º	4º - 5º	5º - 6º
Mann-Whitney U	4761,00	6703,50	8086,50	7792,00
Wilcoxon W	11431,00	15614,50	17956,50	16837,00
Z	-5,12	-4,00	-1,97	-2,06
Asymp. Sig (2-tailed)	,000*	,000*	,048	,039

* Diferencias significativas ($p < .012$) nivel crítico corregido por Bonferroni

nivel de significación .05 entre el número total de comparaciones por pares que vamos a comparar, en nuestro caso son 4, segundo versus tercero, tercero versus cuarto, cuarto versus quinto versus sexto, por lo tanto el nivel crítico corregido por Bonferroni es $.05/4 = .012$ (Pardo & Ruiz, 2002, p. 575).

De los 4 pares de comparaciones que tenemos, solamente en dos no existen diferencias significativas entre los resultados en el TEO sin tiempo. Por lo tanto, existen diferencias significativas ($p < .012$) en el TEO sin tiempo entre los puntajes de segundo versus tercero y tercero versus cuarto.

2.2. Análisis estadístico de los ítems en la aplicación con tiempo (3 minutos)

En la Tabla 6, se muestran la cantidad de sujetos, medias, Mediana, la desviación típica, asimetría, y la prueba de Kolmogorov-Smirnov para una muestra. En los cursos bajos (2º y 3º año) presentan un índice de simetría positiva evidenciando que la frecuencia absoluta es más alta a la izquierda de la media (puntajes bajos). Mientras que para los cursos de 4º, 5º y 6º año

el índice de asimetría es negativo, evidenciando que a la derecha de la media presentan una frecuencia más alta.

En relación al estudio de normalidad (ver Tabla 6), no se debe rechazar la hipótesis de normalidad en los cursos de segundo a sexto año y se puede concluir que las puntuaciones en el TEO, con límite de tiempo de tres minutos, se ajustan a una distribución normal, no siendo así en el puntaje de la muestra total. Esto último, se alinea con la definición que dimos previamente sobre las características de la prueba de tipo de velocidad, cuando los ítems son fáciles y no cuenta el tiempo. Por este motivo es adecuado usar pruebas paramétricas para comprobar la existencia de diferencias entre el puntaje en el TEO con límite de tiempo en la variable curso de tercero a sexto año (procedimiento ANOVA de un factor).

Los resultados obtenidos al comprobar si existen diferencias significativas en el puntaje total con tiempo (3 minutos) en el TEO, entre la variable curso, evidencian que, puesto que el valor del nivel crítico es menor a $p < .05$ (ver Tabla 7), debe rechazarse la hipótesis de igual-

Tabla 6.
Estadísticos descriptivos de la prueba TEO – 3 minutos

	N	Media	Mediana	DE	Asimetría	Normalidad
2º	115	18,53	16,00	10,70	0,61	1,34
3º	133	34,95	33,00	16,36	0,55	1,16
4º	140	49,29	51,00	19,89	-0,03	0,49
5º	134	63,03	65,00	18,02	-0,16	0,59
6º	136	72,59	75,00	19,82	-0,66	1,06
Total	669	48,63	48,00	25,67	0,14	1,67*

* $p < .05$ (Estudio de normalidad, Prueba de Kolmogorov-Smirnov para una muestra)

Tabla 7.
Análisis de varianza de un factor entre el TEO (sin tiempo) por curso

Fuente	gf	SS	MS	F
Entre grupos	4	234969,75	58742,43	193,56***
Intragrupos	653	198171,76	303,47	
Total	657	433141,52		

* $p < .000$

dad de medias. Pudiéndose concluir que las poblaciones definidas por la variable curso no poseen el mismo puntaje en el TEO con tiempo. En consecuencia hay que determinar entre qué grupos existen las diferencias.

Una vez detectado que existen diferencias significativas entre la variable curso, para conocer qué media difiere de qué otra utilizamos el contraste denominado comparaciones múltiples post hoc del método *Tukey* (ver Tabla 8). Al realizar las comparaciones múltiples con la prueba “*post hoc*” de *Tukey*, observamos (ver Figura 2) que existen diferencias significativas entre los grupos de segundo año y el resto de los grupos ($p < .000$); tercero difiere significativamente de cuarto, quinto y sexto año; cuarto año rinde significativamente mejor que quinto y sexto; quinto rinde significativamente mejor que sexto. En suma en cada año existen diferencias significativas con el grupo precedente y el siguiente.

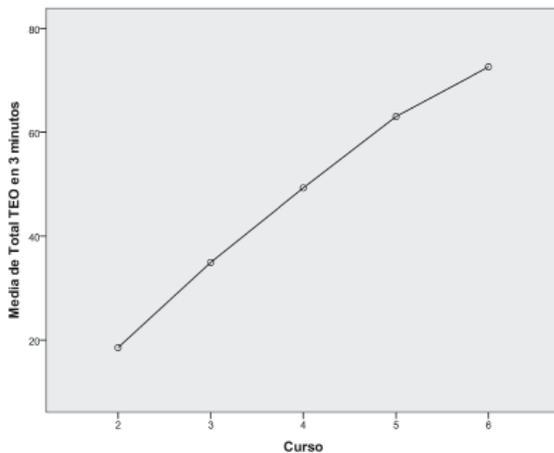


Figura 2. Media del TEO sin tiempo por curso

De todas las comparaciones posibles “*post hoc*” la Tabla 8 ofrece una clasificación basada en el grado de parecido existente entre sus medias, en el subconjunto 1 está incluidos los alumnos de segundo, en el subconjunto 2 los alumnos de tercer año, en el subconjunto 3 los alumnos de cuarto y en el subconjunto 4 los alumnos de quinto y el último subconjunto 5, esta conformado por los alumnos de sexto año.

Tabla 8. Clasificación de subconjuntos homogéneos

Curso	N	1	2	3	4	5
2º	115	18,53				
3º	133		34,95			
4º	140			49,29		
5º	134				60,03	
6º	136					72,59

2.3. Estudio para confirmar que es un test de velocidad

Para efectuar los cálculos de fiabilidad previamente se tiene que estudiar si se trata de una prueba de potencia o de velocidad (ver Tabla 9), esta distinción afecta los procedimientos de análisis de la fiabilidad (Martínez et al., 2006). Se utilizó el método propuesto por Stafford, el mismo, se calcula a través del cociente entre el número de ítems no alcanzados (NI) dividido la suma de los ítems omitidos (O), no acertados (E) y los ítems no alcanzados (NI). El indicador llamado índice de velocidad (IV) oscila entre cero y uno, si se aproxima a cero se trata de un test de poder, en cambio, si el resultado es uno, entonces el test es de velocidad (Lu & Sireci, 2007).

$$IV = \left(\frac{\sum NI}{\sum E + \sum O + \sum NI} \right) * 100$$

Tabla 9. Cálculo del índice de velocidad de Stafford (N = 667)

Curso	NI No alcanzados	E Errores	O Omisión	IV Índice de velocidad
2º	9036	646	1	0,93
3º	7895	623	0	0,92
4º	6437	522	1	0,92
5º	4390	430	0	0,91
6º	3114	391	1	0,89
Total	30872	2613	3	0,92

2.4. Estudio de la dimensionalidad del test

Una vez estudiada la validez de contenido, de evaluar cualitativamente los ítems, y de realizar el análisis métrico, el siguiente paso es la evaluación de la dimensionalidad. El TEO es un test de rendimiento, y los ítems se puntúan como acierto o fracaso, por lo tanto los ítems son dicotómicos. La literatura ha recogido muchas críticas cuando se corren los análisis factoriales clásicos en estas puntuaciones, encontraron que el AF con datos dicotómicos da como resultados factores artificiales o multidimensionales cuando en realidad es unidimensional. Martínez et al. (2006) aconseja en estos casos utilizar los métodos de Análisis Factorial no lineal de McDonald. En consecuencia, se evaluó la dimensionalidad a través del programa NOHARM de Fraser & McDonald (1988) el ajuste para un modelo unidimensional se lo valoró con el índice de Tanaka & Huba (1985), los índices por encima a 0,90 evidencian buen ajuste.

Para el análisis de la dimensionalidad se utilizó una muestra de 693 alumnos, compuesta por alumnos de segundo a sexto año de primaria de ambos sexo. Se consideró la prueba sin límite de tiempo, las omisiones se evaluaron como errores y no se tuvieron en cuenta los sujetos que presentaron ítems no alcanzados. Los resultados obtenidos fueron los siguientes. Para el modelo unidimensional, se obtiene un índice Tanaka bastante alto según la literatura > 0,95 evidenciado que todos los ítems del TEO miden una sola dimensión (habilidad de eficacia ortográfica). Se puede concluir que el TEO cumple con el supuesto de unidimensionalidad. Por lo tanto, podemos aceptar que el **modelo es unidimensional**.

Es importante considerar que las cargas factoriales son todas positivas y mayores a 0,30. Esto evidencia que todos los ítems colaboran a la medición de la habilidad en eficacia ortográfica.

2.5 Fiabilidad

Hay tres métodos clásicos para el estudio de la fiabilidad: test-retest; pruebas paralelas y los coeficientes de consistencia interna. De los tres métodos, el de pruebas paralelas es

imposible porque no tenemos dos versiones del instrumentos, el último (coeficiente de consistencia interna alfa de Cronbach) no es el indicado para datos dicotómicos y además no debe ser utilizado cuando el test en estudio es una prueba de velocidad (Martínez et al, 2006). En consecuencia, el método resultante y adecuado para usarse con datos dicotómicos y en un test de velocidad es el test – retest (Pardo, y Ruíz, 2002).

Este método del cálculo de la fiabilidad test-retest consiste en saber si el test mide el concepto teórico para el cual fue construido y a la vez, si las puntuaciones obtenidas a través del mismo son estables en el tiempo. En especial, interesa saber si el test presenta estabilidad en el tiempo, es decir, si se administra en dos ocasiones el test llega a resultados similares.

Se administró el mismo test en dos momentos distintos (primera administración en septiembre y la segunda en octubre) a una muestra de 137 sujetos. El resultado obtenido indica que presenta una correlación entre formas de .76.

2.6. Validez

La comprobación de la validez del test TEO ha sido realizada teniendo en cuenta la validez de contenido y la validez convergente. La escala resultó válida en cuanto a contenido puesto que los ítems fueron construidos teniendo en cuenta los estudios de Carbonell & Tuana (1978), quienes seleccionaron una lista de palabras de uso frecuente para los escolares de nuestro país a partir de una lista elaborada por García-Hoz (citado en Carbonell & Tuana, 1978) para el español, y los de Carrillo (2010). La misma fue evaluada y filtrada por jueces.

La validez convergente de un concepto valora el grado en el cual dos medidas del mismo constructo muestran correlaciones elevadas. El Test de Eficacia Lectora (TECLE) (Cuadro y col. 2009) evalúa la habilidad lectora a partir de los mecanismos de reconocimiento de palabras (procesamiento fonológico y ortográfico), lo que implica para el uso procesamiento ortográfico el desarrollo del léxico ortográfico. El análisis de convergencia entre el TEO sin tiempo y el Test de Eficacia Lectora (TECLE) (2009) ofrece coeficiente de

correlación altos ($r = .66^{**}$; $R^2 = .43$) y entre el TEO con 3 minutos y el TECLE ($r = .81^{**}$; $R^2 = .66$).

3. DISCUSIÓN

Los resultados obtenidos en este estudio nos han permitido confirmar que el Test de Evaluación del nivel Ortográfico (TEO) para alumnos de segundo a sexto año de nivel primario, presenta adecuadas propiedades psicométricas. Los análisis de validez de contenido, análisis de ítem, dimensionalidad y fiabilidad, han aportado evidencia de la validez de una estructura unidimensional, acorde con los planteamientos teóricos (Sanchez & Cuetos, 1998; Jiménez-Fernández et al, 2010; Perfetti, 2010).

En relación al comportamiento del TEO administrado sin límite de tiempo, los puntajes totales presentan índices de asimetría negativa en todos los cursos y no presentando una distribución normal de los puntajes. En cambio, cuando se introduce la variable tiempo, los resultados en cada año se normalizan. Lo anterior es coincidente con los resultados obtenidos a través del procedimiento ANOVA de un factor con el TEO con tiempo y con el procedimiento de la prueba H de Kruskal-Wallis con el TEO sin tiempo. Los resultados mejoran año a año, de segundo hasta tercero. Pero no se encuentran diferencias significativas entre cuarto, quinto y sexto año cuando se administra el TEO sin tiempo. Estos tres grupos rinden superiormente que los alumnos de segundo y tercer año, aunque entre cuarto y quinto no existen diferencias significativas y lo mismo entre quinto y sexto, pero sí existen diferencias significativas entre cuarto y sexto año. Lo que es consistente con la idea de que en los años superiores (cuarto, quinto y sexto año) hay un desarrollo significativo del léxico ortográfico en las palabras propuestas en la prueba. Al no disponer de tiempo resulta esperable a estos niveles que incluso los alumnos que no disponen de un léxico ortográfico **autónomo** puedan utilizar ciertas estrategias que no los hagan fracasar aunque les insuman mayor tiempo.

Con la introducción del tiempo se logra normalizar la variable de puntaje total del TEO, y se puede observar que los puntajes promedios

en cada grupo crecen año a año, conformándose así, cinco grupos bien diferenciados uno por año. Evidenciando que por curso, frente a la baja dificultad que presenta la prueba, los alumnos de un año rinden significativamente mejor que los alumnos del año anterior y peor que los alumnos de años superiores. Siendo esta una prueba de velocidad con ítems fáciles, entonces, el tiempo límite de tres minutos pone en evidencia el número de representaciones ortográficas disponible en forma **precisa y autónoma**. Marcando una tendencia de crecimiento año a año.

En conclusión, se puede afirmar que este Test de Evaluación del nivel Ortográfico (TEO) constituye un instrumento de medida válido y fiable que permite explorar a partir de un único factor la evaluación del nivel ortográfico. El instrumento puede resultar útil para los maestros ya que les permite, en forma rápida y económica, disponer de información del proceso de adquisición de la ortografía de sus estudiantes; así como para psicopedagogos o psicólogos educacionales al momento de evaluar los aprendizajes en lenguaje escrito de los alumnos que no alcanzan los niveles esperados. Esta prueba aportaría un insumo relevante en el estudio de posibles dificultades específicas.

Los datos obtenidos a nivel grupal informan acerca del funcionamiento de toda la clase en relación con esta habilidad y al mismo tiempo, la posibilidad de obtener un perfil ortográfico del grupo, con fines de intervención es un elemento significativo para el trabajo dentro del aula en relación con esta área de aprendizaje.

Entendemos que esta prueba aportará datos relevantes sobre un aspecto de la ortografía. No es el objetivo de la misma la evaluación completa de esta competencia y los datos que pueda aportar serán a modo de screening, por lo que necesariamente una puntuación afectada deberá complementarse con otras propuestas que aporten mayores datos sobre su competencia en el lenguaje escrito.

Como ya se ha señalado la ortografía constituye un componente sustancial de la escritura, y su correcta adquisición supone un objetivo prioritario de la educación primaria y media.

Para finalizar, es importante destacar que aunque en este trabajo se han aportado evidencias de validez, centradas en la unidimensionalidad del instrumento y validez convergente, queda para investigaciones futuras el estudio de la validez externa. También sería importante analizar la relación que el TEO guarda con otras variables del lenguaje escrito (nivel lector; dictado de palabras; comprensión lectora); con las cuales la psicolingüística cognitiva ha señalado relaciones significativas.

4. REFERENCIAS

- Carbonell, M.A. & Tuana, E. (2004). *La ortografía de uso. Un espacio de reflexión y desafío*. Compilación de escritos de la Prof. Emérita Ma. Angélica Carbonell de Grompone, (pp. 245 -280) (2° Ed.). Montevideo: Sociedad de Dislexia del Uruguay y Coeditores.
- Carretero-Dios, H. & Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales. *International Journal of Clinical and Health Psychology* 5 (3), 521-551.
- Carrillo, M. (2010). La evaluación de la dislexia evolutiva. En C. & Domínguez, *Dislexia y Sordera. Líneas actuales en el estudio de la lectura escrita y sus dificultades* (pp. 111-132). Málaga: Ediciones Aljibe.
- Costa Ball, D., Grundel, M. & Cuadro, A. (2011). Impacto y funcionamiento diferencial de los ítems respecto al género en el Test de Eficacia Lectora (TECLE). *Ciencias Psicológicas V (1)*, 47-57.
- Cuadro, A. (2010). *La lectura y sus dificultades: la dislexia evolutiva*. Montevideo: Magrú, Universidad Católica del Uruguay.
- Cuadro, A., Costa, D., Trias, D. & Ponce de Leon (2009). *Evaluación del nivel lector. Test de eficacia lectora (TECLE)* de J. Marín y M. Carrillo. Manual Técnico. Montevideo: Prensa Médica.
- Fraser & McDonald. (1988). NOHARM: Least squares item factor analysis. *Multivariate Behavioral Research*, 23, 267-269.
- García, J. & González, D. (2000). *Dificultades del Aprendizaje e intervención psicopedagógica. Volumen II. Lectura y Escritura*. Madrid: Eos.
- Jiménez, J., O'Shanahan, M., Tabraue, M., Artilles, C., Guzman, R., Naranjo, F., Muñetón, M. & Rojas, E. (2008). Evolución de la escritura de palabras de ortografía en lengua española. *Psicothema*, 20 (4), 786-794.
- Jiménez-Fenández, G., Defior, S. & Serrano, F. (2010). Adquisición de la escritura de palabras: modelos y metodología de estudio. En Carrillo, M. & Domínguez, A. B. *Dislexia y Sordera. Líneas actuales en el estudio de la lengua escrita y sus dificultades* (pp. 75-97). Málaga, España: Aljibe.
- Lu, Y. & Sireci, S. (2007). Validity Issues in Test Speededness. *Educational Measurement: Issues and Practice*, 26, 29-37.
- Martínez, M.R., Hernández, M.J. & Hernández, M.V. (2006). *Psicometría*. Madrid: Alianza Editorial.
- Pardo, A. & Ruiz, M. (2002). *SPSS 11. Guía para el análisis de datos*. Madrid: McGraw-Hill.
- Perfetti, C. (2010). Decoding, Vocabulary, and Comprehension. The golden Triangle of Reading Skill. En McKeown, M. & Kucan, L. *Bringing reading research to life* (pp. 291-303). New York: The Guilford Press.
- Real Academia Española. (2011). *Ortografía de la lengua española*. Buenos Aires: Espasa.
- Sanchez & Cuetos (1998). Dificultades en la lectoescritura: naturaleza del problema. En González-Pienda, J. & Nuñez, J.C. *Dificultades del aprendizaje escolar* (pp. 263-288). Madrid: Pirámide.
- Sánchez Abchí, V., Diuk, B., Borzone, A. M. & Ferroni, M. (2009). El desarrollo de la escritura de palabras en español: interacción entre el conocimiento fonológico y ortográfico. *Interdisciplinaria*, 26 (1), 95-119.
- Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental child Psychology*, 72, 95-129.
- Share, D. L. (1995). Phonological recoding and self-teaching: sine quanon of reading acquisition. *Cognition*, 55, 151-218.
- Tanaka, J. S. & Huba, G. J. (1985). A fit index for covariance structure models under arbitrary GLS estimation. *British Journal of Mathematical and Statistical Psychology*, 38, 197-201.

Para citar este artículo:

Costa Ball, D., Palombo, A. L. & Cuadro, A. (2011). Propiedades psicométricas de una prueba experimental para la evaluación del nivel ortográfico. *Ciencias Psicológicas V (2)*: 167-178.

Recibido: 06/2011
 Revisado: 10/2011
 Aceptado: 11/2011