

## PROPIEDADES PSICOMÉTRICAS DE UNA PRUEBA EXPERIMENTAL DE DICTADO DE PALABRAS PARA LA EVALUACIÓN DE LA ORTOGRAFÍA EN ESCOLARES URUGUAYOS

### PSYCHOMETRIC PROPERTIES OF AN EXPERIMENTAL TEST OF WORD DICTATION FOR SPELLING ASSESSMENT OF URUGUAYAN SCHOOL CHILDREN

**Ariel Cuadro**

*Universidad Católica del Uruguay*

**Daniel Costa Ball**

*Universidad Católica del Uruguay*

**Ana Laura Palombo**

*Universidad Católica del Uruguay*

**Alexa von Hagen**

*Universidad Católica del Uruguay*

**Resumen:** Se presenta el desarrollo de una prueba de dictado de palabras (TEO-D) para la evaluación del nivel ortográfico de escolares de segundo a sexto grado de primaria, con el objetivo de contribuir a las prácticas educativas y a la valoración neuropsicológica y psicopedagógica de las dificultades en lenguaje escrito. La prueba se construyó a partir de la revisión y análisis de palabras frecuentes de nuestro medio, 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 que contienen el grafema h para el que no hay fonema y los grafemas c y s, que representan el sonido /s/. Se examina la estructura factorial, la fiabilidad y la validez de la prueba, que fue aplicada a una muestra de 488 alumnos de Montevideo, seleccionados mediante muestreo por conveniencia, entre dos colegios de educación privada de nivel socioeconómico medio bajo y medio alto. El análisis factorial confirmatorio, utilizando el programa NOHARM de un modelo de factor único, indica un ajuste adecuado a una estructura unidimensional, obteniendo una muy buena fiabilidad de consistencia interna en la muestra total y por curso.

**Palabras Clave:** *Psicometría, ortografía, evaluación*

**Abstract:** A word dictation test (TEO-D) to assess the spelling level of school children in grade 2 to 6, is presented, to contribute to educational practice, as well as neuropsychological and educational psychology assessment of written language difficulties. The test is constructed considering the review and analysis of frequent words in our environment, containing the alternation of the grapheme b and v, to represent the phoneme /b/, the graphemes j and g to represent the sound /j/, words, that can be written with ll or y, those containing the grapheme, that corresponds to no phoneme: h, and the graphemes c and s, that represent the sound /s/. The factor structure, reliability and validity of the test was examined, after being applied to a sample of 488 students from Montevideo, selected by convenience sampling, between two private schools from medium low and medium high socioeconomic status. The confirmatory factor analysis, using the program NOHARM for a one-factor model, indicates an adequate fit to a one-dimensional structure, achieving very good internal consistency reliability for the total sample and per course.

**Keywords:** *Psychometrics, orthography, assessment*

### Introducción

Formar niños capaces de leer y escribir es una de las funciones principales de la escuela, en tanto estos procesos constituyen habilidades fundamentales para la adquisición de nuevos

aprendizajes, favoreciendo con ello el desarrollo de sus sistemas cognitivos y sus capacidades adaptativas.

La escritura, al igual que la lectura, es una actividad cognitiva compleja, que exige el dominio de un sistema de escritura, con coherencia

interna propia, que guarda estrecha relación con su evolución histórica y su vinculación con la lengua hablada (Pontecorvo y Fabretti, 2003).

La escritura de una lengua supone la aplicación de ciertas reglas específicas, que la determinan y estructuran; lo que implica para su producción la activación de un conjunto de cuatro procesos: la planificación, la construcción sintáctica, la selección de palabras y los procesos motores (Cuetos, 2009).

Desde un punto de vista cognitivo, los mecanismos léxicos y subléxicos son los que se ponen en marcha, al momento de seleccionar grafemas para obtener la forma ortográfica o los grafemas de las palabras, que se quieren escribir (selección de palabras) (Ellis, 1985). El primer mecanismo se pone en marcha cuando se activa directamente la representación ortográfica de la palabra que se pretende escribir, sin realizar ningún análisis intermedio. Esto supone, que el sujeto cuenta con una representación ortográfica definida de la ortografía de la palabra en su léxico mental. El segundo, el mecanismo subléxico, permite obtener la ortografía de la palabra por la aplicación de las reglas de transformación fonema-grafema; se atribuyen secuencias fonológicas a secuencias ortográficas que se ensamblan en una palabra. Este mecanismo de recodificación fonológica ofrece la oportunidad de adquirir información ortográfica de la palabra (Jorm & Share, 1983; Share, 1995), permitiendo progresivamente desarrollar un léxico ortográfico que almacena representaciones ortográficas de palabras escritas, el que a su vez garantiza escribir las palabras en forma correcta (mecanismo léxico).

El mecanismo subléxico resulta por lo general suficiente para asegurar la escritura correcta de palabras en lenguas de ortografía transparente, incluso en castellano donde hay varios fonemas que se representan por grupos de grafemas, como por ejemplo: b, v, s, c, z, ll, y; pero no así para lenguas opacas, como el inglés donde hay una muy baja correspondencia entre fonemas y grafemas (Delattre, Bonin & Barry, 2006).

Perfetti (2010) señala además, que el significado de las palabras facilita la entrada al léxico ortográfico, ya que al decodificar una palabra cuyo significado es conocido, se refuerza la conexión entre la forma ortográfica de la palabra (la ortografía) y su significado. Esto ayuda a establecer una representación específica de esa palabra, especialmente importante en el

caso de palabras con ortografía o pronunciación excepcionales o irregulares.

De este modo, la ortografía se introduce como una variable posible de determinar en la medida en que el sujeto, que ha de apropiarse de las reglas de conversión fonema-grafema, es capaz de considerar las peculiaridades ortográficas de las palabras, guardándolas en su memoria (Costa Ball, Palombo y Cuadro, 2011).

Carbonell y Tuana (2004) entienden, que más precisamente la ortografía incluye todos aquellos aspectos de la escritura, que sobrepasan estrictamente la representación del fonema por medio de un grafema. Estos aspectos, que no tienen que ver con la aplicación directa de estas reglas de conversión fonema grafema y que definen a la ortografía, en el caso del castellano son:

- La escritura de los fonemas para los cuales existe una alternancia de grafemas: b, v, s, c, z, ll, y
- Los grafemas a los que no corresponde fonema: h
- Los grafemas a los que no corresponde ningún fonema de la lengua: x
- Todos los casos en que los fonemas se hacen de difícil percepción, tales como las consonantes b, d, q, p, etc.
- Acento ortográfico
- Los casos de escritura variable según el significado de la palabra o del énfasis que se le de (ej. casa).

Para la evaluación de los procesos léxicos de la escritura, las tareas principales suelen ser la denominación escrita de dibujos, la discriminación ortográfica (elección de la palabra bien escrita entre aquellas palabras y su homófono escrito con el grafema incorrecto) y el dictado de palabras (Cuetos, 2009; Carrillo, 2010). En las pruebas de dictado, por lo general se utilizan palabras de ortografía arbitraria, de ortografía reglada y palabras muy poco frecuentes o pseudopalabras.

En una publicación anterior (Costa Ball et al., 2011) se presentaron las características psicométricas de la prueba: Test de Eficacia Ortográfica, que evalúa el desarrollo del léxico ortográfico en los niños en edad escolar. La prueba consta de 99 ítems formados por una palabra y a cada una de ellas corresponde un distractor homófono (escrito con el grafema incorrecto). El alumno debe elegir la palabra ortográficamente correcta de cada par, disponiendo para ello de un tiempo límite de tres minutos. Fue diseñada

con el objetivo de representar un instrumento sencillo, de fácil aplicación. Los resultados mostraron, que presentaba adecuadas propiedades psicométricas (validez de contenido, análisis de ítems, dimensionalidad y fiabilidad).

El Test de Eficacia Ortográfica (Costa Ball et al., 2011) es una prueba del tipo discriminación ortográfica, cuya principal demanda cognitiva es el procesamiento ortográfico. Exige velocidad en la identificación de la forma ortográfica correcta de las palabras presentadas, que están almacenadas en su léxico ortográfico.

Con el objetivo de proponer una prueba complementaria para el Test de Eficacia Ortográfica (Costa Ball et al., 2011) en la evaluación de la escritura de palabras, se presentan las propiedades psicométricas de una prueba de dictado de palabras (TEO-D) elaborada para escolares uruguayos, de 2º a 6º grado de primaria. Esta prueba, además de exigir la recuperación y selección de la ortografía de la palabra (Delattre et al., 2006), permite observar, a diferencia del Test de Eficacia Ortográfica en la que la palabra correcta está presente, la eficacia del mecanismo ortográfico frente al fonológico para la escritura de palabras ortográficamente correctas; así como detectar si hay deficiencias en el dominio de las reglas de conversión fonema-grafema en palabras de uso frecuente. La recuperación de la ortografía de la palabra, al igual que el deletreo de palabras presentaría mayores dificultades, que el reconocimiento visual de palabras escritas, en tanto exige mayores demandas de la memoria ortográfica relacionadas con la representación de la posición y secuencia de los grafemas (Huntley, Silliman, Berninger & Dow, 2012).

## 1. Método

### 1.1. Participantes

La muestra estuvo compuesta por 488 alumnos de segundo a sexto año de primaria, 246 varones y 240 mujeres, pertenecientes a dos colegios privados de la ciudad de Montevideo emplazados en estratos socioeconómicos de nivel medio-bajo y medio-alto.

La elección de estudiantes de estos grados escolares se basó, en que los mismos ya se iniciaron en la lectura y escritura y al mismo tiempo, al transcurrir la escolaridad, se aprecia un aumento en el 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. Instrumentos

- *Prueba de Eficacia Lectora (TECLE)* de Cuadro, Costa Ball, Trías y Ponce de León (2009): Se trata de un test de velocidad con un tiempo límite de 5 minutos, que evalúa la eficacia lectora a partir de la presentación de 64 ítems. El alumno debe completar cada ítem, que consiste en la presentación de una frase incompleta y, marca la opción correcta de cuatro palabras posibles (tres incorrectas y una correcta). La realización de los estudios psicométricos revelan el carácter unidimensional, fiable (fiabilidad test-retest de .88) y válido de esta prueba. El análisis de funcionamiento diferencial de los ítems (FID) evidenció resultados de al menos un 6% de los ítems con sospecha leve de FID (Costa Ball, Gründel y Cuadro, 2011).

- *Test de Evaluación del nivel Ortográfico (TEO)* de Costa Ball et al. (2011): Se trata de una prueba de velocidad con un tiempo máximo de 3 minutos para medir la eficacia ortográfica por medio de la presentación de 100 ítems dicotómicos (acierto/error). El sujeto debe marcar la opción ortográficamente correcta de dos palabras presentadas: una palabra es el blanco y el distractor es la misma palabra escrita con los grafemas fonéticamente equivalentes.

- Se desarrolló una *Prueba de dictado de palabras para la evaluación del nivel ortográfico (TEO-D)*, que consta de 60 ítems. Los mismos fueron seleccionados de las listas de palabras de uso frecuente para los escolares uruguayos propuesta por Carbonell y Tuana (2004).

Este conjunto de palabras fue sometido, a través del juicio de expertos, a un proceso de actualización de la frecuencia de las mismas, ya que la lista tenía más de 30 años de antigüedad. Cuatro docentes por curso analizaron la adecuación de los ítems mediante una escala Likert de cuatro puntos. Se utilizó el acuerdo inter-jueces para eliminar los ítems más problemáticos.

La lista de Carbonell y Tuana (2004) estaba organizada por vocablos distribuidos por centenas, de tal modo, que las palabras de la centena uno eran las más frecuentes y continuaban de orden decreciente en relación a su frecuencia. Para la elaboración de esta prueba se tomaron en cuenta aquellas palabras que contenían 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 que contienen el grafema para el que no hay fonema: h y los grafemas c

y s que representan el sonido /s/. La razón de ello fue no agregar otra dificultad ortográfica para el mismo sonido, ya que por frecuencia se seleccionaron palabras que contienen ce-ci. No se consideraron las palabras con grafema z, por su baja frecuencia, ni tampoco 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).

De esta forma la lista quedó conformada por 425 palabras, que se distribuyen porcentualmente, según cada dificultad:

Tabla 1  
*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

Como el número era demasiado elevado para someterlo a consideración de los jueces, se decidió acortar la misma a la mitad. Para esto se numeró la lista de palabras en número par e impar al azar y se determinó que se utilizarían las palabras impares para la conformación de la lista del Test de Eficacia Ortográfica (Costa Ball et al., 2011). A su vez se le agregaron las palabras, que componen la prueba de Carrillo (2010) para escolares españoles, de más re-

ciente 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.

La lista final resultó compuesta por 231 palabras, cuyas estructuras ortográficas son representativas del léxico, de 1, 2, 3 y 4 sílabas. Esta fue entregada a los jueces para que asignaran cada ítem a una de las siguientes categorías: Nada frecuente, Poco frecuente, Frecuente y Muy frecuente.

Los criterios selección de jueces fueron los siguientes: tener un mínimo de dos años de experiencia en el grado, tener al menos cinco años de titulación como maestro y experiencia en el trabajo con niños de estratos socioeconómicos de nivel medio o medio-alto.

Del análisis de estos resultados surgieron las palabras que conformarían la prueba del Test de Eficacia Ortográfica (Costa Ball et al., 2011). Analizando las elecciones de los jueces, se estableció que no existían diferencias significativas entre la lista original conformada por Carbonell y Tuana (2004) y la actual. Según lo establecido por los jueces, solamente el 30% de las palabras escogidas por ellos se ubicaban entre la 7ª y 10ª centena. Las demás palabras estaban ubicadas entre la centena 1 y 6, lo que demuestra la vigencia de la clasificación de palabras realizadas por Carbonell y Tuana (2004).

En función de esto y de los objetivos de esta prueba para ser complementaria al Test de Eficacia Ortográfica (Costa Ball et al., 2011), se decide tomar la lista de palabras pares que no se tomaron en este último, respetando la distribución de las palabras de la lista, según el índice de dificultad antes planteado (Tabla 2).

Tabla 2  
*Palabras que conforman la prueba experimental de dictado*

BICICLETA	HABLAR	BESO	SABER	ENVIAR	VERDAD
MUJER	HACER	DISPOSICIÓN	LLAMAR	DELEGACIÓN	ABRAZO
VESTIR	NOVIEMBRE	JUSTICIA	GENTE	PRODUCIR	SERVICIO
BARCO	HERMANO	CENTRO	BAJO	BASTANTE	OFICIAL
VIENTO	LLEVAR	TELEVISIÓN	CONCEDER	EJÉRCITO	POBRE
HIJO	LICEO	TRABAJADOR	NECESARIO	FLORECER	SOCIEDAD
SOBRE	NUEVO	TUYO	PUEBLO	DICIEMBRE	PARECER
HOY	YA	OBRA	ASOCIACIÓN	VIVIR	PALABRA
DIVERTIR	ESTRELLA	RECIBIR	ESPECIAL	VIAJAR	SOCIAL
NACIONAL	PRECIO	ACCION	LLUVIA	DIRIGIR	CIERTO

### 1.3. Procedimiento

Luego de obtener el consentimiento de los colegios, que participaron en el estudio, una psicopedagoga, en coordinación con las maestras de las mismas instituciones, aplicó la prueba en su versión inicial (experimental) junto con las otras dos pruebas señaladas anteriormente, en horario escolar.

Los alumnos fueron informados, que la participación era voluntaria y que sus datos personales y los resultados, se tratarían con total confidencialidad. Una vez asegurado un clima de trabajo adecuado, se procedió a explicar la consigna de la prueba. Se les explicó que debían permanecer en total silencio para poder escuchar con atención la lista de 60 palabras que les dictaría el evaluador. Cada palabra era anunciada con el número de ítem correspondiente y fue pronunciada en forma repetida dos veces. Los estudiantes debían anotar la palabra en un espacio predeterminado en el protocolo, junto con el número del ítem. Una vez que se finalizó con el dictado de la lista completa de 60 palabras, se recogieron los protocolos de aplicación.

### 1.4. Análisis de datos

Para los análisis de datos, se utilizó el Statistical Package for Social Sciences (SPSS) en su versión 1. Se calcularon los estadísticos descriptivos, las correlaciones ( $r$ ) producto-momento de Pearson (validez concurrente), la correlación Biserial Puntual para calcular los coeficientes de discriminación para ítems dicotómicos y una variable continua.

Se realizó el análisis factorial confirmatorio (AFC) con el programa NOHARM (Fraser & McDonald, 1988), para evaluar el ajuste de los datos a un modelo de factor único. Este programa se utilizó para ajustar modelos unidimensionales o multidimensionales. Para la evaluación de la bondad de ajuste se utilizó el índice de Tanaka, que debía ser mayor, que el valor de referencia (0.95) para establecer un nivel adecuado de ajuste (Tanaka & Huba, 1985).

## 2. Resultados

**Estudios de las propiedades psicométricas de los ítems.** El análisis de la calidad métrica o análisis de ítems, que componen un test, es una etapa relevante en el desarrollo de instrumentos. Por este motivo, estos análisis, comienzan luego de los análisis de validez de contenido, evaluación de jueces, etc. (Carretero-Dios y Pérez, 2005). El TEO-D es un test dicotómico, en el que la evaluación de la calidad métrica de los ítems de la prueba, se presenta en función de las respuestas del sujeto al ítem (acierto o fracaso).

**Índice de Dificultad del ítem.** Este aspecto es definido por la proporción de sujetos que responden correctamente a ese ítem. Los resultados con índices cercanos a 1 evidencian un ítem difícil de responder correctamente, mientras los índices próximos a 0 suponen un ítem sencillo. Se seleccionaron ítems con índices de dificultad bajos -atendiendo a los diferentes grados escolares- para que el test se tornara discriminativo. En la Tabla 3 se muestran los índices de dificultad para todos los ítems de la prueba de dictado de palabra (TEO-D).

Tabla 3  
Índice de dificultad del ítem

Ítems	ID	Ítems	ID	Ítems	ID	Ítems	ID
1	,71	16	,89	31	,75	46	,71
2	,91	17	,89	32	,62	47	,83
3	,77	18	,90	33	,75	48	,93
4	,92	19	,92	34	,83	49	,81
5	,89	20	,86	35	,70	50	,43
6	,91	21	,86	36	,74	51	,79
7	,93	22	,93	37	,92	52	,61
8	,98	23	,77	38	,40	53	,57
9	,81	24	,84	39	,80	54	,84
10	,94	25	,33	40	,81	55	,91
11	,70	26	,79	41	,73	56	,72
12	,81	27	,76	42	,84	57	,74
13	,78	28	,55	43	,85	58	,92
14	,92	29	,84	44	,73	59	,76
15	,73	30	,70	45	,69	60	,57

**Discriminación del ítem.** Es la capacidad del ítem para diferenciar sujetos de baja, media y alta puntuación en un criterio (Barbero, Vila y Suarez, 2006). Se utilizó como criterio interno, la puntuación de los sujetos en la prueba de dictado de palabras (TEO-D). Este índice de discriminación, al emplear un criterio interno, se convierte en un índice de homogeneidad. Específicamente se utilizó el procedimiento de Correlación Biseral Puntual. El mismo sirve para conocer el grado de asociación entre el ítem y la prueba como criterio interno. En la Tabla 4 se muestran los índices de discriminación de todos los ítems de la prueba.

En la Tabla 5, se muestra la cantidad de sujetos, medias, mediana, desviación típica, asimetría, y la prueba de Kolmogorov-Smirnov para una muestra para el estudio de la normalidad. Como se observa en la Figura 1, el promedio de acierto en los 60 ítems que conforman la prueba, aumenta en la medida que se avanza en el curso. Del estudio de los puntajes mínimos y máximos obtenidos por año, se aprecia que en segundo

grado hay alumnos, que se acercan a los puntajes máximos, evidenciando, que para los alumnos competentes en las habilidades ortográficas, (incluso en segundo año) es posible alcanzar el techo de la prueba. En el estudio del coeficiente de asimetría, se puede apreciar una distribución normal en segundo año. A medida que avanzan los cursos, la distribución se aparta del criterio de normalidad, reflejando un aumento del índice de asimetría con valores negativos, indicando que muchos alumnos obtienen puntajes altos.

En la última columna de la Tabla 5, se presentan los resultados del estudio de normalidad univariante de los datos (dictado de palabras), en el que se utiliza el contraste de Kolmogorov-Smirnov para valorar la normalidad de los datos. Valorando los datos estadísticos, se observa, que a partir de cuarto año se encuentran valores significativos, por lo que se debe rechazar la hipótesis de normalidad en 4º, 5º y 6º año. En cambio, para los cursos inferiores, de segundo y tercer año, los resultados en la prueba de dictado tienen una distribución normal.

Tabla 4  
Índices de discriminación de ítems (Correlación Biseral Puntual)

ítem	r <sub>bp</sub>						
1	0,58	16	0,56	31	0,26	46	0,56
2	0,43	17	0,50	32	0,52	47	0,62
3	0,48	18	0,43	33	0,61	48	0,50
4	0,22	19	0,33	34	0,44	49	0,60
5	0,47	20	0,56	35	0,48	50	0,26
6	0,51	21	0,29	36	0,58	51	0,56
7	0,33	22	0,35	37	0,45	52	0,63
8	0,34	23	0,56	38	0,43	53	0,59
9	0,56	24	0,40	39	0,61	54	0,60
10	0,50	25	0,34	40	0,59	55	0,43
11	0,72	26	0,66	41	0,52	56	0,68
12	0,67	27	0,60	42	0,64	57	0,57
13	0,55	28	0,29	43	0,63	58	0,47
14	0,59	29	0,51	44	0,28	59	0,69
15	0,54	30	0,41	45	0,58	60	0,45

Tabla 5  
Medidas de tendencia central y dispersión del puntaje total por curso

Curso	N	M	DE	As	Min.	Max.	Normalidad
Segundo	96	34.48	12.37	-0.13	3	59	.821
Tercero	92	43.09	11.25	-0.67	17	60	1.19
Cuarto	105	47.36	11.46	-2.43	0	60	1.65**
Quinto	86	52.55	6.43	-2.00	21	60	1.44*
Sexto	109	54.34	6.18	-2.13	27	60	1.88*
Total	488	46.50	12.18	-1.28	0	60	3.42***

Nota: \* p < .05; \*\* p < .01; \*\*\* p < .001

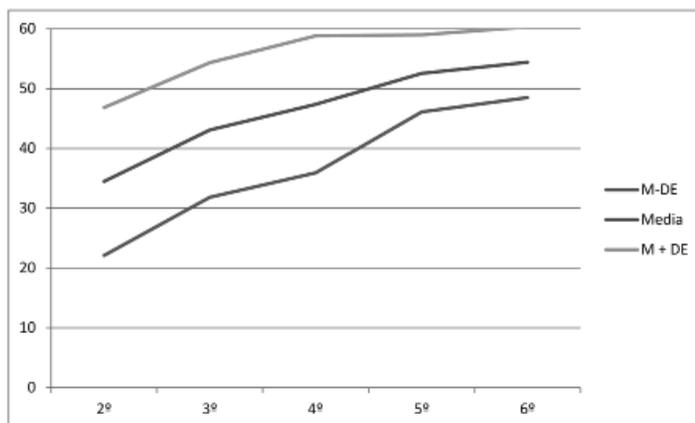


Figura 1. Medidas de tendencia central y dispersión por curso en el TEO-D

**Estudio de dimensionalidad.** El siguiente paso implicó la exploración de la estructura interna de la prueba, a través de la validez de constructo. Esto supone la obtención de evidencias de validez interna. Esta etapa se debe iniciar, luego de evaluar la validez de contenido, el análisis de ítems. Una vez obtenida la dimensionalidad del instrumento, es posible reportar la fiabilidad y no antes (Carretero-Dios y Pérez, 2005).

El TEO-D es un test de rendimiento, con ítems dicotómicos que se puntúan en términos de acierto o fracaso. Martínez, Hernández y Hernández (2006) aconsejan utilizar los métodos de Análisis Factorial no lineal de McDonald, cuando se quiera realizar un análisis factorial confirmatorio con datos dicotómicos. Por este motivo, se evalúa la dimensionalidad de la prueba de dictado TEO-D a través del programa NOHARM de Fraser & McDonald (1988) y del índice de bondad de ajuste de Tanaka, que debe ser mayor que el valor de referencia (0.95) y lo más cerca posible de su valor máximo (1) para lograr un buen ajuste al modelo valorado.

El análisis se realizó en dos etapas: en la primera se corrió el análisis con toda la muestra (2º - 6º año) para comprobar la unidimensionalidad del instrumento de medida con los 488 alumnos. En una segunda etapa, se volvió a correr el análisis, pero en cuatro bloques solapados entre ellos. En el primer grupo los cursos de 2º y 3º; en el segundo los cursos de 3º y 4º; en el tercero los cursos de 4º y 5º y en el último los cursos de 5º y 6º año.

Los resultados obtenidos sugieren que la prueba TEO-D es unidimensional en toda la población estudiada y en cada uno de las agrupaciones realizadas. Se obtuvo un índice Tanaka mayor de .95 en todos los análisis. Se puede concluir que el TEO-D cumple con el supuesto de unidimensionalidad.

**Estudio de Fiabilidad.** Una vez conocida la dimensionalidad de la prueba de dictado de palabras (TEO-D), que responde a un modelo de factor único, se estudió la fiabilidad de la misma. El método adecuado para este tipo de estudio de la fiabilidad debe basarse en las fórmulas de consistencia interna de Kuder-Richardson, sugerida para datos dicotómicos de acierto y fracaso, como es el caso de la prueba presentada. La fiabilidad total para toda la muestra es de .94, para segundo y tercer año es de .93, para cuarto .94, para quinto .88 y para sexto año .90 (ver diagonal en la Tabla 6 donde se muestra la fiabilidad de cada prueba).

**Estudio de validez.** Para evaluar la validez de la prueba de dictado de palabras (TEO-D) se utilizó el procedimiento de validez de constructo y concurrente. El primero de los dos, fue desarrollado anteriormente, cuando se presentaron los estudios de validez de constructo con los estudios sobre la dimensionalidad de la prueba, a través del programa NOHARM, para ajustar modelos unidimensionales y multidimensionales. En segundo lugar, se empleó la validez concurrente, para evaluar la prueba de dictado de palabras (TEO-D) y el criterio al mismo tiempo

(concurrentemente), con el objetivo de validar el TEO-D por comparación con otras dos pruebas, ya validadas previamente: TECLE (Cuadro et al., 2009) y TEO (Costa Ball et al., 2011). La Tabla 6, muestra las correlaciones entre las puntuaciones en el puntaje total en la prueba de dictado de palabras (TEO-D). Se observan valores significativos en sentido positivo con el puntaje en eficacia lectora (TECLE;  $r = .85^{**}$ ) y con el puntaje de la prueba de eficacia ortográfica (TEO;  $r = .76^{**}$ ).

Tabla 6  
Validez Concurrente entre el TEO-D, el TECLE y el TEO

Escalas	1	2	3
1. TEO	<b>.76</b>		
2. TECLE	.85**	<b>.88</b>	
3. TEO-D	.76**	.73**	<b>.94</b>

Nota: Los coeficientes alfa se presentan en negritas a lo largo de la diagonal

**Análisis de varianza.** Para comprobar si existen diferencias significativas en el puntaje total sobre la variable curso, se empleó el análisis de varianza de un factor (ANOVA). Anteriormente, en la Tabla 5 se mostró, que el puntaje total en dictado de palabras sólo cumple con el supuesto de normalidad en los cursos de segundo y tercer año, no así en el resto de los cursos, no pudiéndose entonces utilizar el ANOVA. Pero cuando los tamaños de los grupos son grandes, el estadístico  $F$  se comporta razonablemente bien, incluso con poblaciones alejadas de la normalidad (Pardo y Ruiz-Díaz, 2002).

Primero se estudió la homocedasticidad o igualdad de varianza. Los resultados muestran, que se debe rechazar la hipótesis de igualdad de varianza (Estadístico de Levene = 59.7;  $gl = 2$ ;  $p < .000$ ) y concluir que en las poblaciones definidas por la variable curso, las varianzas del puntaje en la prueba de dictado de palabras no son iguales. Como no se pudo utilizar la prueba ANOVA, por el cumplimiento del supuesto de homocedasticidad, se utilizó el estadístico de Welch como alternativa. Puesto que este estadístico es menor que .05, se pudo rechazar la hipótesis de igualdad de medias y concluir, que el puntaje total en el TEO-D difería para la variable curso. Se utilizó el procedimiento de comparaciones post hoc para saber qué media difería de otra, a través del método de Tukey. No se registraron diferencias significativas, con excepción de las que se señalan entre 5º y 6º año en el puntaje total en la prueba TEO-D.

En el resto de las comparaciones posibles, se registraron diferencias entre los grados. Se puede establecer la existencia de cuatros subgrupos basados en el grado de parecido existente entre sus medias. En la Tabla 7, se muestra la relación entre los distintos subgrupos: el primer subgrupo conformado por los puntajes más altos en TEO-D en los cursos de 5º y 6º año; en segundo lugar, los puntajes de cuarto año, que son significativamente menores, que los anteriores, pero más altos, que los de 3º año; en tercer lugar, los alumnos de segundo año, que rinden por debajo de los alumnos de 4º, 5º y 6º año pero por encima de los de 2º, y por último, los alumnos de segundo año con un rendimiento por debajo del resto de los grupos ( $2^\circ < 3^\circ < 4^\circ < 5^\circ = 6^\circ$ ).

Tabla 7  
Clasificación de subconjuntos homogéneos

Curso	N	1	2	3	4
2º	96	34.48			
3º	92		43.07		
4º	105			47.36	
5º	86				52.55
6º	109				54.34

### 3. Discusión

Los resultados obtenidos en este estudio, permiten confirmar tanto las adecuadas propiedades psicométricas de los ítems, como la validez y fiabilidad del *Test de Evaluación del Nivel Ortográfico (TEO-D)* para alumnos de 2º a 6º de nivel primario. Los análisis de validez de contenido, análisis de ítem, dimensionalidad y fiabilidad, han aportado evidencia de la validez de una estructura unidimensional de la prueba.

El desempeño de los alumnos mejora progresivamente año a año, de segundo hasta quinto año. Algunos alumnos de segundo año logran un rendimiento sin errores, evidenciando el hecho de disponer a ese nivel escolar, de las representaciones ortográficas en su léxico mental.

No se encontraron diferencias significativas en el rendimiento de quinto y sexto año, lo que era esperable, ya que en estos niveles de escolaridad, las palabras dictadas son de alta frecuencia y de significado conocido, lo que ayuda a establecer una representación específica de las palabras. (Perfetti, 2010).

En relación con los índices de discriminación, solamente 5 de 60 ítems presentaron un índice de correlación biseral puntual inferior a .29, demostrando que estos ítems correlacionan en forma baja con el total de la escala, que es unidimensional. Los ítems a revisar son: barco, beso, saber, bastante y dirigir.

La prueba presentó una correlación significativa (0,76) con el Test de Eficacia Ortográfica (Costa Ball et al., 2011), que también mide el desarrollo del léxico ortográfico en los niños de edad escolar, a partir de la elección de homófonos/pseudohomófonos. Este resultado sugiere, que no habría diferencias significativas entre ambos tipos de pruebas en la evaluación de los procesos léxicos de la escritura en escolares. Si bien, la prueba de dictado de palabras parece exigir mayores demandas cognitivas, en tanto el sujeto tiene que recuperar, seleccionar y producir la ortografía de la palabra (Delattre et al., 2006; Huntley et al., 2012). En futuros estudios se podría analizar si en el caso de los estudiantes con dificultades o trastornos en lenguaje escrito aparecen o no diferencias en la ejecución de esta prueba, considerando el menor desarrollo del léxico ortográfico que presentan estos sujetos (Cuadro, 2010).

De igual forma, la prueba TEO-D presentó una correlación positiva alta (0,85) con el Test de Eficacia Lectora (Cuadro et al., 2009), que evalúa el nivel lector, considerando las variables de precisión y fluidez en el reconocimiento de la palabra escrita. Una forma de interpretar esta relación es considerando la incidencia del desarrollo del léxico ortográfico para el reconocimiento rápido y preciso de la palabra (Share, 1995), como para la escritura correcta de palabras; en particular para las palabras de ortografía arbitraria o irregulares. El léxico ortográfico implica entonces, que cada representación de la palabra almacenada tenga un alto nivel de activación, para que su recuperación al momento de escribirla o leerla sea automática (Cuadro, 2010).

En suma, la prueba de dictado de palabras (TEO-D) constituye un instrumento de medida válido y fiable, que permite explorar, a partir de un único factor, la evaluación del nivel ortográfico. 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.

Al mismo tiempo la tarea de escritura de palabras dictadas, permite, en función de los errores, no solo evaluar el desarrollo del léxico

ortográfico, sino verificar deficiencias en el dominio de las reglas de transformación fonema-grafema (sustituciones, omisiones, alteraciones, agregados), contribuyendo de esta forma al diagnóstico de las dificultades o trastornos en escritura.

#### 4. Referencias

- Barbero, M., Vila, E. y Suárez, J. C. (2006) *Psicometría*. Unidades Didácticas. Madrid: UNED.
- Carbonell, M. A., y Tuana, E. (2004). *La Ortografía de uso. Un espacio de reflexión y desafío*. Compilación de escritos de la Prof. Emérita Psicóloga Ma Angélica Carbonell de Grompone, (pp. 245- 280.) Montevideo: Sociedad de Dislexia y Coeditores.
- Carretero-Dios, H. y 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 lengua escrita y sus dificultades*. (pp.111-132) Málaga: Aljibe.
- Costa Ball D., Gründel, M. y Cuadro, A. (2011). Impacto y funcionamiento diferencial de los ítems respecto al género en el Test de Eficiencia Lectora (TECLE). *Ciencias Psicológicas* V (1), 47-57.
- Costa Ball, D., Palombo, A. L. y Cuadro, A. (2011). Propiedades psicométricas de una prueba experimental para la evaluación del nivel ortográfico. *Ciencias Psicológicas* V (2), 167-178.
- Cuadro, A., Costa Ball, D., Trías, D. y Ponde de León, P. (2009). Evaluación del nivel lector. Test de eficacia lectora (TECLE) de J. Marín y M.Carrillo. Manual Técnico. Montevideo: Prensa Médica.
- Cuadro, A. (2010). *La lectura y sus dificultades. Dislexia Evolutiva*. Montevideo: Grupo Magro
- Cuetos, F. (2009) *Psicología de la Escritura*. Madrid: Wolters Kluwe España.
- Delattre, M., Bonin, P., & Barry, Ch. (2006) Written Spelling to Dictation: Sound-To-Spelling Regularity Affects Both Writing Latencies and Durations. *Journal of Experimental Psychology: Learning, Memory and Cognition*, (32)6, 1330-1340.
- Ellis, A. W. (1985). The cognitive neuropsychology of developmental (and acquired) dyslexia: A critical survey. *Cognitive Neuro psychology*, 2, 169-205.
- Fraser, C. & McDonald, R. P. (1988) NOHARM: Least squares item factor analysis. *Multivariate behavioral research* 23, 267 – 269.
- Huntley B. R., Silliman, E., Berninger, V. & Dow, M. (2012). Linguistic Pattern Analysis of Misspellings of Typically Developing Writers in Grades 1-9. *Journal of Speech, Language, and Hearing Research*, 55, 1587-1599.
- Jorm, A., & Share, D. (1983). Phonological recording and reading acquisition. *Applied Psycholinguistics*, 4, 103-147.
- Martínez, M.R., Hernández, M.J. & Hernández, M. V. (2006). *Psicometría*. Madrid: Alianza.

- Pardo, A., y Ruiz-Díaz, M. (2002). SPSS 11 Guía para el análisis de datos. Madrid: McGraw-Hill/ Interamericana de España.
- Perfetti, C. A. (2010). Decoding, Vocabulary, and Comprehension. The golden Triangle of Reading Skill. En M. & Kucan, Bringing reading research to life (pp. 291-303). New York: The Guilford Press. Recuperado de <http://www.pitt.edu/~perfetti/PDF/Decoding%20Beck%20festschrift%20chapt.pdf>
- Pontecorvo, C., y Fabretti, D. (2003) Aprender un sistema de escritura, aprender una lengua. En C. Pontecorvo, *Manual de Psicología de la Educación*. Madrid: Popular.
- Real Academia Española (2011). "Ortografía de la lengua española". (1ª ed.). Buenos Aires: Espasa.
- 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:

Cuadro, A., Costa Ball, D., Palombo, A. L. y von Hagen, A. (2013). Propiedades psicométricas de una prueba experimental de dictado de palabras para la evaluación de la ortografía en escolares uruguayos *Ciencias Psicológicas VII (1)*: 57 - 66.

*Recibido: 03/2013*  
*Revisado: 04/2013*  
*Aceptado: 05/2013*