



Consejo General  
de la **Psicología**  
ESPAÑA



Miembro de la Federación Europea  
de Asociaciones de Psicólogos

Conde Peñalver, 45. 5º Izq.  
28006 Madrid – España  
Tel.: +34 91 444 90 20  
Fax: +34 91 309 56 15  
secop@cop.es  
www.cop.es

**EVALUACIÓN DE LA PRUEBA DE  
RENDIMIENTO ORTOGRÁFICO 1-2  
(PRO 1-2)  
PRUEBA DE RENDIMIENTO  
ORTOGRÁFICO 3 (PRO 3)**

## RESUMEN DE LA VALORACIÓN DEL TEST

### Descripción general

Características	Descripción
Nombre del test	PRO prueba de rendimiento ortográfico 1-2; PRO prueba de rendimiento ortográfico 3
Autor	Galve Manzano, J.L.; Trallero Sanz, M.; Martínez Arias, R. y Dioses Chocano, A.S.
Autor de la adaptación española	-
Variable(s)	Rendimiento ortográfico
Área de aplicación	Psicología educativa
Soporte	Papel y lápiz e Informatizado

### Valoración general

Características	Valoración	Puntuación
Materiales y documentación	Buena	4
Fundamentación teórica	Excelente	5
Adaptación	--	--
Análisis de los ítems	Buena	4
Validez: contenido	Adecuada-buena	3,5
Validez: relación con otras variables	Adecuada	2,5
Validez: estructura interna	Adecuada	3
Validez: análisis del DIF	--	--
Fiabilidad: equivalencia	--	--
Fiabilidad: consistencia interna	Buena	4
Fiabilidad: estabilidad	--	--
Fiabilidad: TRI	--	--
Fiabilidad: inter-jueces	--	--
Baremos e interpretación de puntuaciones	Adecuada-buena	3,5

## Comentarios generales

De acuerdo con la documentación aportada la prueba de rendimiento ortográfico PRO 1-2-3 pretende constituirse en un instrumento importante para la predicción del desempeño académico en nivel elemental y medio de educación, así como también para la evaluación de la aplicación del conocimiento ortográfico adquirido por los alumnos en primaria. Cabe destacar que la revisión y fundamentación teórica que soporta el test permite que el usuario comprenda claramente los aspectos a evaluar, el papel que ellos juegan en el rendimiento académico general y las teorías subyacentes a ellos.

Las instrucciones para los usuarios son claras y están redactadas para cada una de las subpruebas. El tiempo de aplicación no es fijo y debe adaptarse de acuerdo al desempeño general del alumno o la mayoría de alumnos (cuando se trata de una aplicación colectiva), aunque se establece un tiempo medio entre 80 y 90 minutos para la prueba completa. La recomendación de los autores es presentar la prueba en dos sesiones. El diseño de los cuadernillos de aplicación de la prueba es adecuado, aunque sería conveniente que se eliminaran los apartados de instrucciones para el evaluador y se adiccionaran ejemplos para todas las subpruebas. Como punto fuerte, la fundamentación teórica es buena, y los ítems muestrean de forma bastante exhaustiva el constructo a medir y sus facetas

El ámbito de uso del instrumento es claramente el sector educativo, con respecto al rendimiento, pero también puede ser una herramienta de apoyo en la identificación de dificultades específicas de aprendizaje y del procesamiento del lenguaje escrito.

En futuras ediciones de la herramienta se debería añadir nueva información, como por ejemplo: a) la técnica de muestreo empleada para seleccionar la muestra que hizo parte del estudio definitivo así como criterios de inclusión y exclusión; b) una descripción detallada de la consulta a expertos que permitió aportar evidencias de validez de contenido; c) descripción operacional de los niveles en los que los profesores han de clasificar el desempeño de los alumnos y los niveles de acuerdo alcanzados cuando varios docentes evaluaron a un mismo alumno; d) incorporar instrucciones y módulos para el manejo de poblaciones especiales; y e) nuevos estudios psicométricos (p.ej., análisis factorial, DIF, estimación de la fiabilidad, etc.).

# ANÁLISIS DETALLADO DE LA PRUEBA

## 1. DESCRIPCIÓN GENERAL DEL TEST

### 1.1. Nombre del test:

- PRO prueba de rendimiento ortográfico 1-2; PRO prueba de rendimiento ortográfico 3.

### 1.2. Nombre del test en su versión original:

- --

**1.3. Autor del test original:**

- Galve Manzano, J.L.; Trallero Sanz, M.; Martínez Arias, R. y Dioses Chocano, A.S.

**1.4. Autor de la adaptación española:**

- --

**1.5. Editor del test en su versión original:**

- CEPE.

**1.6. Editor de la adaptación española:**

- --

**1.7. Fecha de publicación del test original:**

- 2010.

**1.8. Fecha de la publicación del test en su adaptación española:**

- --

**1.9. Fecha de la última revisión del test en su adaptación española:**

- --

**1.10. Área general de la(s) variable(s) que pretende medir el test:**

- Rendimiento ortográfico, rendimiento académico/competencia curricular.

**1.11. Breve descripción de la(s) variable(s) que pretende medir el test:**

- Es una prueba objetiva de rendimiento ortográfico para niños desde nivel medio hasta superior. Consiste en una serie de escalas que miden diferentes componentes de dicho rendimiento (Procesamiento léxico-ortográfico, Sintáctico y Semántico) y permite también obtener una puntuación Global.
- El test PRO-1 pretende evaluar el nivel ortográfico de alumnos de primaria, específicamente en nivel elemental (3ro y 4to EP). El test PRO-2 evalúa la misma variable en alumnos de nivel medio (5to y 6to EP). El test PRO 3 pretende evaluar el nivel de rendimiento ortográfico de alumnos de secundaria, específicamente desde primer a tercer curso de ESO. El test PRO evalúa el nivel del rendimiento ortográfico mediante tres aspectos: Procesamiento léxico, procesamiento sintáctico y procesamiento semántico. El procesamiento léxico está compuesto por la escritura de palabras que aporta una valoración global del uso de las reglas ortográficas, la escritura de pseudopalabras permite evaluar el uso y aplicación de reglas y el uso de normas de acentuación, formada por acentuación reglada y monosílabos, uso correcto de tildes y acentuación de monosílabos y relativos/interrogativos, que valora el conocimiento del uso de la tilde y la memoria ortográfica. El procesamiento sintáctico se refiere al uso de signos de puntuación en frases, así como el uso de signos de interrogación y exclamación. Finalmente, el procesamiento semántico se evalúa con un dictado de homófonos. Este test es de aplicación colectiva, aunque también admite la aplicación individual, los autores sugieren dividir la aplicación en dos sesiones: en la primera sesión el

procesamiento léxico (escritura de palabras, pseudopalabras y acentuación reglada). En la segunda sesión las demás pruebas de acentuación de monosílabos, puntuación de textos y escritura de homófonos.

**1.12. Áreas de aplicación:**

- Psicología educativa.

**1.13. Formato de los ítems:**

- Respuesta construida, y respuesta semiconstruida.

**1.14. Número de ítems:**

- PRO-1: 105 ítems; PRO-2: 125 ítems; PRO-3: 126 ítems
- PRO-1 Escritura de palabras: 64 ítems, escritura de pseudopalabras: 12 ítems, uso de signos de puntuación: 7 ítems, uso de signos de interrogación y exclamación: 7 ítems, y dictado de homófonos: 15 ítems. Total 105 ítems.
- PRO-2 Escritura de palabras: 64 ítems, escritura de pseudopalabras: 12 ítems, uso de acentuación de monosílabos y relativos/interrogativos: 20 ítems, uso de signos de puntuación: 7 ítems, uso de signos de interrogación y exclamación: 7 ítems, y dictado de homófonos: 15 ítems. Total 125 ítems.
- PRO-3: Escritura de palabras: 64 ítems, escritura de pseudopalabras: 12 ítems, uso de acentuación de monosílabos y relativos/interrogativos: 20 ítems, uso de signos de puntuación: 8 ítems, uso de signos de interrogación y exclamación: 7 ítems, y dictado de homófonos: 15 ítems. Total 126 ítems.

**1.15. Soporte:**

- Papel y lápiz e informatizado.

**1.16. Cualificación requerida para el uso del test de acuerdo con la documentación aportada:**

- Otra: Administradas por el profesor de lenguaje.

**1.17. Descripción de las poblaciones a las que el test es aplicable:**

- Alumnos de Lengua Española de 3º a 6º curso de primaria (PRO 1 y 2) y de 1º a 3º de ESO (PRO 3).

**1.18. Existencia o no de diferentes formas del test y sus características:**

- Existen 3 versiones del test administrables tanto en papel y lápiz como por ordenador: PRO-1 (3º y 4º de primaria); PRO-2 (5º y 6º de primaria) y PRO-3 (1º, 2º y 3º de ESO). La administración informatizada no precisa requisitos inusuales.

**1.19. Procedimiento de corrección:**

- Manual y Automatizada por ordenador.

**1.20. Puntuaciones:**

- Las puntuaciones directas son la suma del número de errores ortográficos cometidos. Al ser ítems de respuesta construida, no se precisan correcciones por azar.

**1.21. Escalas utilizadas:**

- Centiles, decatipos y puntuaciones T.

**1.22. Posibilidad de obtener informes automatizados:**

- Sí.

**1.23. Tiempo estimado para la aplicación del test:**

- En aplicación individual: 80-90 minutos.
- En aplicación colectiva: 80-90 minutos.

**1.24. Documentación aportada por el editor:**

- Manual, Discos u otros dispositivos magnéticos y cuadernillos de trabajo.

**1.25. Precio de un juego completo de la prueba:**

- PRO 1-2: 140.78 euros; PRO-3: 116.64 euros (consulta 30-5-2016).

**1.26. Precio y número de ejemplares del paquete de cuadernillos:**

- Precio por cuadernillo: 2.04 euros (consulta 30-5-2016).

**1.27. Precio y número de ejemplares del paquete de hojas de respuesta:**

- --

**1.28. Precio de la corrección y/o elaboración de informes por parte del editor:**

- --

## 2. VALORACIÓN DE LAS CARACTERÍSTICAS DEL TEST

### 2.1. Aspectos generales

Contenido	Valoración	Puntuación
2.1. Calidad de los materiales del test	Buena	4
2.2. Calidad de la documentación aportada	Buena	4
2.3. Fundamentación teórica	Excelente	5
2.4. Adaptación del test	--	--
2.5. Desarrollo de los ítems del test	Buena	4
2.6. Calidad de las instrucciones para el participante	Buena	4
2.7. Calidad de las instrucciones (administración, puntuación, interpretación)	Buena	4
2.8. Facilidad para registrar las respuestas	Buena	4
2.9. Bibliografía del manual	Buena	4
2.10. Datos sobre el análisis de los ítems	Buena	4

### 2.11. Validez

#### 2.11.1. Evidencias de validez de contenido

Contenido	Valoración	Puntuación
2.11.1.1. Calidad de la representación del contenido o dominio	Buena	4
2.11.1.2. Consultas a expertos	-	-

## 2.11.2. Evidencias de validez basadas en la relación entre las puntuaciones del test y otras variables

### 2.11.2.1. Evidencias de validez basadas en la relación con otras variables

Contenido	Valoración	Puntuación
2.11.2.1. Diseños empleados	-	
2.11.2.1.1. Tamaño de las muestras	-	-
2.11.2.1.2. Procedimiento de selección de las muestras		-
2.11.2.1.3. Calidad de los tests empleados como criterio o marcador	-	-
2.11.2.1.4. Promedio de las correlaciones con otros tests que miden constructos similares	-	-
2.11.2.1.5. Promedio de las correlaciones con otros tests que miden constructos no relacionados	-	-
2.11.2.1.6. Resultados de la matriz multirrasgo-multimétodo	-	-
2.11.2.1.7. Resultados de las diferencias intergrupo	-	-

### 2.11.2.2. Evidencias de validez basadas en la relación entre las puntuaciones del test y un criterio

Contenido	Valoración	Puntuación
2.11.2.2.1. Criterios empleados	Profesor de lengua sobre el nivel de aprendizaje en este de lengua	
2.11.2.2.2. Calidad de los criterios empleados	Adecuada	3
2.11.2.2.3. Relación temporal entre test y criterio	Concurrente	
2.11.2.2.4. Tamaño de las muestras	Varios estudios con muestras grandes o estudios meta-analíticos apropiados	5
2.11.2.2.5. Procedimiento de selección de las muestras	Incidental	
2.11.2.2.6. Promedio de las correlaciones del test con los criterios	Buena	3



### 2.11.3. Evidencias de validez basadas en la estructura interna

Contenido	Valoración	Puntuación
2.11.3.1. Resultados del análisis factorial	Adecuada	3
2.11.3.2. Funcionamiento diferencial de los ítems	-	-

### 2.11.4. Acomodaciones en la administración del test

Contenido	Valoración	Puntuación
2.11.4. El manual del test informa sobre las acomodaciones en la administración del test	No	

## Comentarios generales sobre evidencias de validez

En la evidencia de validez de contenido no se especifica claramente el procedimiento seguido, si bien es de destacar que la selección de los dominios de la prueba está basada en la revisión de los currículos estatales (que se asume han sido suficientemente discutidos y evaluados por expertos previamente), hace falta una revisión por expertos de los contenidos del test y cómo éstos se articulan para evaluar el rendimiento ortográfico siguiendo los contenidos de los currículos estatales, así como también hace falta una evaluación de los niveles de dificultad de los ítems, con los correspondientes índices de acuerdo.

Con respecto a la evidencia de validez de criterio, aunque se presentan los índices de correlación de la valoración del profesor con respecto al rendimiento en Lengua y los resultados del test de los alumnos en los distintos niveles educativos, no se presenta información suficiente sobre la calidad de la escala empleada por los profesores para hacer su valoración y tampoco se encuentra mínimamente una definición de los niveles en que clasificaron el rendimiento de los alumnos.

En el manual técnico del test se presenta alguna evidencia en cuanto a la validez de contenido, de criterio y de estructura del test. Sin embargo, sería interesante la presentación más detallada de los resultados o análisis, de tal manera que se permita al lector tener a mano la evidencia de las afirmaciones que se hacen con respecto a la validez.

Finalmente, en cuanto a la validez de la estructura del test, hace falta un resumen de los resultados del análisis factorial. Se limita por una parte a evaluar las correlaciones entre pruebas y a factorizarlas mediante componentes principales y, por otra, a correlacionar las

puntuaciones con una valoración de los profesores. Sin embargo, al ser una prueba de tipo muy objetivo, la estructura factorial es clara, y las correlaciones con el criterio, aún con su fuerte componente de subjetividad y potencial error, son buenas. En definitiva, el instrumento cumple los mínimos pero se podrían mejorar en futuras ediciones algunos aspectos.

## 2.12. Fiabilidad

Contenido	Valoración	Puntuación
2.12.1. Datos aportados sobre fiabilidad	Un único coeficiente de fiabilidad y dos mitades	

### 2.12.2. Equivalencia formas paralelas

Contenido	Valoración	Puntuación
2.12.2.1. Tamaño de las muestras	-	-
2.12.2.2. Puesta a prueba de los supuestos de paralelismo	-	-
2.12.2.3. Promedio de coeficientes de equivalencia	-	-

### 2.12.3. Consistencia interna

Contenido	Valoración	Puntuación
2.12.3.1. Tamaño de las muestras	Un estudio con una muestra grande	3
2.11.3.2. Coeficientes de consistencia interna presentados	Coeficiente alfa o KR-20	
2.11.3.3. Promedio de los coeficientes de consistencia	Buena	4

### 2.12.4. Estabilidad (test-retest)

Contenido	Valoración	Puntuación
2.12.4.1. Tamaño de las muestras	-	-
2.12.4.2. Coeficientes de estabilidad	-	-

### 2.12.5. Cuantificación de la precisión mediante TRI

Contenido	Valoración	Puntuación
2.12.5.1. Tamaño de las muestras	-	-
2.12.5.2. Coeficientes proporcionados	-	-
2.12.5.3. Tamaño de los coeficientes	-	-

### 2.12.6. Fiabilidad inter-jueces

Contenido	Valoración	Puntuación
2.12.6.1. Tipos de coeficientes presentados	-	-
2.12.6.2. Promedio de los coeficientes	-	-

### Comentarios generales sobre fiabilidad

Se ha evaluado tan sólo la fiabilidad como consistencia mediante el coeficiente alfa. Las muestras ya se han descrito anteriormente: el muestreo es incidental, pero pueden considerarse apropiadas. En el documento se menciona que se realizaron estudios de fiabilidad de consistencia interna y por el método de las dos mitades, sin embargo, del último no se presenta resultado alguno.

En cuanto a los resultados de la fiabilidad de consistencia interna y en términos generales, la fiabilidad se encontró en un buen nivel. Los autores calcularon la fiabilidad de la prueba global: PRO1 obtuvo un alfa de 0,937 y PRO2 uno de 0,945 que sugieren un nivel excelente en cuanto a la fiabilidad. En el PR-3 de 0,95. En la medida en que se van subdividiendo los niveles de educación y los contenidos de la prueba, se encuentran índices de fiabilidad que tienden más al nivel entre 0,80 y 0,85. El procesamiento semántico fue el componente con una consistencia interna por debajo de 0,80. En cuanto a los subcomponentes de la prueba, la escritura de pseudopalabras del componente léxico-ortográfico fue el que obtuvo el índice de fiabilidad más bajo, oscilando entre 0,438 – 0,562. Aunque el análisis de fiabilidad es claramente mejorable, los resultados son muy buenos dada la naturaleza de las variables a medir: tanto el constructo como los indicadores como la variable de respuesta (nº de errores) son muy objetivos.

### 2.13. Baremación e interpretación de las puntuaciones

#### 2.13.1. Interpretación normativa de las puntuaciones

Contenido	Valoración	Puntuación
2.13.1.1. Calidad de las normas	Varios baremos dirigidos a diversos estratos poblacionales	4
2.13.1.2. Tamaño de las muestras	Moderado	3
2.13.1.3. Aplicación de estrategia de tipificación continua	No	
2.12.1.4. Procedimiento de selección de las muestras	Incidental	
2.12.1.5. Actualización de baremos	Excelente	5

### 2.13.2. Interpretación referida a criterio

Contenido	Valoración	Puntuación
2.13.2.1. Adecuación del establecimiento de los puntos de corte	-	-
2.13.2.2. Procedimiento empleado para fijar los puntos de corte	-	-
2.13.2.3. Procedimiento de obtención del acuerdo inter-jueces	-	-
2.13.2.4. Valor del coeficiente de acuerdo inter-jueces	-	-

## Comentarios generales sobre baremación e interpretación de las puntuaciones

Se trata de una baremación normativa estándar. Proporciona tanto transformaciones lineales (T) como no lineales (centiles) y las tablas se establecen por cursos escolares. El grupo normativo sobre el que se elaboran las tablas puede considerarse apropiado.