

**EVALUACIÓN DEL CESPRO**  
**BATERÍA PARA LA EVALUACIÓN DE LAS**  
**ESTRUCTURAS SINTÁCTICO-SEMÁNTICAS**  
**QUE COMPONEN LOS ENUNCIADOS DE**  
**LOS PROBLEMAS MATEMÁTICOS Y DE LA**  
**UTILIZACIÓN DE ESTRATEGÍAS**  
**ALGORÍTMICAS PARA SU RESOLUCIÓN**

## RESUMEN DE LA VALORACIÓN DEL TEST

### Descripción general

Características	Descripción
<b>Nombre del test</b>	CESPRO-Batería para la evaluación de la comprensión de las estructuras sintáctico-semánticas que componen los enunciados de los problemas matemáticas y de la utilización de estrategias algorítmicas para su resolución
<b>Autor</b>	Manuel Trallero Sanz, José Luis Galve Manzano, María Rosario Martínez Arias, Cristina Trallero de Lucas, Alejandro Segundo Dioses Chocano, Luis Fidel Abregú Tueros, Carmen Isabel Inca Maldonado
<b>Autor de la adaptación española</b>	-
<b>Variable(s)</b>	Resolución de problemas, Cálculo y Numeración
<b>Área de aplicación</b>	Psicología educativa
<b>Soporte</b>	Papel y Lápiz, Informatizado

### Valoración general

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

## Comentarios generales

La prueba denominada “Batería para la evaluación de la comprensión de las estructuras sintáctico-semánticas que componen los enunciados de los problemas matemáticas y de la utilización de estrategias algorítmicas para su resolución” pretende proporcionar una herramienta válida y fiable para la evaluación de la competencia en resolución de problemas matemáticos, de numeración y cálculo, tal y como indican los autores. Para ello, la batería consta de tres factores:

1. Resolución de problemas: Evalúa la capacidad de plantear y resolver problemas matemáticos de distintos niveles de dificultad. Evaluando 3 aspectos específicos de éstos: (a) Comprensión del enunciado; (b) Resolución de las operaciones; (c) Elección de la solución correcta.

2. Cálculo: Evalúa la realización de operaciones de adición, sustracción, multiplicación y división. También si es capaz de identificar los errores más frecuentes que se producen en la realización de estas operaciones.

3. Numeración: Evalúa la comprensión, interpretación y uso del significado de los números y su escritura (cardinalidad, ordinalidad, números romanos, mayor, menor, mitad, decimales, etc.)

En cuanto a las evidencias de validez de la prueba se proporcionan tanto de contenido, interna y externa (relaciones con otras variables). En cuanto a la validez interna los datos presentados indican la unidimensionalidad de las diferentes pruebas que conforman el CESPRO. En cuanto a evidencias de validez externa, se presentan datos de correlaciones de las puntuaciones del CESPRO con la valoración del profesor en el rendimiento en matemáticas, una prueba de aptitudes intelectuales generales (BADyG-r) y con la batería INVE de Inteligencia Verbal. En cuanto al rendimiento en matemáticas el promedio de las correlaciones fue de 0,51, y se han utilizado muestras muy grandes tanto españolas como peruanas de diversos centros y que muestrean toda la geografía nacional. Por otro lado, las correlaciones entre el CESPRO y los factores del BADyG-r y el INVE se realizan con muestras de entre 42 y 84. El promedio de las correlaciones de las diferentes puntuaciones y versiones del CESPRO y la puntuación de CI obtenida con el cuestionario BADyG-r fue de 0,51. El promedio de las correlaciones de las diferentes puntuaciones y versiones del CESPRO con las puntuaciones en INVE-2 Inteligencia Verbal, Comprensión Semántica y Comprensión Lógica-Verbal fue de 0,57.

En cuanto a las evidencias de fiabilidad se proporcionan diferentes coeficientes de fiabilidad, como consistencia interna, tales como alfa de Cronbach, coeficiente de Spearman-Brown y dos mitades de Guttman. Estos coeficientes en la mayoría de las pruebas son superiores a 0,80. Por otro lado, se realiza análisis de ítems basado en la TRI, se proporcionan

los parámetros de los ítems estimados (dificultad y discriminación), y sería interesante que se aportaran datos de la fiabilidad de la medida según este modelo.

Se valora positivamente que los baremos se presenten para los distintos niveles educativos y atendiendo a la muestra (española y peruana). Sería de interés que en el Manual se proporcionara información más detallada acerca de cómo se han establecido los criterios de interpretación de puntuación en términos de la escala de valoración cualitativa. (Muy Bajo, Bajo, Medio-bajo, Medio, Medio-alto, Alto y Muy Bajo). También sería apropiado dar orientaciones sobre qué puntuación sería considerada problemática para el alumno y, por tanto, requeriría intervención específica.

El punto fuerte del cuestionario es que permite evaluar el nivel matemático del alumno, o al menos algunos aspectos del mismo. Los ítems están graduados atendiendo al nivel educativo del alumno y a su dificultad, lo cual proporciona una herramienta valiosa que permite evaluar el nivel de competencia curricular del alumno en el ámbito matemático, así como detectar dónde se producen las mayores dificultades. Así, puede ser utilizado como una herramienta de cribado (screening) para conocer el nivel del alumno al iniciar el curso o para analizar sus conocimientos al finalizarlo. Así como para la detección de dificultades concretas en matemáticas. Quizá un aspecto a mejorar sería añadir nuevos factores que permitieran evaluar lo que el DSM-V considera “trastorno específico del aprendizaje” con dificultad en matemáticas (discalculia), lo que permitiría utilizarlo para realizar dicho diagnóstico.

La batería resulta valiosa por la escasez de este tipo de pruebas en España, proporcionando a los profesionales de la educación y la psicología una herramienta objetiva que permite evaluar el nivel matemático del alumno. Esta herramienta se proporciona en formato de lápiz y papel, pero también informatizada. La aplicación informatizada del CESPPO permite de manera eficaz poder seleccionar las pruebas que interesa aplicar en cada momento, el formato de presentación es muy cuidado y adaptado al nivel y edad del examinado, y posibilita la obtención inmediata de los resultados del test.

Además en el manual se proporcionan datos referidos a población peruana.

# ANÁLISIS DETALLADO DE LA PRUEBA

## 1. DESCRIPCIÓN GENERAL DEL TEST

### 1.1. Nombre del test:

- **CESPRO.** Batería para la evaluación de la comprensión de las estructuras sintáctico-semánticas que componen los enunciados de los problemas matemáticos y de la utilización de estrategias algorítmicas para su resolución.

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

- 

### 1.3. Autor del test original:

- Manuel Trallero Sanz, José Luis Galve Manzano, María Rosario Martínez Arias, Cristina Trallero de Lucas, Alejandro Segundo Dioses Chocano, Luis Fidel Abregú Tueros, Carmen Isabel Inca Maldonado.

### 1.4. Autor de la adaptación española:

- -

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

- Instituto de Orientación Psicológica EOS

### 1.6. Editor de la adaptación española:

- -

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

- 2016

### 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:

- 2016

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

- Aptitudes, Habilidades

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

- El CESPRO evalúa aspectos de la inteligencia numérica y aptitudes matemáticas. En concreto las facetas de resolución de problemas, cálculo numérico y numeración o familiarización con aspectos de las propiedades cuantitativas:
  1. Resolución de problemas: Evalúa la capacidad de plantear y resolver problemas matemáticos de distintos niveles de dificultad. Evaluando 3 aspectos específicos de éstos: (a) Comprensión del enunciado; (b) Resolución de las operaciones; (c) Elección de la solución correcta.

2. Cálculo: Evalúa la realización de operaciones de adición, sustracción, multiplicación y división. También si es capaz de identificar los errores más frecuentes que se producen en la realización de estas operaciones.

3. Numeración: Evalúa la comprensión, interpretación y uso del significado de los números y su escritura (cardinalidad, ordinalidad, números romanos, mayor, menor, mitad, decimales, etc.).

Cálculo y numeración son pruebas complementarias.

#### 1.12. Áreas de aplicación:

- Psicología Educativa

#### 1.13. Formato de los ítems:

- Respuesta construida, elección múltiple.

#### 1.14. Número de ítems:

- CESPRO 1:
  - Resolución de problemas: 20
  - Cálculo: 15
  - Numeración: 10
- CESPRO 2:
  - Resolución de problemas: 25
  - Cálculo: 15
  - Numeración: 15
- CESPRO 3:
  - Resolución de problemas: 25
  - Cálculo: 20
  - Numeración: 16
- CESPRO 4:
  - Resolución de problemas: 24
  - Cálculo: 20
  - Numeración: 16
- CESPRO 5:
  - Resolución de problemas: 30
  - Cálculo: 20
  - Numeración: 10
- CESPRO 6:
  - Resolución de problemas: 30
  - Cálculo: 20
  - Numeración: 10
- CESPRO 7:
  - Resolución de problemas: 35

- Cálculo: 25
- Numeración: 10

**1.15. Soporte:**

- Papel y Lápiz, Informatizado

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

- Nivel B. Orientadores, psicólogos, pedagogos, profesores de matemáticas y de pedagogía terapéutica familiarizado con el enfoque cognitivo.

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

- Alumnado desde 1º de Educación Primaria hasta 1º de Educación Secundaria Obligatoria.

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

- Existen 7 versiones del test según el curso en el que se encuentre el evaluado.
  - CESPPO 1- Primer curso/Nivel de Educación Primaria
  - CESPPO 2- Segundo curso/Nivel de Educación Primaria
  - CESPPO 3- Tercer curso/ Nivel de Educación Primaria
  - CESPPO 4- Cuarto curso/Nivel de Educación Primaria
  - CESPPO 5- Quinto curso/Nivel de Educación Primaria
  - CESPPO 6- Sexto curso/Nivel de Educación Primaria
  - CESPPO 7- Primer Curso/Nivel de Educación Secundaria
- Estas versiones se pueden aplicar tanto en papel y lápiz como en aplicación informatizada con corrección automática (no se especifica hardware o software específico).

**1.19. Procedimiento de corrección:**

- Manual, Automatizada por ordenador

**1.20. Puntuaciones:**

- Para la obtención de las puntuaciones directas en **Resolución de Problemas** se contrasta la respuesta del alumno con una tabla de respuestas correctas. Se asigna un punto por acierto. Esta prueba está formada por tres puntuaciones (Comprensión del Enunciado, Resolución de Operaciones y Elección de Solución). La suma de las puntuaciones directas de estas tres puntuaciones proporciona una puntuación total en resolución de problemas.

Puntuación Directa en **Comprensión del Enunciado**: Sumatorio de aciertos de la Columna A1 y A2 de la Tabla de corrección correspondiente.

Puntuación Directa en **Resolución de Operaciones**: Sumatorio de los aciertos de la Columna B1 y B2 de la Tabla de corrección correspondiente.

Puntuación Directa en **Elección de la Solución Correcta**: Sumatorio de los aciertos de la Columna C de la Tabla de corrección correspondiente.

Puntuación Directa Global en **Resolución de Problemas**: Sumatorio de las puntuaciones obtenidas en cada una de las anteriores subpruebas.

- Puntuación Directa en **Cálculo**: se contrasta la respuesta del alumno con una tabla de valoración que ofrece puntuaciones de 2, 1 y 0 puntos, así se puntúa con 2, 1 ó 0 puntos según la respuesta dada por el alumno.
- Puntuación Directa en **Numeración**: se contrasta la respuesta del alumno con una tabla de respuestas correctas. Se asigna un punto por cada ítem contestado correctamente.
- Según el nivel de la prueba elegido (CESPRO 1, 2, 3, 4, 5, 6, ó 7) se utilizará las tablas correspondientes para la corrección de la prueba.

#### 1.21. Escalas utilizadas:

- Puntuaciones Directas, Centiles, Decatipos, puntuaciones T

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

- Según señalan los autores, se utiliza un programa informático denominado PICESPRO que permite realizar la corrección y la obtención de un informe. El informe proporciona un perfil de rendimiento de la prueba, presentando los tres factores evaluados y una breve descripción de los mismos. Existe tres modalidades u opciones de perfil, a elegir por el evaluador, con puntuaciones centiles, decatipos y puntuaciones T. Además, proporciona una valoración cualitativa (muy bajo, bajo, medio, alto y muy alto). Resulta muy visual y orientativo.

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

- En aplicación individual: Variable. Se da por finalizada cuando acabe la mayoría (el 85% del grupo evaluado). El tiempo aproximado es de dos sesiones de 50 minutos para toda la batería.
- En aplicación colectiva: Idéntico al de la aplicación individual

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

- Manual, y 7 cuadernos de aplicación (uno por cada nivel)

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

- 140,92 Euros (fecha consulta: 26 de julio de 2017)
- 31,00 euros (precio Manual)

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

- 15,60 euros por cada 10 unidades independientemente del curso (fecha consulta: 26 julio de 2017).
- 2,05 euros por 1 cuadernillo+corrección online
- 1,95 euros por 1 uso de Autoaplicación informatizada CESPRO

#### 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:

- Información no encontrada.

## 2. VALORACIÓN DE LAS CARACTERÍSTICAS DEL TEST

### 2.1. Aspectos generales

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

## 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
2.11.2.1.1. Diseños empleados	Correlaciones con otros tests, rendimiento escolar	
2.11.2.1.2. Tamaño de las muestras	Varios estudios con muestras de tamaño moderado o con alguna muestra grande y otras pequeñas	4
2.11.2.1.3. Procedimiento de selección de las muestras	Incidental	
2.11.2.1.4. Calidad de los tests empleados como criterio o marcador	Buenos	4
2.11.2.1.5. Promedio de las correlaciones con otros tests que miden constructos similares	Buena	4
2.11.2.1.6. Promedio de las correlaciones con otros tests que miden constructos no relacionados	-	-
2.11.2.1.7. Resultados de la matriz multirasgo-multimétodo	-	-
2.11.2.1.8. 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	Rendimiento escolar en el área de matemáticas (evaluado por profesores)	
2.11.2.2.2. Calidad de los criterios empleados	Bueno	4
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	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	Muy Buena	4

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

Contenido	Valoración	Puntuación
2.11.3.1. Resultados del análisis factorial	Buena	4
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

La validez del cuestionario se examina a través de la estructura interna por medio de Análisis Factorial Exploratorio de los ítems. Así, se puso a prueba el supuesto de unidimensionalidad, y aunque los resultados obtenidos con el criterio del autovalor mayor de 1 retenía tres factores, la razón de la varianza explicada por el primer factor respecto del segundo apuntó la unidimensionalidad. También se presentan las correlaciones entre las diferentes puntuaciones del CESPPO, esta información se presenta como un estudio de “validez estructural”. Las correlaciones obtenidas son elevadas, pero en el caso de las puntuaciones de las diferentes “subpruebas” cabe esperar este resultado pues forman parte de la puntuación total en Resolución de problemas.

Por otro lado, se realizan correlaciones con criterios externos: la valoración del profesor en el rendimiento en matemáticas, una prueba de aptitudes intelectuales generales (BADyG-r) y con la batería INVE de Inteligencia Verbal.

Los profesores valoran a sus alumnos con respecto a los siguientes indicadores: (a) Nivel de lectura; (b) nivel de aprendizaje global en el área de matemáticas; (c) nivel de aprendizaje en resolución de problemas aritméticos; y (d) nivel de aprendizaje de cálculo. En este caso las muestras utilizadas son de tamaño grande (N>500). Y las correlaciones se sitúan en un rango de 0,30 a 0,70, siendo el promedio de las mismas de 0,51. No se aporta información sobre la validez de las puntuaciones en Numeración, pero hay que considerar que son asimiladas desde las Baterías PAIB. Se han utilizado muestras muy grandes españolas y peruanas de diversos centros y, al menos en España, muestrean toda la geografía nacional. Las muestras han sido obtenidas tanto en centros públicos como concertados guardando una

correspondencia de cuotas por tipo de centro y localización geográfica, de forma que sean representativas de la población.

Las correlaciones entre el cuestionario y los factores del BADyG-r y el INVE se realizan con muestras de tamaño pequeño (entre 42 y 84 estudiantes). Los valores de estas correlaciones se encuentran entre 0,40 y 0,70. El promedio de las correlaciones de las diferentes puntuaciones y versiones del CESPRO y la puntuación de CI obtenida con el cuestionario BADyG-r fue de 0,51. El promedio de las correlaciones de las diferentes puntuaciones y versiones del CESPRO con las puntuaciones en INVE-2 Inteligencia Verbal, Comprensión Semántica y Comprensión Lógica-Verbal fue de 0,57.

En general, los datos de validez proporcionados son buenos.

## 2.12. Fiabilidad

Contenido	Valoración	Puntuación
<b>2.12.1. Datos aportados sobre fiabilidad</b>	Varios coeficientes de fiabilidad (para cada escala o subescala), coeficientes de fiabilidad para diferentes grupos de personas.	

### 2.12.2. Equivalencia formas paralelas

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

### 2.12.3. Consistencia interna

Contenido	Valoración	Puntuación
<b>2.12.3.1. Tamaño de las muestras</b>	Varios estudios con muestras grandes	5
<b>2.11.3.2. Coeficientes de consistencia interna presentados</b>	Coeficiente alfa, Spearman-Brown y Guttman	
<b>2.11.3.3. Promedio de los coeficientes de consistencia</b>	Excelente	5

### 2.12.4. Estabilidad (test-retest)

Contenido	Valoración	Puntuación
-----------	------------	------------

<b>2.12.4.1. Tamaño de las muestras</b>	-	-
<b>2.12.4.2. Coeficientes de estabilidad</b>	-	-

### 2.12.5. Cuantificación de la precisión mediante TRI

Contenido	Valoración	Puntuación
<b>2.12.5.1. Tamaño de las muestras</b>	Varios estudios con muestras grandes	5
<b>2.12.5.2. Coeficientes proporcionados</b>	Otros: Parámetros de dificultad y discriminación de los ítems	
<b>2.12.5.3. Tamaño de los coeficientes</b>	-	-

### 2.12.6. Fiabilidad inter-jueces

Contenido	Valoración	Puntuación
<b>2.12.6.1. Tipos de coeficientes presentados</b>	-	
<b>2.12.6.2. Promedio de los coeficientes</b>	-	-

### Comentarios generales sobre fiabilidad

Se proporcionan diferentes coeficientes de fiabilidad, como consistencia interna, tales como alfa de Cronbach, coeficiente de Spearman-Brown y dos mitades de Guttman. Estos coeficientes se muestran para cada uno de los factores, para cada curso y en muestra española y peruana. Estos índices en la mayoría de las pruebas son superiores a 0,80. En promedio el coeficiente alfa fue de 0,94 para la prueba de “Resolución de problemas”, de 0,90 para la de “Cálculo/Operaciones” y de 0,81 para “Numeración. El promedio total fue de 0,88. Aunque se proporciona información sobre el tamaño de la muestra en la que se han obtenido los diferentes coeficientes de fiabilidad. Pudiera entenderse que es la misma que para la muestra total. Por otro lado, se realiza análisis de ítems basado en la TRI, se proporcionan los parámetros de los ítems estimados (dificultad y discriminación), pero no se aportan datos de la fiabilidad de la medida según este modelo.

## 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 poblaciones	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

Aparecen baremos para los distintos niveles educativos y atendiendo a la muestra (española y peruana). Los baremos se encuentran al final del manual y se ofrecen en percentiles, decatipos y puntuaciones T.

En los modelos de perfiles de presentación de datos que se muestran en la Parte IV del Manual, aparece una escala de valoración cualitativa (Muy Bajo, Bajo, Medio-bajo, Medio, Medio-alto, Alto y Muy Bajo) para la interpretación de puntuaciones, estos niveles se establecieron a partir de los percentiles 20, 40, 60 y 80, pero sería de interés que en el Manual se explicara de manera más explícita.