

EVALUACIÓN DEL CUESTIONARIO

ESPERI

RESUMEN DE LA VALORACIÓN DEL TEST

Descripción general

Características	Descripción
Nombre del test	ESPERI. Cuestionario para la detección de los trastornos del comportamiento en niños y adolescentes.
Autor	María José Parellada, Javier San Sebastián y Rosario Martínez
Autor de la adaptación española	
Editor de la adaptación española	
Fecha de la última revisión del test en su adaptación española	2009
Constructo evaluado	Trastornos del comportamiento
Área de aplicación	Psicología Clínica, Educativa y Forense
Soporte	Lápiz y papel e informático

Valoración general

Características	Valoración	Puntuación
Calidad de los Materiales y documentación	Adecuado	3
Fundamentación teórica	Adecuado	3
Adaptación Española		
Análisis de los ítems	Adecuado	3
Validez de contenido	Adecuada pero con algunas carencias	2
Validez de constructo	Buena	4
Análisis del sesgo	No se aporta	-
Validez predictiva	No se aporta	-
Fiabilidad: equivalencia	No se aporta	-
Fiabilidad: consistencia interna	Buena	4
Fiabilidad: estabilidad	No se aporta	-
Baremos	Buena	4

Comentarios generales

El test ESPERI consta de cuatro cuestionarios, uno para niños de 8 a 11 años, otro para jóvenes de 12 a 17 años y otros dos que responden los padres y profesores. Su principal utilidad es la detección precoz de los trastornos del comportamiento en el aula de niños y adolescentes. El test permite la obtención de medidas en factores como Inatención-Impulsividad-Hiperactividad, Disocial, Predisocial, Oposicionismo desafiante, y la comparación de resultados procedentes de auto y heteroinformes realizados por los padres y profesores.

Entre sus puntos fuertes, señalaríamos los siguientes:

1. Es fruto del trabajo de un grupo multidisciplinar de expertos. El grupo revisó los instrumentos disponibles para la detección temprana de los trastornos comportamentales, encontró varias carencias y estableció como objetivo, dentro del proyecto ESPERI, la elaboración de cuestionarios que respondiesen a esas carencias y tuviesen las propiedades psicométricas apropiadas.
2. El Manual ofrece los resultados de estudios de validez de constructo y varios indicadores de la fiabilidad (consistencia interna) de las subescalas y escalas totales, con resultados buenos.
3. Gracias a los apreciables tamaños muestrales, el test aporta varios baremos centiles calculados sobre una amplia base y cuya gestión se facilita mediante el uso del CD que se ofrece con la edición en papel. La versión informatizada facilita considerablemente la tarea de corrección y de obtención de puntuaciones.
4. La prueba permite combinar autoinformes de los chicos evaluados con heteroinformes (de padres y/o profesores), en respuesta a la importancia de informadores múltiples para la detección precoz de los trastornos en el aula en estas poblaciones.

En el proceso de revisión se han apreciado algunas insuficiencias que convendría subsanar en las futuras ediciones del test:

1. La justificación del proceso de creación de los ítems (validez de contenido) y su posterior análisis es algo limitado y no se detalla suficientemente en el Manual. Apenas se ofrecen estudios de validación que comparen el nuevo test con variables externas a la propia prueba. Convendría añadir estudios de validez de todas las escalas, con otras variables, y completar los estudios de la estructura interna con otros análisis factoriales (confirmatorios, validación cruzada...). Los datos ofrecidos sobre la estructura interna de la versión de profesores plantean dudas. Dado que existen otras pruebas de amplia difusión y carácter gratuito, con objetivos parecidos al ESPERI, su comparación con ellas tendría mucho interés.
2. Los indicadores de los estudios de fiabilidad han sido calculados con los datos disponibles en el proceso de construcción de las escalas. Pudieran resultar algo más

- bajos al obtenerlos en otras aplicaciones del test. No se han generado formas paralelas ni realizado estudios sobre la estabilidad de las puntuaciones.
3. A efectos aplicados, una prueba que pretende ayudar en el diagnóstico de trastornos de conducta debería aportar más evidencias de su utilidad diagnóstica, predictiva o concurrente con otros datos que un profesional/usuario reconozca claramente como criterio de referencia.
 4. Desde el punto de visto sustantivo, hay un cierto sesgo biomédico tanto en la construcción como en la fundamentación teórica de las escalas, que puede tener su utilidad en la evaluación del Trastorno por déficit de atención con hiperactividad, pero que no debiera extenderse al resto de trastornos del comportamiento.

En resumen, los resultados que ofrece el manual sugieren la utilidad del test ESPERI para la detección temprana de los trastornos comportamentales en el aula de niños y adolescentes.

ANÁLISIS DETALLADO DE LA PRUEBA

1. DESCRIPCIÓN GENERAL DEL TEST

1.1. Nombre del test

ESPERI. Cuestionario para la detección de los trastornos del comportamiento en niños y adolescentes.

1.2. Nombre del test en su versión original

No procede

1.3. Autor

M^a José Parellada Redondo, Javier San Sebastián y Rosario Martínez Arias
Javier Martín Babarro y Sergio Martín López son coautores del programa informático y del correspondiente Manual Técnico de uso que se despliega al trabajar con la versión informática del test.

1.4. Autor de la adaptación española

No procede

1.5. Editor del test en su versión original

Instituto de Orientación Psicológica EOS

1.6. Editor de la adaptación española

No procede

1.7. Fecha de publicación del test original

2009

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

No procede

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

No procede

1.10. Área general de la variable que pretende medir el test

- Escalas Clínicas
- Trastornos de conducta

1.11. Breve descripción de la variable que pretende medir el test

Pretende medir los "trastornos del comportamiento en el aula", específicamente el de atención-impulsividad, el oposicionista, el predisocial y el disocial. Incluye cuatro escalas: una para niños entre 8 y 11 años, otra para los adolescentes entre 12-17 años y dos para los adultos con los que interactúan cotidianamente: padres y profesores. Estas dos últimas se hallan en proceso de construcción, por lo que en la edición revisada solo se incluyen algunos resultados preliminares. Padres y profesores han de responder en relación a su hijo o alumno evaluado.

La escala para niños de 8 a 11 años mide cuatro factores de primer orden y dos macrofactores de segundo orden. Los primeros son: 1) Inatención-Impulsividad-Hiperactividad, 2) Disocial, 3) Predisocial y 4) Oposicionismo-Negativismo-Desafiante.

En la escala para adolescentes, 12 a 17 años, los factores de primer orden son el 1) Disocial, 2) Inatención-Impulsividad, 3) Predisocial, 4) Psicopatía y 5) Hiperactividad.

La versión del cuestionario para el profesor mide tres factores: 1) Inatención-Impulsividad-Hiperactividad, 2) Oposicionismo-Negativismo Desafiante y 3) Predisocial-Disocial. En la de los padres se sugiere obtener puntuaciones en los mismos factores que se han aislado en la versión para profesores.

1.12. Áreas de aplicación

- Psicología clínica
- Psicología educativa
- Psicología forense
- Servicios sociales

1.13. Formato de los ítems

- Tipo Likert

1.14. Número de ítems

En todas las escalas se incluyen ítems que evalúan aspectos de conducta positivos, que no se utilizan en para obtener la puntuación.

Versión 8-11 años (Educación Primaria): 48 ítems. Se indica (p. 54 del Manual) que solo 40 ítems intervienen en la obtención de las puntuaciones, pero son 42 los ítems involucrados en la obtención de las puntuaciones en los 4 factores.

Versión 12-17 años (Educación Secundaria): 58 ítems. Se indica que solo 48 ítems intervienen en la obtención de las puntuaciones, si bien son de hecho 47 los ítems involucrados en la obtención de las puntuaciones en los 5 factores.

Versión padres. 65 ítems (niños hasta 11 años) y 4 ítems adicionales si el chico es mayor de 11 años. La escala incluye 4 preguntas abiertas sobre la historia escolar del evaluado. Se generan las puntuaciones en los 3 factores a partir de 59 ítems.

Versión profesores: 65 ítems (niños hasta 11 años) y 4 ítems adicionales si el chico es mayor de 11 años. La escala incluye 5 preguntas abiertas sobre la historia escolar del alumno. Se generan las puntuaciones en los 3 factores a partir de 59 ítems.

1.15. Soporte

- Papel y lápiz
- Informatizado

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

- Ninguna

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

Se aplica a niños escolarizados entre 8-11 años y 12-17 años. No tiene otras restricciones. En las versiones de padres y profesores son estos quienes responden en relación a su hijo o alumno.

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

ESPERI incluye 4 versiones distintas del cuestionario. La infantil y la de adolescentes tienen doble soporte en papel (cuadernillo) y pantalla. El cuestionario para profesores solo se administra por ordenador mediante la ejecución del programa que viene en el CD que acompaña al Manual. La versión para padres también se administra solo por ordenador. No obstante los ítems de todas las versiones se muestran en el Manual.

1.19. Procedimiento de corrección

- Manual mediante plantilla
- Automatizada por ordenador

1.20. Procedimiento de obtención de las puntuaciones directas

Las puntuaciones se obtienen sumando las de los ítems o elementos que integran cada una de las subescalas. En los factores de segundo orden se suman las puntuaciones directas de las subescalas que los forman.

1.21. Transformación de las puntuaciones

- Normalizada

1.22. Escalas utilizadas

- Centiles

1.23. Posibilidad de obtener informes automatizados

- Si

1.24. Oferta por parte del editor de un servicio para la corrección y/o elaboración informes

- No

1.25. Tiempo estimado para la aplicación del test

No se indica en el manual. En otros manuales de EOS aparece esta información en la ficha técnica, que no se ha incluido en este caso. Una duración máxima estimada podría ser 20 minutos.

1.26. Documentación aportada por el editor

- Manual
- Disketes/CD
- Cuadernillos de las 2 versiones para chicos

1.27. Precio de un juego completo de la prueba

- 125 euros (incluye 600 administraciones por ordenador).

1.28. Precio y número de ejemplares del paquete de cuadernillos

- Versiones para 8-11 y 12-17 años: 6.60 euros (10 unidades).

1.29. Precio y número de ejemplares del paquete de hojas de respuesta

1.30. Precio de la corrección y/o elaboración de informes por parte del editor

- El importe de 600 usos es 106.81 euros.

1.31. Bibliografía básica acerca del test aportada en la documentación

No se ofrece bibliografía básica sobre el test, si bien se aportan 33 referencias en la bibliografía del Manual, en las que se ha fundamentado la creación del test.

2. VALORACIÓN DE LAS CARACTERÍSTICAS DEL TEST

Aspectos generales

Contenido	Valoración	Puntuación
2.1. Calidad de los materiales del test	Adecuada	3
2.2. Calidad de la documentación aportada	Adecuada	3
2.3. Fundamentación teórica	Adecuada	3
2.4. Adaptación del test	No procede	-
2.5. Calidad de las instrucciones	Adecuada	3
2.6. Facilidad para comprender la tarea	Adecuada	3
2.7. Facilidad para registrar las respuestas	Buena	4
2.8. Calidad de los ítems (aspectos formales)	Buena	4
2.9. Datos sobre el análisis de los ítems	Adecuada	3

Validez de contenido

Contenido	Valoración	Puntuación
2.10. Calidad de la representación del contenido o dominio	Adecuada pero con algunas carencias	2
2.11. Consultas a expertos	Se ha consultado de manera informal a un pequeño número de expertos	2

Validez de constructo

Contenido	Valoración	Puntuación
2.12. Diseños empleados	<ul style="list-style-type: none"> ▪ Correlaciones con otros tests ▪ Análisis factorial exploratorio 	
2.13. Tamaño de las muestras	Varios estudios con muestras grandes	5
2.14. Procedimiento de selección de las muestras	No se aporta	
2.15. Mediana de las correlaciones del test con otros tests similares	Buena ($0.50 \leq r < 0.60$)	4
2.16. Calidad de los test empleados como criterio o marcador	Adecuada pero con algunas carencias	2
2.17. Datos sobre el sesgo de los ítems	No se aporta	-

Validez predictiva

No se aporta propiamente un estudio de validez predictiva, si bien se hacen contrastes con variables personales de los sujetos (edad, sexo...).

Contenido	Contenido	Puntuación
2.18. Diseño de selección del criterio		
2.19. Tamaño de las muestras	No se aporta	
2.20. Procedimiento de selección de las muestras	No se aporta	
2.21. Correlaciones del test con los criterios	No se aporta	

Comentarios generales sobre validez

Como estudio de validez se hace un análisis de la estructura interna, mediante análisis factoriales exploratorios, para verificar la estructura de los elementos y las escalas. También se realiza un estudio de validez convergente y discriminante, mediante la correlación de las puntuaciones en la escala (solo versión 12 a 17 años) con la prueba CBCL (Child Behavior Checklist, de Achenbach). Se confirman las correlaciones hipotetizadas entre las 5 subescalas del ESPERI y las 8 dimensiones que mide la prueba de Achenbach. Los autores también comparan los resultados de las escalas con las variables edad y sexo a fin de identificar perfiles y tipologías de severidad de los trastornos. Ofrecen además un estudio de los niveles de gravedad aplicando el análisis de conglomerados.

Convendría repensar en futuras ediciones algunos de los resultados ofrecidos: 1) Los 3 factores de la versión del profesor correlacionan entre sí más de 0.81, como reconoce el Manual. Estas elevadas correlaciones cuestionan que sean realmente 3 los factores que esta versión /del cuestionario mide. 2) Las correlaciones de los mismos rasgos entre la versión de los estudiantes y del profesor están en el rango (0.41 a 0.45), que parecen bajas como medidas de mismo rasgo con distinto método.

Fiabilidad

Contenido	Valoración	Puntuación
2.22. Datos aportados sobre fiabilidad	<ul style="list-style-type: none"> ▪ Un único coeficiente de fiabilidad ▪ Un único error típico de medida 	

Equivalencia formas paralelas

Contenido	Valoración	Puntuación
2.23. Tamaño de las muestras	No se aporta	
2.24. Coeficientes de equivalencia	No se aporta	

Consistencia interna

Contenido	Valoración	Puntuación
2.25. Tamaño de las muestras	Un estudio con una muestra grande ($N \geq 500$)	3
2.26. Mediana de los coeficientes de consistencia	Excelente ($r \geq 0.85$)	5

Estabilidad (test-retest)

Contenido	Valoración	Puntuación
2.27. Tamaño de las muestras	No se aporta	
2.28. Coeficientes de estabilidad	No se aporta	

Comentarios generales sobre fiabilidad

Los autores han utilizado como principal indicador de la fiabilidad, entendida como consistencia interna, el coeficiente alfa de Cronbach. Los valores encontrados son altos y son mayores en los heteroinformes (versiones de padres y profesores) que en los autoinformes (las dos versiones para chicos). En la escala para profesores, los coeficientes alfa de las subescalas están por encima de 0.95 (incluso en la escala Oposicionismo de solo 10 ítems), lo que plantea que puede haber ítems redundantes. Se ofrecen, para cada una de las subescalas, otras informaciones de la fiabilidad, como el intervalo de confianza de alfa, la mediana de los índices de discriminación de los ítems, la correlación media entre los ítems, y el error típico de medida. No incluyen otros procedimientos para el cálculo del coeficiente de fiabilidad, como son los indicadores de estabilidad y equivalencia.

Baremación

Contenido	Valoración	Puntuación
2.29. Calidad de las normas	Amplio rango de baremos en función de la edad, el sexo, el nivel cultural y otras características relevantes	5
2.30. Tamaño de las muestras	Moderado (300 ≤ N < 600)	3
2.31. Procedimiento de selección de las muestras	Incidental	

Comentarios generales sobre baremación

En relación a las versiones para niños y adolescentes, el Manual ofrece baremos centiles de cada subescala y de la escala total, por sexos y para los rangos de edad entre los que se han encontrado diferencias estadísticamente significativas. En relación a las versiones para los padres y profesores, se ofrece un solo baremos en cada caso, por no haberse encontrado diferencias significativas; si bien, como el Manual comenta, la ausencia de diferencias pudiera deberse a los menores tamaños muestrales en estos casos.

Los autores afirman que es conveniente considerar como “caso en riesgo” al chico que tenga un rango centil superior a 85 y marcan en negrita en las tablas las puntuaciones que tienen asociados tales rangos centiles. Dada la especial relevancia de este rango centil 85, debiera explicarse más en el Manual las razones que justifican su elección.

No se detalla en el Manual si en la obtención de los centiles se ha seguido algún proceso de ajuste a la distribución normal o se han obtenido a partir de la distribución empírica. No se menciona en el manual el papel del error típico de medida en la correcta interpretación de las puntuaciones. El tamaño de las muestras infantil, adolescentes y profesores es notable. La de padres es reducida (80 casos).