

EVALUACIÓN DEL PENSAMIENTO CRÍTICO

M^a Teresa Rodríguez*, Carlos Saiz*, Silvia F. Rivas**

* *Universidad de Salamanca*, ** *Universidad SEK de Segovia*

RESUMEN

La razón para desarrollar una prueba de pensamiento crítico (PENCRISAL) nace de la necesidad de evaluar la intervención llevada a cabo en este ámbito y de la carencia de un instrumento adecuado para tal fin. La mayoría de los tests que miden pensamiento crítico (Ennis, 2003) manifiestan una limitación importante, son herramientas cerradas que sacrifican la indagación de los mecanismos que se utilizan para responder a la comodidad de la corrección mecánica. La prueba HCTAES (Halpern, 2006) permite solventar esa imposibilidad de explorar los procesos implicados en la tarea de responder a un test. Para ello, Halpern y su grupo desarrollan un test, cuyos ítems son situaciones que describen problemas cotidianos que se deben resolver mediante respuestas abiertas y cerradas. Estos dos rasgos (situaciones-problema y respuestas abiertas) hacen que este instrumento sea muy valioso en la evaluación e indagación del pensamiento crítico. Sin embargo, en nuestro estudio del HCTAES hemos encontrado problemas serios en el diseño de las situaciones cotidianas. En él, muchas de las situaciones-problema, en unos casos, inducen varias estrategias o vías de solución, que dificultan el saber qué proceso de pensamiento se está utilizando. En otros ítems, ni siquiera se estimulan procesos de pensamiento. La prueba PENCRISAL se ha diseñado manteniendo parte del formato original del HCTAES, en lo referente a las situaciones-problema y en las respuestas abiertas. El utilizar situaciones cotidianas como ítems nos permite, por un lado, implicar más a la persona que realiza la prueba y, por otro, saber qué operación, estrategia o mecanismo está utilizando para resolverlo. Para lograr esto último, hemos diseñado situaciones que se deban resolver de una única forma, de modo que podamos anticipar qué proceso sigue la persona que la contesta. Este rasgo hace del PENCRISAL una herramienta precisa de medida en razonamiento, solución de problemas y toma de decisiones, y un utensilio poderoso de investigación en dichos mecanismos. Los primeros datos obtenidos con esta prueba apoyan sus virtudes.

Palabras Clave: Pensamiento crítico, evaluación, instrucción, razonamiento, solución de problemas y toma de decisiones.

INTRODUCCIÓN

La evaluación del pensamiento crítico se viene realizando desde hace varias décadas, pero sus dificultades y limitaciones siguen siendo las mismas que las encontradas en sus comienzos (Ennis, 2003; Halpern, 2003). El origen de estos problemas es doble: conceptual y metodológico. El primero proviene de la falta de consenso sobre lo que se entiende por pensamiento crítico, y el segundo procede de la naturaleza de los tests que mayoritariamente se han empleado. No vamos a ocuparnos del problema conceptual porque no entra dentro de los objetivos de esta comunicación. Nuestro estudio se dirigirá al segundo tipo de dificultades. La mayoría de las pruebas que se han utilizado han optado por un formato de respuesta cerrado, con el fin de reducir el tiempo de corrección. Muchos instrumentos utilizan pruebas de elección múltiple, que impiden saber realmente los pasos o procesos que se siguen para responder. Este formato, como sabemos, consta de varias alternativas entre las que se debe elegir una. Esto presenta al menos dos inconvenientes, uno que quien responde no genera solución alguna al problema planteado, y otro que las soluciones propuestas pueden elegirse sin necesidad de recurrir a proceso alguno de pensamiento.

Para evitar estas limitaciones importantes, Halpern (2006) propone un instrumento muy innovador y muy acertado, el HCTAES (*Halpern Critical Thinking Assessment Using Everyday Situations*). Con él evalúa el pensamiento crítico, utilizando ítems o problemas que sean situaciones cotidianas. La utilización de tareas de la vida diaria se ha empleado, por ejemplo, en el área de la psicología social y de la personalidad (Cantor, 1994). El planteamiento en el que se basan estas iniciativas, como la prueba que nos ocupa, es el *análisis de tareas* (ver, como muestra, Kirwan y Ainsworth, 1992). Esta metodología consiste en diseñar problemas o situaciones a los que podamos anticipar las operaciones necesarias para su resolución. El HCTAES se inspira en esta técnica, aunque no se aplica todo lo bien que se debiera, algo que mostraremos en la valoración que hemos realizado de la prueba más adelante, pero antes conviene describirla, para conocerla en sus aspectos más relevantes.

La prueba consta de 25 ítems en los que se plantean problemas cotidianos, para los que hay que proponer una solución y su justificación. Esta primera parte de la tarea define la parte abierta del test, en la que se exige que se produzcan soluciones, decisiones o reflexiones. En una segunda parte, se vuelve a presentar la misma situación, pero ahora lo que se demanda es que se elija una opción de las varias ofrecidas. En esta parte, se mantiene un formato cerrado como en la mayoría de los tests desarrollados. El test, pues, consta de dos partes: una abierta y otra cerrada.

El HCTAES posee rasgos importantes que lo diferencia de los demás: a) el incorporar una parte abierta y b) utilizar como ítems tareas cotidianas. La aportación que esto supone desde el punto de vista de la evaluación del pensamiento crítico es primordial. El obligar a los participantes a generar respuestas permite recoger información esencial sobre los procesos fundamentales de pensamiento, algo imposible de captar con un formato de respuesta cerrado. Y el emplear situaciones cotidianas como problemas, por un lado, hace que la prueba sea más interesante y, por otro, se consigue aproximar el contexto de la evaluación y el funcionamiento cotidiano; en definitiva, se refuerza la validez ecológica de la prueba.

Sin embargo, a pesar de estas importantes ventajas, encontramos algunos problemas serios en el diseño de las tareas. El primero, especialmente grave, es el hecho de que muchas situaciones-problema se pueden resolver empleando diferentes mecanismos de pensamiento. Si nos enfrentamos a un problema que se puede abordar de diferente manera no podremos saber qué pasos realmente se han seguido para solucionarlo. Si con ítems abiertos intentamos desvelar los mecanismos de pensamiento desencadenados, pero resulta que son más de uno los que se inducen, no podremos saber cuál se ha utilizado, de modo que perdemos el poder identificar las operaciones mentales activadas. Si esta deficiencia es por exceso, encontramos una segunda por defecto, esto es, observamos que algunas situaciones problema no inducen ningún mecanismo de pensamiento. Si no se necesita un mecanismo de pensamiento para resolver el problema, entonces ¿qué es lo que mide el ítem? Si no podemos responder con claridad a esta pregunta tenemos un serio problema de validez de constructo, no estamos midiendo lo que pretendemos medir, y esto debe corregirse. Estas dos limitaciones nos llevan a una tercera complicación, que se refiere a la aplicación de los criterios de corrección, que en ocasiones no son excesivamente rigurosos a la hora de puntuar las respuestas generadas. Al ser respuestas abiertas, no siempre se han recogido las posibles contestaciones formulables por los que contestan; y esto se ve aumentado al activar algunas preguntas varios procesos de pensamiento, como hemos indicado anteriormente. Estas deficiencias serias del test no invalidan, en todo caso, el diseño innovador del mismo, eso sí, obliga a perfeccionarlo, por esta razón, hemos desarrollado una prueba alternativa que corrige estas deficiencias, pero que conserva la dos aportaciones importantes del HCTAES. Provisionalmente, a esta prueba la hemos denominado PENCRISAL.

MATERIALES Y MÉTODO: TEST PENCRISAL

Esta prueba tiene su origen, como decíamos, en la necesidad de disponer de un instrumento adecuado de evaluación del pensamiento crítico, siguiendo las aportaciones importantes realizadas en la prueba de Halpern, antes comentada. Pero también nace de la necesidad de evaluar la intervención en este ámbito, llevada a cabo con los alumnos de 4º psicología de la Universidad de Salamanca. De este doble origen surge el diseño de esta nueva prueba de evaluación.

El PENCRISAL consta de 35 situaciones-problema que poseen tres características fundamentales. La primera se refiere a que cada uno de los ítems hace referencia a situaciones personales encontradas en la vida diaria. Con ello se pretende conseguir una mayor motivación para la realización de la prueba, ya que pueden ver parte de sus problemas y preocupaciones reflejadas en dichas situaciones, a la vez que aumenta la posibilidad de que transfieran más y mejor las habilidades aprendidas en un contexto académico a su vida. La segunda de estas características tiene que ver con el formato de respuesta abierto, pues nuestro deseo es que la persona que se enfrenta a cada una de las 35 tareas genere o produzca soluciones, no que elija entre varias que le pudiéramos ofrecer. Como hemos comentado anteriormente, esta característica nos ayuda a recoger gran cantidad de información acerca de los procesos de pensamiento que sustentan las respuestas a las situaciones-problemas. Y la última de las características del test, es probablemente la más importante por su implicación práctica.

Las situaciones-problemas se han planteado teniendo en cuenta la metodología del análisis de la tarea, mencionado antes. Esto es, las situaciones han sido diseñadas de forma que solo se puedan resolver si activamos la operación o estrategia adecuada. La tarea requiere un proceso bien definido para poder ser resuelta, si el que responde no lo hace siguiendo ese proceso, la respuesta que genere va a ser incorrecta. El hecho de que solo se pueda responder de una única manera es muy importante, ya que así estamos seguros de que medimos lo que realmente queremos medir, y de esta forma solucionamos el problema encontrado en el HCTAES referido a la validez de constructo.

Con el PENCRISAL evaluamos las tres habilidades fundamentales del pensamiento crítico: razonamiento, solución de problemas y toma de decisiones (Saiz y Nieto, 2002). Dentro del razonamiento, incorporamos tres habilidades diferentes: las de razonamiento deductivo, inductivo y de razonamiento práctico o informal. Los procesos de razonamiento deductivo e inductivo serían habilidades básicas y necesarias para poner en marcha el proceso más general de razonamiento práctico. Cuando solucionamos problemas, a veces razonamos, a veces decidimos, o ambas cosas. Esta habilidad, en cierto modo, necesita de las otras dos, pero éstas son capacidades tan diferenciadas que pueden valorarse individualmente, sin necesidad de integrarlas en solución de problemas.

Halpern defendía la necesidad de mantener las dos formas de corrección: preguntas abiertas y cerradas (Halpern, 2006). Tras el análisis del HCTAES esto parece innecesario, de modo que en nuestra prueba hemos considerado esas dos formas de preguntar en una sola, con el objetivo fundamental de que el tiempo utilizado para la realización de la prueba sea menor; además, la información que nos ofrecían los dos formatos de respuesta no eran muy diferentes. En cada ítem se puede obtener una puntuación de 0, 1 ó 2.

Para ilustrar los criterios de corrección vamos a utilizar una de las situaciones-problema del test:

“Si la pena de muerte es una medida preventiva, este carácter debería producir un menor número de asesinatos en los países donde existe, en comparación con aquellos donde no se aplica. Pero el porcentaje de asesinatos es el mismo en unos países y otros. Si una medida como la pena de muerte no es preventiva es innecesaria. Y los datos nos indican que no es preventiva, ya que los asesinatos no disminuyen. Por consiguiente, si estos no disminuyen, la pena de muerte es innecesaria”

¿Es correcto este razonamiento? ¿Por qué?

Si la respuesta hubiera sido que el razonamiento es correcto, porque se puede seguir un procedimiento deductivo como el citado: $P \rightarrow D, \neg D, \neg P \rightarrow I, \neg P \Rightarrow \neg D \rightarrow I$ o expresa lo mismo mediante sentido común, la puntuación sería de 2 puntos. Esto es igual a las preguntas abiertas de la prueba de Halpern. La persona que respondiera indicando que la afirmación es correcta y nos ofreciera una explicación aproximada, obtendría 1 punto, ya que nos indica que ha puesto en marcha un proceso de pensamiento que se acerca al correcto. Y si la respuesta fuera que la afirmación es incorrecta, o la explicación aportada lo fuera también, se puntuaría con 0.

El PENCRISAL se puede aplicar de dos modos. El primero, sería la aplicación de la prueba mediante lápiz y papel. Y el segundo, sería la aplicación del test a través de Internet. Hemos optado por esta última opción por las ventajas que ofrece. La realización del test por este medio permite cumplimentarla en varias ocasiones, si se desea, distribuyéndose así el trabajo de una forma cómoda y menos cansada. Por este medio también se logra que se disminuya el tiempo de realización de la prueba, algo que ayuda a que cuando se conteste a las últimas situaciones el cansancio o la desmotivación sea menor.

Los rasgos del PENCRISAL lo hacen una herramienta precisa de evaluación de tres habilidades fundamentales del pensamiento crítico como hemos señalado.

Como objetivo específico de este momento de la construcción de la prueba nos planteamos realizar una validación de constructo que nos permita verificar qué habilidades de pensamiento pueden explicar la variabilidad total que se observe en la prueba.

RESULTADOS

Se dispone de una muestra de 469 sujetos procedentes de tres universidades diferentes. Concretamente: 212 alumnos del 2º curso de Preparatoria del Instituto Tecnológico de Monterrey (México), 160 alumnos de 4º de Psicología de la USAL y 97 de 4º de Psicología de la Universidad de Málaga.

Durante el estudio exploratorio previo de los datos, se detectaron mediante el diagrama *box-plot*, 14 casos *outliers*. Debido a lo extremo de sus puntuaciones, que podrían afectar a las correlaciones, se decidió excluirlas de los análisis posteriores. De esta forma, la muestra final se ha visto reducida a 445 participantes, que son suficientes para el objetivo propuesto.

Para la evaluación de la validez de constructo se ha optado por emplear el Análisis Factorial, uno de los métodos más poderosos que existen para este fin (Kerlinger, 2002). Se ha utilizado el procedimiento de Componentes Principales del Análisis Factorial Exploratorio (AFE) para detectar la estructura de la prueba. El proceso se ha completado con la rotación ortogonal mediante el método Varimax debido a que es el procedimiento que ofrece una separación más clara de los factores. El análisis estadístico se ha llevado a cabo con el conocido paquete SPSS® versión 15.

Los datos han cumplido satisfactoriamente con la prueba KMO de adecuación muestral ($,773$). El determinante de la matriz de correlaciones ($,007$) nos permite deducir la existencia de un gran número de intercorrelaciones entre los ítems. Y la significación ($p=,000<,050$) de la prueba de Esfericidad de Bartlett nos permite rechazar la hipótesis nula de matriz de identidad. En consecuencia se cumplen satisfactoriamente las condiciones previas que legitiman el uso del AFE.

La extracción inicial determina un conjunto de factores que explica el 54,662% de la variabilidad empírica total observada. Tasa que aunque es mejorable, es cuando menos meritoria.

Debido al modelo teórico de partida, esto es, que las habilidades de pensamiento crítico las debemos analizar a diferentes niveles ya que hay habilidades que contienen otras más básicas, hemos creído idóneo analizarlas por separado.

Así, podemos ver en la tabla 1, que los ítems que correspondían a la habilidad de deducción formaban todos un único factor.

Tabla 1

	Component
	1
16. Deductivo Proposicional	,750
8. Deductivo Proposicional	,673
1. Deductivo Proposicional	,640
28. Deductivo categórico	,627
23. Deductivo categórico	,534

Extraction Method: Principal Component Analysis.
a 1 components extracted.

La escala de inducción refleja dos factores, como podemos observar en la tabla 2. Por un lado, encontramos los ítems correspondientes al razonamiento causal y, por otro lado, los pertenecientes al razonamiento analógico. Estos dos factores los hemos obtenido tras una rotación ortogonal.

Tabla 2

	Component	
	1	2
6. Inductivo Causal	,701	
2. Inductivo Causal	,674	
10. Inductivo Causal	,558	
24. Inductivo: razonamiento analógico		,893
9. Inductivo Razonamiento analógico	,398	,466

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a Rotation converged in 3 iterations.

Al analizar la escala de razonamiento crítico, volvemos a encontrar dos factores bien definidos, después de usar el método de rotación Varimax. Por una parte, los referidos a la habilidad de argumentación y, por otra, los implicados en la búsqueda de falacias (ver tabla 3).

Tabla 3

	Component	
	1	2
25. Razonamiento práctico	,724	
7. Razonamiento práctico	,712	
21. Razonamiento práctico	,692	
30. Razonamiento práctico	,595	
34. Razonamiento práctico: falacias		,727
11. Razonamiento práctico: falacias		,655
31. Razonamiento práctico: falacias	,401	,424

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a Rotation converged in 3 iterations.

En cuanto a las dos escalas restantes, es decir, la de toma de decisiones y de solución de problemas (ver tablas 4 y 5 respectivamente), encontramos un patrón común. Ambas escalas se componen de dos factores. Uno de ellos se refiere a las habilidades específicas y, el otro, a las habilidades más generales. Pero en el caso de la habilidad de toma de decisiones general creemos que sería viable agrupar el segundo factor con el tercero, ya que comparten parte de la varianza.

Tabla 4

	Component		
	1	2	3
20. TD disponibilidad	,759		
17. TD probabilidad	,714		
19. TD representatividad	,554	,462	
35. TD general		,824	
14. TD general			,799
27. TD general		,439	,674

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a Rotation converged in 5 iterations

Tabla 5

	Component	
	1	2
33. SP: MF	,811	
12. SP búsqueda de regularidades	,673	
26. SP general		,854
13. SP general		,614

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a Rotation converged in 3 iterations.

En consecuencia y a modo de conclusión final, podemos inferir a partir de nuestros resultados que la prueba demuestra suficientemente su validez de constructo en cuanto a lo que se pretende medir con ella. Los ítems correspondientes a cada habilidad miden los aspectos más relevantes en ella, puesto que discriminan claramente las diferentes capacidades, cuya medición abordamos desde esta prueba.

BIBLIOGRAFÍA

- Cantor, N. (1994). Live task problem solving: situational affordances and personal needs. *Personality and social psychology bulletin*, 20, 235-243.
- Ennis, R.H. (2003). Critical thinking assessment. En D. Fasko (Ed.), *Critical thinking and reasoning. Current research, theory, and practice* (págs. 293-313). New York: Hampton press.
- Halpern, D.F. (2003). The "How" and "Why" of critical thinking assessment. En D. Fasko (Ed.), *Critical thinking and reasoning. Current research, theory, and practice*. (págs. 355-366). New York: Hampton press.
- Halpern, D.F. (2006). *Halpern Critical Thinking Assessment Using Everyday Situations (HCTAES): Background and Scoring standards (2º Report.)* unpublished manuscript. Claremont, CA: Claremont McKenna Collage.
- Kerlinger, F.N. y Lee, H.B. (2002). *Investigación del comportamiento: métodos de investigación en ciencias sociales* (4ª Edición). México: McGraw Hill.
- Kirwan, B. y Ainsworth, L.K. (Eds., 1992). *A Guide to Task Analysis*. London: Taylor and Francis.
- Saiz, C. y Nieto, A. M. (2002). Pensamiento crítico: capacidades y desarrollo. En C. Saiz (Ed.), *Pensamiento crítico: conceptos básicos y actividades prácticas* (p. 15-19). Madrid: Pirámide.