

# ***PRUEBA DE PENSAMIENTO CRÍTICO PENCRISAL***

**\*Silvia F. Rivas<sup>1</sup>, Carlos Saiz<sup>1</sup>, y Patricia Morales<sup>2</sup>**

**Universidad de Salamanca<sup>1</sup> y Pontificia Universidad Católica de Perú<sup>2</sup>**

**EJE TEMÁTICO: evaluación del pensamiento crítico**

**COMUNICACIÓN: oral**

## **RESUMEN**

En los últimos tiempos, hay un creciente interés en la educación superior por la mejora de las competencias transversales de pensamiento crítico. El nuevo enfoque por competencias de la enseñanza universitaria le otorga un papel central en los nuevos planes de estudio. Y aunque todavía son escasos, está aumentando el desarrollo de programas que mejoren estas habilidades dentro de diferentes áreas de estudio universitaria. Pero de todos es sabido, que a la implementación de cualquier metodología le debe seguir necesariamente la evaluación de su eficacia, pues ésta es la que nos permite mejorarla y cambiarla, y esto no siempre se hace. Y para ello no debe valernos cualquier instrumento. La evaluación debe caminar de la mano de la intervención. Necesitamos sistemas de evaluación eficaces y válidos que nos permitan evaluar de la manera más precisa las habilidades de pensamiento crítico y que se ajusten a los objetivos de nuestras intervenciones. De nada nos sirve una buena iniciativa sin una adecuada herramienta de evaluación. Es en este punto donde existe un gran vacío, sobre todo en nuestro idioma. Por esta razón, nuestros esfuerzos en investigación van encaminados en llenar esta laguna en el ámbito de la evaluación del pensamiento crítico. En los últimos años hemos desarrollado una prueba de evaluación de las habilidades de pensamiento crítico (PENCRISAL), que se fundamenta en el enfoque del pensamiento crítico, que tiene como referentes más representativos a autores de diferentes disciplinas (Ennis, 1996; Facione, 2011; Halpern, 1998, 2014; Walton, 2006). Por todo ello, lo que presentamos en esta comunicación son los fundamentos y las características principales de este instrumento, que ha sido validado en población española (Rivas y Saiz, 2012), y recientemente adaptado en población universitaria peruana (Rivas, Morales y Saiz, 2014)

Palabras clave: pensamiento crítico, instrucción, evaluación y psicología educativa

## I. INTRODUCCIÓN

En el presente trabajo, vamos a intentar responder a la cuestión de si es posible evaluar el pensamiento crítico, en todas sus facetas o, en las más relevantes. Lo que expondremos a lo largo de esta comunicación será la defensa de que es posible construir un sistema de medida multidimensional. De lo primero que tenemos que ocuparnos es de las necesidades y dificultades de la propia evaluación, con el fin de proponer una que cumpla con los objetivos que se buscan y que además sea rigurosa. A continuación, debemos tratar el origen de toda evaluación, los programas de intervención, con el propósito de dejar muy claro que toda medida depende de ellos, o que sin ellos no tiene sentido. Una medida de aprendizaje, cuando no la hay, es un contrasentido obvio. Todo sistema de evaluación, dirigido a medir lo que se ha adquirido, siempre tropieza con el ingrato problema de la generalización, en especial en los procesos de pensamiento. Es pues imprescindible que nos detengamos en él, pues es el núcleo de toda evaluación de competencias intelectuales. Y finalmente, después de describir nuestra prueba, resumiremos algunos de los resultados que hemos obtenido con ella. Terminaremos señalando el estado actual de la evaluación en pensamiento crítico y lo que nuestra prueba aporta, frente a las existentes.

## II. NECESIDAD, IMPORTANCIA Y DIFICULTADES DE LA EVALUACIÓN

La necesidad e importancia de la evaluación del pensamiento crítico proviene de si social o personalmente se desea que estas competencias se mejoren. Sin esta condición, nada importa. Reflexionar de un modo eficaz exige mejorar dicha capacidad, y saber si existe dicha mejora impone cuantificar o cualificarla de algún modo. En definitiva, tener como objetivo el desarrollo de nuestras habilidades de pensamiento obliga a medirlas o valorarlas, con el fin de constatar algún cambio o progreso en esas competencias. Pero la importancia de pensar críticamente supone que esto nos hace más eficaces a la hora de lograr nuestras metas o simplemente que nos adaptemos bien a nuestro entorno. Por su parte, su necesidad asume que no pensamos tan bien como pudiéramos. Hay pues dos ideas que conviene separar, una, que pensar bien nos beneficia, y otra, que podemos hacerlo mejor. No es el lugar aquí para justificar con datos o razones, de los muchos de los que se disponen, lo necesario e importante que es pensar críticamente. Sí procede tratar las *dificultades* a la hora de evaluarlo.

Las dificultades en la evaluación del pensamiento crítico son *conceptuales y metodológicas*. Las primeras, provienen de las diferentes formas que hay de entender lo que es pensar críticamente. Para no distraernos de los objetivos de nuestro trabajo en valorar las ideas que existen, diremos solo lo que se entiende como la idea más consensuada y aceptada. *El pensamiento crítico es un proceso de búsqueda de conocimiento, a través de las habilidades de razonamiento, solución de problemas y toma de decisiones que nos permite lograr con la mayor eficacia los resultados deseados* (Saiz, 2017). Esta es una idea con la que estamos trabajando en intervención y evaluación. Esta definición intenta recoger lo común a todo proceso de pensar, buscar conocimiento (Halpern, 2014), junto con sus mecanismos fundamentales, encaminados al logro de nuestras metas.

Las dificultades metodológicas (ver Ennis, 2003) tienen su origen principalmente en el uso de pruebas con un formato de respuesta cerrado. Una prueba que solo pida a quien la responda que marce unas casillas que se dan como opciones de respuesta impide captar lo esencial del pensamiento, esto es, los procesos que llevamos a cabo para realizar una buena reflexión. Se sacrifica el conocimiento a la rapidez en la corrección. Ya se sabe que es muy costoso, en tiempo, corregir los formatos de respuesta abiertos, pero es el único modo de conseguir adentrarse en la mente de quien responde. Si realmente deseamos saber, si estamos midiendo pensamiento, debemos dar la oportunidad de que se exprese. Una marca en una casilla difícilmente puede lograr esto. Son pocas las iniciativas de evaluación en esta dirección. Una de las más recientes es la de Halpern. Esta autora ha desarrollado el test HCTA (Halpern, 2010). Esta prueba se centra precisamente en la eficacia del aprendizaje, en los procesos. En ella, se plantean problemas que se deben resolver y explicar. Se pide que se expresen los pasos que se siguen para afrontar esos problemas, que se detallen las estrategias o procesos seguidos. Esta medida, muy innovadora, junto con la que hemos desarrollado y que describiremos después, son las únicas pruebas de pensamiento crítico, enfocadas hacia los procesos de pensamiento. La razón obvia ya se ha dicho, el coste tan grande de tiempo en su corrección.

### III. EVALUACIÓN DE PROGRAMAS

Mejorar nuestra capacidad de pensar, como cualquier otra habilidad, se aprende. Y los procesos de aprendizaje necesitan evaluarse para conocer si se está consiguiendo desarrollar las competencias adiestradas. Como decíamos al principio, pensar es nuestra herramienta esencial de adaptación. Su importancia y necesidad son difíciles de cuestionar. La razón de ser, pues, de la evaluación del pensamiento crítico es cuantificar el proceso de aprendizaje, sea este reglado o no (académico o no). En nuestro caso, llevamos un tiempo desarrollando el programa de instrucción ARDESOS (Saiz y Rivas, 2011), lo que nos ha obligado a desarrollar una evaluación del mismo. Sin embargo, conviene saber que este proceso de evaluación es de naturaleza *conceptual* y *empírica*. En el presente trabajo se requiere de nosotros una mayor atención para la segunda. Ésta es de naturaleza *cuantitativa* y va dirigida a probar la *eficacia* de la instrucción, esto es, si después de la enseñanza se ha conseguido algún cambio. Es, pues, cuantitativa y se realiza *después* de que la intervención haya terminado. Dediquemos algo de tiempo a su descripción. No obstante, es necesario que una evaluación pueda, además, servir de diagnóstico sobre las competencias de pensamiento crítico, con independencia de que se deba aplicar siempre junto con una instrucción. Ahora bien, debemos de reconocer que su aplicación más común es la de evaluar la mejora del aprendizaje.

Cuando la evaluación empírica es para detectar una mejora en la enseñanza se impone un diseño singular. Lo que necesitamos son medidas antes y después del aprendizaje. La organización de la medida de evaluación que conviene realizar aquí, se denomina *diseño antes-después con grupo control* (al menos uno). Con esta metodología, conseguimos valorar si se ha producido un cambio o una mejora en el proceso de aprendizaje. A este cambio se le denomina *magnitud del efecto*. Ahora se puede apreciar mejor la importancia del diseño que hemos descrito antes. Ahora bien, si se pone en marcha algo tan costoso por esfuerzo y tiempo como es mejorar los procesos de pensamiento, deseamos que se mantengan o permanezcan el mayor tiempo posible. A esto nos referiremos como *grado de persistencia* del efecto. Pensar es una habilidad generalizable, susceptible de usar en cualquier ámbito a dominio. El poder decir que las competencias adquiridas se aplican con independencia del contexto del aprendizaje es otra de los modos de decir que la enseñanza ha funcionado (Perkins y Grotzer, 1997)

#### IV. PRUEBAS DE EVALUACIÓN

La evaluación de la eficacia del pensamiento crítico debes ser *cuantitativa* fundamentalmente, por varias razones que iremos exponiendo a lo largo de este apartado. Nuestra forma de entender la evaluación pasa por incorporar *pruebas de rendimiento específicas y pruebas estandarizadas*. De las que nos ocuparemos es de las segundas. Para ello hemos desarrollado y validado una prueba de esta naturaleza, PENCRISAL (Rivas y Saiz, 2012; Saiz y Rivas, 2008). Las razones para utilizar pruebas estandarizadas para evaluar el pensamiento crítico, además de las de rendimiento, son sobradamente conocidas, pero conviene recordarlas. Su uso nos permite *evaluar el constructo de pensamiento crítico*, nos proporciona un indicador fiable de validez, esto es, que estamos midiendo lo que deseamos y no otra cosa. Esta es la razón fundamental, la que realmente importa, el disponer de una *medida fiable y válida* de pensamiento. Hay otras razones importantes, aunque no tanto. Una de estas es que nos proporciona *indicadores de diferencias individuales*, entre personas y poblaciones. La tercera razón es que una medida de esta clase nos permite *compartir el conocimiento* con otros investigadores, al emplear medidas semejantes. Una penúltima razón es que nos posibilita el *comparar estudios* de diferentes investigaciones y poder aumentar nuestro conocimiento y capacidad de explicación. Y finalmente, con un instrumento de esta clase, podemos evaluar las competencias del pensamiento crítico y programas de instrucción.

La prueba PENCRISAL tiene su origen en el test HCTA (Halpern, 2010). Este instrumento es realmente innovador, de modo que hemos mantenido parte de sus principios en nuestras pruebas y hemos modificado otros que no resultan muy apropiados (ver Saiz y Rivas, 2008). Los principios que sustentan nuestras pruebas son: 1) *la utilización de ítems que sean situaciones cotidianas*, 2) *el uso de diferentes dominios*, 3) *un formato de respuesta abierto*, y 4) *el empleo de problemas de respuesta única*. El primero mantiene una de las aportaciones más importantes del test de Halpern. El segundo y el cuarto son la aportación de nuestra prueba, en especial, el uso de problemas de única respuesta. El formato abierto de las pruebas es la forma de permitir que los procesos de pensamiento se expresen. Finalmente, el uso de diferentes dominios nos permite valorar el grado de generalización de las habilidades.

Los ítems formados por situaciones cotidianas o situaciones problema intentan evaluar el pensamiento crítico en la vida diaria y hacer que la tarea sea más interesante, motive más. Al usar *problemas cotidianos*, intentamos simular el comportamiento diario de las personas, de modo que podamos apreciar hasta qué punto las habilidades son de uso generalizado en nuestras tareas del día a día. Por otra parte, afrontar problemas que simulen la realidad le añade un plus de motivación importante a la hora de realizar la prueba.

## Ejemplo 1

**DEDUCCIÓN****Ítem 16: Prueba de Habilidad de PC de California**

Si Alex quiere a alguien, él quiere a Bárbara. Hay muchas personas a las que Bárbara no quiere, y Alex es una de ellas. Pero todo el mundo quiere a alguien.

**Situación 1: PENCRISAL**

El entrenador personal de Juan le ha dicho que solo si entrena dos horas al día aprobará las pruebas del cuerpo de bomberos. Juan está preocupado porque, al no haber entrenado las dos horas requeridas, cree que no va a superar ese examen.

En el ejemplo 1, se muestran dos ítems, uno del test de California y otro del PENCRISAL. Se puede observar que la situación 1 de nuestra prueba plantea un problema de deducción de manera menos estereotipada y natural que el ítem 16. Aquí, además, se responde marcando una casilla, y en nuestra prueba se debe explicar si es correcta la reflexión del protagonista sobre problema cotidiano. Se busca, con las situaciones cotidianas, conseguir la mayor proximidad entre el contexto de la evaluación y el funcionamiento natural y cotidiano de las personas.

Un segundo rasgo de nuestras pruebas es el uso de *diferentes dominios* (se emplean diferentes contextos como salud, deportes, educación, ocio...) El propósito de usar diferentes dominios en la intervención y en la evaluación es poder medir la transferencia. Recordemos que esta hace referencia a la capacidad de generalización de las habilidades aprendidas en la esfera del aprendizaje a otras. La tercera característica de nuestras pruebas es el uso de un *formato de respuesta abierta*. Este rasgo tiene más importancia de lo que pueda parecer. Un formato cerrado, en realidad es solo una prueba de *comprensión* del problema que se nos plantea con el fin de discriminar entre las opciones que se nos proponen. En realidad, solo tenemos que reconocer o identificar la opción que consideramos correcta, pero solo señalar, sin más explicación. No tenemos que justificar por qué elegimos esa opción y no otra, y esto nos impide averiguar qué procesos de pensamiento se están empleando en cada situación. El formato cerrado ofrece muy poca información, y lo que es más problemático, que con lo que da, nos imposibilita para averiguar

qué mecanismo se pone en marcha con cada situación-problema. Por el contrario, un formato de respuesta abierto obliga a exteriorizar los pasos seguidos para resolver el problema, en realidad hace que la prueba sea de *producción*, no de comprensión. Obliga a generar o producir respuestas que justifiquen la forma de resolver cada situación. Cada problema cotidiano que planteamos, activa un proceso de pensamiento determinado. Y esto nos permite saber lo que más importa, cómo se resuelve ese problema, esto es, con qué proceso de pensamiento concreto. De este modo, podemos disponer de la información que importa, a saber, la forma de pensar en cada problema o situación. Esto, sencillamente, es imposible obtenerlo con un formato de respuesta cerrado.

En el ejemplo 2, podemos ver estas diferencias. En el ítem 6 del test Cornell, podemos apreciar que solo se debe marcar una opción, y que es imposible saber con esta respuesta si se está “viendo” el principio de transitividad o no. En la situación 5, la respuesta debe justificarse, de modo que podemos ver fácilmente si está resolviendo el problema usando ese principio o no. Además de decir si es correcto lo que se afirma, debe explicarlo. El formato de respuesta no sólo es una cuestión de economía (de ahorro de tiempo), es mucho más, es establecer la diferencia entre saber qué se piensa frente a los problemas o, por el contrario, adivinar o especular sobre lo que se piensa. No es nada baladí esta cuestión.

### Ejemplo 2

**R. CATEGÓRICO**

**CORNELL (Ennis y Millman, 1985): Ítem 6 Formato Rp cerrada**

El señor García propone que admitamos a todos los extranjeros que quieran entrar a nuestro amado país. Pero los extranjeros siempre han ocasionado problemas y siempre los darán. La mayoría de ellos ni siquiera pueden hablar castellano. Como todo grupo que ocasiona problemas es malo, de esto se sigue que los extranjeros son un colectivo perjudicial

**Contesta al ítem de la forma siguiente:**  
 Si la conclusión *se sigue necesariamente* de las dos afirmaciones, **marca A**  
 Si la conclusión *contradice* las afirmaciones expuestas, **marca B**  
 Si la conclusión *ni se sigue ni contradice* las afirmaciones, **marca C**

Todos los extranjeros (E) ocasionan problemas (P)	$E \rightarrow P$	<b>Respuesta: A</b>
Todo grupo que ocasiona problemas (P) es malo (M)	$P \rightarrow M$	
Todos los extranjeros (E) son malos (M)	$E \rightarrow M$	

**PENCRISAL: Situación 5 Formato Rp abierta**

La función de comprender envuelve el lenguaje, y solo los hombres lo usan, de modo que podemos concluir que solo ellos comprenden.

**¿Es correcto el razonamiento? Explica el por qué**

Comprender (C) envuelve el lenguaje (L)	$C \rightarrow L$	<b>Respuesta: Correcto</b> <b>Explicación: Principio transitividad</b>
Solo Los hombres (H) usan el lenguaje (L)	$L \rightarrow H$	
Los hombres (H) comprenden (C)	$C \rightarrow H$	

En este formato está algo tan esencial como poder saber si realmente estamos midiendo lo que pretendemos y no otra cosa. Los números pueden “tranquilizarnos”, pero no convencernos. La

estadística puede decir que existen determinadas dimensiones en una prueba, pero que esas sean de pensamiento o no, no puede, porque lo que se responde a una prueba cerrada no me permite saberlo. Son demasiado complejos nuestros mecanismos de pensamiento como para captarlos con una simple marca en una casilla.

El último rasgo de nuestras pruebas consiste en plantear situaciones-problema que solo se puedan responder de una manera, esto es, diseñar ítems que sean *problemas de respuesta única*. El test predecesor de nuestras pruebas, el HCTA, consta de situaciones cotidianas que se pueden responder de diferentes formas, no plantea un problema para cada proceso. Veámoslo en el ejemplo 3.

### Ejemplo 3

**SOLUCIÓN PROBLEMAS**

**HALPERN (CTA) : Ítem 24**

Supón que estás cuidando del perro de tu vecino y una de las tareas que tienes que hacer es darle una pastilla voluminosa y aparentemente amarga. Se trata de un perro de presa grande que mordió a un niño el año pasado. ¿Cómo te las arreglarías para darle la medicina?.

**Propón dos buenas soluciones para este problema**

**2 puntos:** Si hace uso de métodos creativos **CORRECCIÓN (0-1-2)**

Pedir ayuda, consultar experto, mezclar pastillas, etc.

**PENCRISAL: Situación 22**

Eres propietaria de un bar familiar donde trabaja tu madre en la cocina y tu hija ayudando como camarera de forma ocasional. Ahora necesitas contratar a un nuevo trabajador para poder mantener el bar abierto ya que han ingresado a tu madre en el hospital y los médicos te han dicho que la estancia en él será prolongada. Además, tu hija ha empezado las clases en una universidad que está fuera de la ciudad. Tienes que hacer una selección entre varios candidatos, pero no tienes tiempo y no quieres pedirle a tu hija que te ayude en esto pues tendría que desplazarse. Una amiga te habla de su hija que está en paro y necesita trabajar, pero tienes tus dudas, ya que no tiene experiencia laboral en la hostelería.

**¿Cuál sería la mejor forma de actuar por parte de Rocío para resolver su situación? Justifícalo**

**2 puntos:** Si aplica bien un método **CORRECCIÓN (0-1-2)**

**Identificar problema:** madre e hija no atienden el bar y necesitan alguien que trabaje en el bar

**Subproblemas o metas:**

- Hacer una selección de candidatos para elegir trabajador adecuado, le llevará tiempo y tendrá que cerrar el bar.
- Contratar a la hija de su amiga provisionalmente mientras busca un candidato adecuado

En el ítem 24, se observa que las respuestas que se dan y puntúan pueden ser de lo más variado, y algunas nada tienen que ver con procesos de pensamiento, pues pedir información o consultar a un experto, que se sepa, no requiere demasiada inferencia. Esto nos plantea un problema, pues nos deja en una situación parecida a la generada por las respuestas cerradas, esto es, que no sabemos qué está haciendo mentalmente quien responde el problema. Para solucionar esta grave limitación, en nuestras pruebas hemos incorporado la metodología del *análisis de tareas* (ver Saiz y Rivas, 2008). Con este procedimiento, conseguimos fijar qué proceso se usa para cada situación problema, de modo que quien lo resuelve bien, sabemos con certeza cómo lo ha hecho, qué operaciones mentales a empleado. La idea es diseñar problemas que sólo se puedan resolver



de una manera, con el fin de “atrapar” el mecanismo de pensamiento correspondiente. Esta técnica nos proporciona ítems altamente informativos sobre los procesos y facilita la cuantificación de los ítems, al permitir solo una forma de respuesta. En realidad, con esta técnica, al puntuar la prueba, conseguimos una precisión muy próxima a la de las respuestas cerradas, una virtud nada desdeñable.

Una vez expuestos los principios en los que se fundamentan nuestras pruebas, vamos a describirlas brevemente. En cuanto a la *validez de contenido*, nuestras pruebas constan de un conjunto de ítems que son representativos de las operaciones o procesos de pensamiento crítico, y están redactados de un modo sencillo y claro, y en un lenguaje coloquial. Nuestras pruebas están formadas por 35 *situaciones-problemas*, con formato de respuesta abierta, que no exige respuestas técnicas, que supongan alguna formación previa en lógica o estadística, por ejemplo (ver ejemplo 4).

#### Ejemplo 4

TAREA DEDUCTIVA	
Razonamiento Proposicional	
<b>Situación 20: PENCRISAL</b>	
<p>Juan necesita utilizar el transporte público todos los días para ir a trabajar y tarda aproximadamente unas dos horas. Estos últimos días, con la huelga de autobuses, ha habido problemas de tráfico, por lo que siempre ha llegado tarde. Hoy tiene una reunión muy importante y su jefe está intranquilo por si llegará a tiempo. Le pregunta a un compañero por Juan y éste le dice que no se preocupe que hoy no hay huelga, luego no tendrá problemas de tráfico, así que llegará a tiempo para la reunión.</p>	
<b>FORMULACIÓN DEL PROBLEMA</b> H: hay huelga; T: llegar tarde al trabajo $H \rightarrow T, \neg H \Rightarrow \neg T$ <b>FALACIA NEGACIÓN ANTECEDENTE (NA)</b>	<b>RESPUESTA CORRECTA</b> El razonamiento <i>no es válido</i> porque el hecho de que no haya huelga no implica que no pueda llegar tarde, pues pueden darse otras circunstancias para que se retrase.

Los procesos de pensamiento representados, lo están en los *cinco factores o dimensiones* que constituyen las pruebas, a saber: *razonamiento deductivo, inductivo, práctico, toma de decisiones y solución de problemas* (ver tabla 1). El *orden de presentación* de los ítems o situaciones cotidianas es *aleatorio*.

Tabla 1

FACTORES PENCRISAL/PENTRASAL				
DEDUCCIÓN	INDUCCIÓN	R.PRÁCTICO	TD	SP
Proposicional =4	Causal=3	Argum. =4	General=2	General=4
Categorico =3	Comprobación hipótesis =1	Falacias=3	Probabilidad=2	Regularidades=2
<b>TOT=7</b>	Analógico =2	<b>TOT=7</b>	Coste Inversión=1	Medio Fin=1
	Generalizac =1		Representación=1	<b>TOT=7</b>
	<b>TOT=7</b>		Disponibilidad=1	
			<b>TOT=7</b>	

Ejemplo 5

**PENTRASAL**

**Situación 33: Toma de decisiones: representatividad**

En un estudio realizado en dos institutos de secundaria, uno con 2000 alumnos y otro de 200, se encuentra que las calificaciones medias de los chicos y chicas son muy semejantes. Además en ambos centros hay aproximadamente el mismo número de varones que de mujeres. Sin embargo, en una evaluación, el 60 % de las chicas sacan mejores notas que los chicos ¿Dónde es más probable que suceda este resultado?

**CORRECCIÓN**

- **0 puntos:** “En ambos por igual, porque las probabilidades para ambos son las mismas independientemente del tamaño de las muestras” **Incorrecto**
- **1 punto:** “Es más fácil en el instituto pequeño, porque habrá menos chicas” **Explicación imprecisa**
- **2 puntos:** “En el pequeño, el de 200, porque la probabilidad será más real en las muestras grandes que en las pequeñas” **Rp y explicación correctas**

El *proceso de valoración* de los ítems o situaciones-problema de respuesta única es sencillo. Se puntúa cada ítem de cero a dos. Cero, para las respuestas incorrectas, uno, para las respuestas correctas con una justificación/explicación incompleta, y dos, cuando es completa (ver ejemplo 5).

La puntuación global es de 70 (2 X 35), y la de las subescalas de 14 (7 X 2). Las *características psicométricas* de las pruebas son el ser un test de ejecución máxima y de potencia. El tiempo medio para terminar la prueba es de ciento veinte minutos. La *forma de aplicación* de las pruebas puede ser *individual y colectiva, y a través de internet*. La administración colectiva permite un control de la influencia de factores externos, evita la suplantación de identidad, y estandariza la aplicación. En la aplicación individual se pierde el control de la situación, no se aplica de modo estandarizado, aunque se gana en flexibilidad y disponibilidad de quien la realiza. La aplicación informatizada ofrece la ventaja de que impide dejar ítems en blanco y se puede realizar en varias sesiones.

La forma de puntuar las pruebas busca la cuantificación más precisa; la aplicación informatizada y de manera colectiva reúne el mayor número de ventajas.

## VI. RESULTADOS Y ESTADO ACTUAL DE LA INVESTIGACIÓN

En la actualidad se ha validado la versión 3 del PENCRISAL (ver Rivas y Saiz, 2012, para la fundamentación psicométrica; ver Rivas, Morales y Saiz, 2014, para la adaptación peruana). Esta versión ha sido modificada teniendo en cuenta los primeros resultados de la aplicación de dos versiones anteriores. Los resultados de los que disponemos tienen que ver con la *medida de la eficacia* de nuestra intervención y con las *propiedades psicométricas*. Estas son razonablemente buenas. Los factores del test explican el sesenta por ciento de la varianza y la fiabilidad entra dentro de los índices aceptables. Además, al utilizar este test en un diseño antes-después, nos ha permitido corregir, modificar y desarrollar nuestra metodología de instrucción, siendo la cuarta versión, la que empleamos en la actualidad. Adicionalmente y desde hace tiempo, se está aplicando a decenas de personas, en varios países latinoamericanos, con fines de diagnóstico e investigación.

## VII. CONSIDERACIONES FINALES

A lo largo de esta comunicación, nos hemos ocupado de justificar nuestro sistema de evaluación en pensamiento crítico, en contextos cotidianos. Las dificultades que acompañan a una empresa tan compleja como esta, son enormes, por lo que cualquier propuesta debe recorrer un

periplo de mucho tiempo hasta que se perfila como estable. Nosotros llevamos trabajando ya tiempo en este sistema de evaluación. De momento, hemos conseguido disponer de un instrumento fiable y válido para medir pensamiento crítico, de manera estandarizada. El trabajo es mucho y los recursos pocos, por esta razón, vamos avanzando con más parsimonia de la deseada. Lo importante, para nosotros, es comprobar, a partir de los resultados que vamos obteniendo, que nuestras propuestas van por el buen camino, aunque lo recorramos con lentitud. Creemos que hemos tomado una de las bifurcaciones buenas, en nuestra senda por mejorar y constatar el desarrollo del pensamiento crítico.

## VI. BIBLIOGRAFÍA

- Ennis, R. H. (2003). Critical thinking assessment. En D. Fasko (Ed.), *Critical thinking and reasoning. Current research, theory, and practice* (pp. 293-313). Cresskill, NJ: Hampton Press.
- Halpern, D.F. (1998). Teaching critical thinking for transfer across domains - *Dispositions, skills, structure training, and metacognitive monitoring*. *American Psychologist*, 53 (4), 449-455.
- Halpern, D. F. (2010). *Halpern Critical Thinking Assessment*. Austria: Schuhfried, Vienna Test System.
- Halpern, D.F. (2014). *Thought and knowledge: An introduction to critical thinking* (5th edition). New York: Psychology Press.
- Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, 16 (3), 235-266.
- Perkins, D. N. (1999). ¿Qué es la comprensión? In M. S. Wiske (Ed.), *La enseñanza para la comprensión. Vinculación entre la investigación y la práctica* (pp. 69-92). Buenos Aires: Paidós.
- Perkins, D. N. y Grotzer, T. A. (1997). Teaching intelligence. *American Psychologist*, 52 (10), 1125-1133.
- Rivas, S.F. Morales, P. y Saiz, C. (2014). Propiedades psicométricas de la adaptación peruana de la prueba de pensamiento crítico PENCRISAL. *Avaliação Psicológica*, 13 (2), pág. 257-268.
- Rivas, S.F. y Saiz, C. (2012). Validación y propiedades psicométricas de la prueba de pensamiento crítico PENCRISAL. *Revista Electrónica de Metodología Aplicada*, Vol. 17 (1), 18-34.
- Saiz, C. y Rivas, S.F. (2008). Evaluación en pensamiento crítico: una propuesta para diferenciar formas de pensar. *Ergo, Nueva Época*, 22-23, 25-66.
- Saiz, C. y Rivas, S.F. (2011). Evaluation of the ARDESOS program: an initiative to improve critical thinking skills. *Journal of the Scholarship of Teaching and Learning*, Vol. 11, No. 2, 34-51.
- Schank, R. C. (2005). *Lessons in Learning, e-Learning, and Training: Perspectives and Guidance for the Enlightened Trainer*. San Francisco, CA: Pfeiffer/Wiley.