

## VALIDACION DE CONSTRUCTO EN EL AMBITO PEDAGOGICO

por JOSÉ MANUEL GARCÍA RAMOS

*Universidad Complutense de Madrid*

### *Introducción*

Todos los estudiosos de nuestro campo coinciden en estimar que el concepto de validez es central en la investigación científica. Dicho concepto se vincula a otros, tales como el concepto de contrastación efectiva, el concepto de verdad, el concepto de adecuación a la realidad, el concepto de utilidad, etc. En este sentido se habla de validez con referencia a *conceptos, modelos, hipótesis, diseños experimentales, medidas, juicios o valoraciones, evaluaciones, procesos*, etc. (Bartolomé, 1983).

Lógicamente, la validación en todos los contextos anteriores exige la observación, la medición, por lo que suele decirse que el avance científico en cualquier disciplina se encuentra vinculado estrechamente con el avance en la elaboración y posibilidades de utilización de instrumentos adecuados y útiles a la medición de los fenómenos estudiados.

En general, validez de un método es la exactitud con que pueden hacerse medidas significativas y adecuadas con él, en el sentido de que midan realmente los rasgos o variables que pretenden medir. Esto es, «un test o instrumento de medición cuya fiabilidad es conocida será válido, si mide aquello para lo que fue construido» (Magnuson, 1966) (p. 153). Es evidente que cuando queremos estimar la validez de un instrumento de medida, necesitamos saber qué rasgo deseamos que mida. A este rasgo suele denominársele *variable de criterio*. Nos interesa saber en qué grado se corresponden las posiciones de los individuos en la distribución de las puntuaciones obtenidas en el test a validar y los datos que usamos, con un grado conocido de certeza, como índices para las puntuaciones de los individuos en la variable criterio.

De lo anterior, parece deducirse la gran importancia del estudio de la validez, pero aún hay algunas razones más que justifican y exigen, de forma más clara, los procesos de validación. En la mayoría de las mediciones de variables psicopedagógicas, «la 'distancia' entre el objeto o variable a medir y el instrumento, es lo suficientemente grande como para que no sea evidente, ni siquiera fácilmente demostrable, la relación entre ambos» (Pérez Juste, 1981). Es decir, cuando el fenómeno a medir es muy complejo o bien las variables a medir son de tipo organísmico, la tarea de construir buenos operadores de conducta adecuados a tales variables y que representen suficientemente bien lo que pretenden o dicen medir se va haciendo más y más complicada. Como afirma Bartolomé (1983), «si la relación entre el instrumento y las variables es evidente, el tema de la validez carece de importancia. Es justamente cuando tal evidencia no aparece cuando debemos investigar profundamente el proceso de validación» (p. 3).

El tema de la validez, como vemos, es ciertamente muy complejo. Por otra parte, son muchas las dificultades con que nos encontramos cuando queremos estudiar cualquier rasgo o variable. Vamos a comentar brevemente algunos de estos problemas ya implícitos en lo anteriormente tratado. Un primer problema, es el que se deriva del carácter indirecto de la medición, ya que nunca medimos un rasgo sino algunas de sus manifestaciones conductuales, las cuales, suponemos, nos permiten estimar el grado en que un individuo «posee» un rasgo o variable. En este sentido nunca podrá decirse que un instrumento es válido o no, sino en qué grado es válido para medir lo que pretende medir. Una restricción más, cuando pretendemos asignar un grado de validez a un instrumento, es la población en que dicha validez es estimada, dado que un instrumento podrá ser válido (en un cierto grado) para una determinada población de sujetos y no ser válido (en el mismo grado) para otra población de sujetos distinta.

De las restricciones anteriores parece derivarse claramente que la validez no es un concepto único, sino que adopta muy variadas formas e interpretaciones dependiendo de los diversos factores de restricción ya comentados. Por otra parte, el énfasis hasta el momento lo estamos poniendo en los tipos de validez tradicionales en que supuestamente es posible obtener un índice del grado de validez de un instrumento, para lo cual es preciso contar con una medida de las variables de criterio. Pero ¿qué ocurre cuando en estos tipos de validez (predictiva y concurrente fundamentalmente) los criterios finales exigen medidas que habrán de recogerse tras un lapso de tiempo considerable? (con las dificultades «administrativas» consiguientes), o ¿cuando ni siquiera es posible contar con un «criterio intermedio»? o ¿cuando la propia relación con un criterio supuesto es difícil de establecer o bien, no es fácil la definición de un universo de contenido determinado? (que nos permita estimar la validez de contenido del instrumento).

Muchas definiciones se han dado del concepto de «validez», sobre todo por entenderla diferencialmente en función del tipo concreto de

validez de que hablemos. Y también, son diferentes estos tipos de variables según la terminología usada por los autores que la definen; las más comúnmente aceptadas son: validez aparente, validez de contenido, validez predictiva, validez concurrente, validez de constructo y validez de utilidad; si revisamos la bibliografía sobre el tema, veremos que otros adjetivos se unen al sustantivo «*validez*» (tales como: incremental, convergente, discriminante, factorial, etc.), pero la mayoría de estos adjetivos quieren significar algunos de los aspectos esenciales que cada autor señala dentro de la validez de constructo. En este trabajo no vamos a definir todos los tipos de validez, sino que vamos a tratar el tema específico de la validez de constructo (también denominada de otras formas: de construcción, de hipótesis de trabajo, diferenciadora, etcétera), sin duda alguna la más importante desde el punto de vista científico, ya que va a la raíz del problema, al tratar de estudiar la naturaleza de la variable objeto de medida.

En todo caso, las principales cuestiones a que pretende responder el estudio de la validez podrían sintetizarse de la forma en que lo hacen Martínez Arias (1981) y Cronbach (1971) [1]:

- *Adecuación* con lo que mide la conducta bajo estudio.
- Capacidad de *diagnosticar* la conducta real de una persona.
- Sensibilidad para detectar un objetivo.
- Seguridad en las *decisiones*.
- Utilidad.

En esta línea, algunas de las conclusiones que pueden extraerse de todo lo dicho anteriormente son las siguientes:

1. La validez no es un concepto único; adopta diversas y variadas formas e interpretaciones según sea la variable a validar, los objetivos del instrumento de medida, la población a que se aplique dicho instrumento, etc.

2. Lo que se valida no es estrictamente el instrumento, sino la interpretación de los datos obtenidos por un procedimiento específico (Cronbach, 1971) (Magunson, 1966).

3. La validez no debe entenderse como el cálculo de un simple índice que indique la relación entre test y criterio, sino más bien como un *proceso continuo* (Nunnally, 1978) que requiere la acumulación de datos procedentes de distintas investigaciones empíricas y de procesos lógicos.

4. La validez no puede determinarse en términos absolutos respecto de ninguna variable de nuestro campo de estudio, sino que es una cuestión *de grado*.

5. La elección de uno u otro tipo de validez o de varios tipos a la vez dependerán de las funciones típicas u objetivos que cumplirá el instrumento de medición particular que sea estudiado.

### 1. *Qué se entiende por validez de constructo. Problemática específica*

Con el desarrollo de las técnicas orientadas a evaluar rasgos básicos de la personalidad, muchos de los cuales eran considerados estados inconscientes, llegó a ser obvio que los procedimientos que se venían utilizando (predictiva, concurrente, aparente, etc.) eran inadecuados debido a los problemas de criterio que presentaban. El problema era sin duda, difícil al comprobarse que los rasgos no se relacionaban con la conducta abierta de una manera simple y directa. Algunos psicólogos se inclinaron por abandonar los intentos de validar estos instrumentos y se contentaron con «la validez de fe». No obstante, el abandono del rigor científico era tan evidente, que no podía admitirse tal solución. En un esfuerzo por salvar este obstáculo, se propuso el concepto de validez de constructo en las «recomendaciones técnicas» de la A.P.A. (1954). Este tipo de validación se propuso para una situación en la que «... el examinador no tenga una medida definitiva de criterio ante la cualidad que le interesa y deba usar medidas indirectas para validar la teoría» (p. 214).

Como vemos, este nuevo concepto originó —como señala Megargee (1971)— una distinción fundamental entre *la pertinencia* y *la equivalencia de la conducta*. Es decir, si es imposible encontrar una medida de criterio equivalente al rasgo, el investigador puede recurrir a otras conductas observables que, aunque no son equivalentes al rasgo, están relacionadas con él. De lo que se trata, pues, es de ir verificando hipótesis sobre relaciones de las variables con otras variables (emanadas de la teoría), para así ir teniendo una confianza creciente en el instrumento, a la vez que se trata también de ir validando la red de relaciones teóricas que dicho rasgo estableció.

Para Cronbach y Meehl (1967), la validación de constructo es importante no sólo en variables de personalidad, sino a veces para cualquier clase de constructo; de aptitud, de rendimiento, de intereses, de actitudes, etc. La validez de constructo interviene en la respuesta a preguntas como: ¿hasta qué punto, determinado test de inteligencia está libre de instrucción?, ¿ese test de «interpretación de datos» mide la habilidad lectora, el razonamiento cuantitativo o las respuestas de grupo?, ¿cómo difiere una persona en el Strong Accountant con la puntuación A y con la puntuación B en el Strong C.P.A., de una persona que tiene puntuaciones inversas? (pp. 243-270). O también: ¿qué factores explican los resultados de un instrumento de evaluación?, ¿las medidas resultantes son consistentes con las pronosticadas sobre la base del constructo?; ¿hay otros constructos que pueden explicar el resultado observado?, etc. (Martínez Arias, 1981, p. 184).

Cronbach y Meehl (1967) definen un constructo como «algún atributo postulado acerca de la persona, que supuestamente se refleja en la ejecución de un test». Presumimos con ello que una persona posea o no posea en cierto momento, un atributo cualitativo o estructural, o un grado de atributo cuantitativo. Un constructo es, pues, un concepto, una abstracción y en este sentido, el constructo de personalidad mismo, des-

cansa en la creencia de que las consistencias conductuales de los individuos, existen con amplitud y justifican gran parte de la varianza de su conducta. Ahora bien, «los constructos siempre se hallan insertos en una teoría científica, es decir, para ser válidos necesitan anclarse en un sistema» (Bunge, 1976).

Para Cronbach y Meehl (1967), la lógica de la validación de constructo se invoca «cuando el constructo está altamente sistematizado o indeterminado, se usa en una teoría ramificada o en unas pocas proposiciones sencillas o interviene en proposiciones absolutas y declaraciones de probabilidad».

Siguiendo el ejemplo del constructo de personalidad, es preciso aclarar que, como afirma Mischel (1977), «las consistencias de la conducta empíricamente establecidas, no parecen suficientemente grandes como para justificar las creencias de que los rasgos de la personalidad están muy ampliamente generalizados» (p. 56). Es decir, la validez de constructo —el único tipo de validez completa, al decir de Loevinger, 1957— tiene como propósito fundamental, *validar la teoría subyacente al sistema de evaluación o medida*, sea ésta una teoría de personalidad, o una teoría conductual y *a la medida misma* [2].

Ciertamente, los constructos son los que determinan qué conductas han de ser relacionadas para su observación y la situación en que será aplicada. Un instrumento de medida estará, pues, ligado al sistema de constructos dentro del que fue construido y es *este tipo de validez* el que nos indicará *el grado en que el instrumento de evaluación, es una medida adecuada del constructo y en qué medida las hipótesis derivadas del constructo pueden confirmarse mediante la utilización del instrumento*. En este sentido Fiske (1971) afirma que *el investigador debe describir «la única cualidad a la cual su constructo se refiere»*, es decir, debe establecer el «corazón» del constructo, en semejanza a un camino que integre todas las manifestaciones que el conceptualizador desee incluir bajo la variable (p. 98). Dicho de otra manera, en el momento de la construcción del instrumento de evaluación, «es necesario establecer las especificaciones de los objetivos del mismo y el dominio conductual al que han de referirse los elementos del instrumento de evaluación» (Martínez Arias, 1981, p. 285).

Hoy en día, se acepta la validez de constructo (con sus múltiples técnicas y procedimientos) como un camino fundamental para la validación de instrumentos de medida, a la par que la teoría que subyace bajo él mismo; no obstante, algunos autores como Bechtoldt (1959) se oponen a la idea de validar tests por medio de este tipo de validación, al sostener que esto es poco más que un nuevo nombre del proceso básico de la construcción de la teoría y tiene el peligro, no sólo de introducir una confusión indeseable, sino más grave aún, «nos aleja de una adhesión estricta a la definición operacional», que considera esencial para el progreso científico de la psicología. En todo caso, esta opinión hoy en día no es aceptada y de hecho es una práctica común la utilización de la validez de constructo en la validación de instrumentos de medida.

El mismo Campbell (1960), publicó un análisis ulterior de la validación de constructo y una crítica de la posición de Bechtold, en la que «niega que la validación de constructo representa el abandono del operacionismo o estimule la cosificación de los rasgos» (Megargee, 1971, p. 103).

No obstante, los constructos, como interpretaciones de los tests —según Cronbach y Meehl, 1967— no son indudablemente psicológicos, y así subrayan que «la determinación, de si una interpretación de las propiedades o relaciones de un test implica cuestiones de validez de constructo, se decide al examinar el conjunto de la evidencia ofrecida, junto con la que se sostiene acerca del test en el mismo contexto». A este respecto, remitimos a la discusión de Cronbach y Meehl (1967, pp. 110-112), sobre la inadecuación de la validación en términos de criterios específicos en la cual se enfrentan las posiciones de Anastasi y Cronbach sobre si los tests pueden validarse, sin una medida de un criterio definido específicamente.

En todo caso, coincidimos con Martínez Arias (1981), quien afirma que cuando especificamos los objetivos del instrumento de medida en la primera fase de la validación de un constructo —lo que Fiske (1971) llamaría el «corazón» del constructo— hacemos realmente una elección entre diferentes formas de operativizar el dominio de conductas en las que estamos interesados, lo cual es una tarea similar a lo que Kerlinger (1973) denomina «definición operativa del constructo». Por otra parte, y como veremos más adelante, los procedimientos utilizados en la validación de constructos no se agotan en el análisis correlacional, como ocurre con la validez de un test respecto de un criterio externo. En todo caso, coincidimos con Nunnally (1978), quien afirma que los científicos no pueden trabajar sin constructos; en el fondo, se puede afirmar que la mayoría de las variables que medimos y en las cuales estamos interesados en nuestro campo, son constructos aún cuando las «distancias» entre dichas variables o constructos y sus definiciones operacionales respectivas puedan variar en una amplia escala.

En resumen, la mayoría de los autores (Cronbach, 1967; Campbell, 1960; Meehl, 1967; Martínez Arias, 1981; Fiske, 1971..., etc.) creen un imperativo el que los investigadores de las ciencias sociales conceden un lugar en su pensamiento metodológico a la validación de constructo, de manera que —como afirman Cronbach y Meehl (1967)— «su fundamento, legitimidad y sus peligros lleguen a ser explícitos, familiares, en lugar de la esparcida tendencia a ocuparse de lo que no es sino investigación práctica de validación de constructos y de usar los constructos en la práctica de los tests, mientras se habla de una metodología «operacional» que, de ser adoptada, violentaría la investigación al meterla en un molde al que no se ajusta» (p. 130). Parece claro que, para estos autores, determinados constructos que no admiten una metodología operacional, no deberían ser violentados por la investigación con tal soporte metodológico, pues podrían suponer un reduccionismo del propio constructo.

2. Aspectos importantes en la validación de constructo

De lo dicho anteriormente se deduce claramente que este tipo de validación requiere de todo el soporte metodológico y útil para llevar a cabo sus objetivos. Parece claro que no hay un tipo de técnica de validación concreta que responda a un solo tipo de procedimiento de validación. De hecho, son diversos y muy complejos, en ocasiones, los procedimientos utilizados para tal fin, llegándose incluso a crear nuevos soportes metodológicos específicos (técnicas) para validar algún aspecto de la red de leyes lógicas planteadas en la definición de un constructo.

Siguiendo a Nunnally (1978), tres son los aspectos más importantes a tener en cuenta en la validación de constructos:

1. Especificar el dominio de observables relativos al constructo.
2. A partir de investigación empírica y análisis estadístico, determinar hasta qué punto los observables tienden a medir la misma cosa, algunas o muchas cosas diversas.
3. Llevar a cabo estudios diferenciales y/o experimentos controlados para determinar hasta qué punto las medidas supuestas del constructo producen resultados que son predecibles a partir de hipótesis teóricas altamente aceptadas, concernientes al constructo.

Indudablemente, suele ser necesario evaluar la validez de constructo por medio de la integración de fuentes diversas (ver tabla 1).

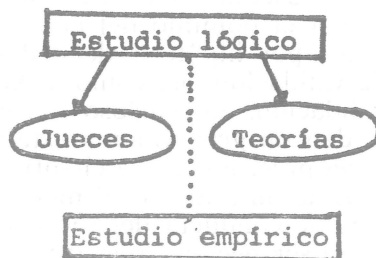
Vamos a explicitar un tanto estos tres aspectos señalados por Nunnally (1978).

2.1. Definición del dominio del constructo

Este aspecto hace referencia a la especificación de los «contenidos» o conductas presumiblemente constitutivas del mismo. Sería algo similar a la definición de la red nomológica del constructo; lo que Cronbach y Meehl (1967) denominan «elucidación de lo que es un objeto y exposición de las leyes bajo las cuales ocurre».

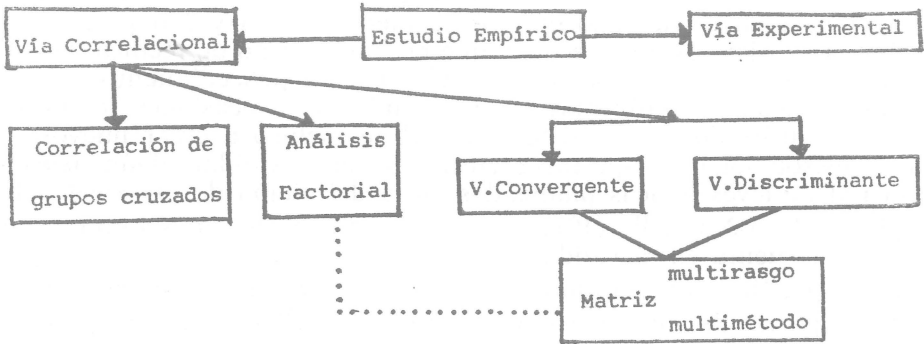
1. Especificar el dominio de observables relativas al constructo

— ¿Cómo definir? .....



— ¿Una especificación empírica? .....

2. ¿Hasta qué punto los observables tienden a medir el mismo constructo o constructos diferentes?



3. Las medidas supuestas del constructo, ¿se comportan como tales en el curso de investigaciones científicas basadas en la teoría del constructo?

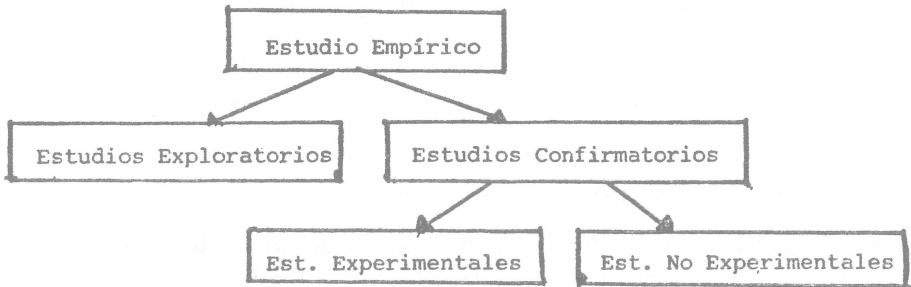


TABLA 1.—Aspectos importantes de la validez de constructo (Nunnally)

Sin duda alguna, este aspecto es de una importancia capital en la investigación sobre variables de construcción, dado que de cómo se defina el dominio del constructo y de cuáles sean sus contenidos definidos por hipótesis dependerá en gran parte todo el trabajo posterior de validación, así como la importancia de las conclusiones de la propia validación. Es curioso observar que aunque todos los investigadores le dan un gran valor a este aspecto, no siempre se realiza con la necesaria precisión y con el suficiente rigor. Obviamente, cuando no tenemos todavía un criterio definido, en su lugar pueden utilizarse diversas muestras de conducta del rasgo fundamental o variable de constructo. Esto suele ser necesario en las primeras definiciones de un constructo y ello está perfectamente justificado, ya que cuando la teoría sobre algún

constructo no está suficientemente establecida en sus orígenes o cuando la orientación del propio constructo es eminentemente práctica, «las primeras formulaciones de las relaciones de la red pueden ser meramente descriptivas» (Cronbach y Meehl, 1967); por otra parte, aún las más avanzadas ciencias físicas no pueden establecer con rigor y precisión el dominio del constructo en sus orígenes. Evidentemente, en nuestras ciencias (sociales en general, pedagógicas en particular) en la mayoría de las ocasiones no disponemos de una red de relaciones consistente, ni siquiera de formulaciones absolutamente explícitas que nos permitan la definición inicial del dominio de un constructo. A medida que investigamos, en las siguientes fases del proceso de validación, podemos ir definiendo con creciente precisión dicho dominio, así como las relaciones que se presumen entre distintos observables implicados en el constructo. En esta fase de continuas definiciones y redefiniciones del constructo los modelos exploratorios de investigación pueden orientar la definición progresiva de hipótesis con creciente precisión.

## 2.2. *¿Los observables tienden a medir la misma cosa, algunas cosas o muchas cosas diversas?*

Para resolver la segunda serie de objetivos propuestos en el segundo aspecto señalado por Nunnally (1978) se utilizan básicamente métodos analíticos y correlacionales. Algunos de los procedimientos de la psicometría tradicional tales como: la validez predictiva y la validez concurrente; estudios sobre consistencia interna de las pruebas, estudios de fiabilidad, estudios de índices de homogeneidad, dificultad o validez de elementos..., etc., apoyan también con frecuencia determinados aspectos de la red teórica emanada del constructo al permitir una mayor precisión en la definición de hipótesis sobre el dominio del constructo. Cronbach y Meehl (1967) señalan que al profesional que le interese la validez con fines prácticos (pedagogo, psicólogo, etc.) le concierne «encontrar no un criterio imperfecto, sino un criterio de cualquier manera», en especial en el campo diagnóstico, ya que, bien no puede esperar un claro y unitario criterio o bien parece difícil que pueda encontrarse, más aún cuando existen otros instrumentos de una dimensión o constructo en estudio. Ciertamente, esta opinión de Cronbach y Meehl contrasta con la posición más clásica de Anastasi (1959), quien afirma que «sólo como medida de un criterio definido específicamente, un test puede ser validado con objetividad..., pretender que un test mide algo más que su criterio es pura especulación»; sin embargo, este mismo autor apoya la validez de constructo al indicar que «los tests pueden ser interpretados provechosamente si conocemos la relación entre la conducta examinada y otras muestras de conducta, sin que ninguna de estas muestras ocupe necesariamente la posición preeminente de un criterio».

En la respuesta a la cuestión planteada en este aspecto, puede utilizarse evidencia procedente de muchos tipos de estudios, tanto exploratorios como confirmatorios, pero pensamos que son estos últimos los que más luz pueden arrojar sobre el constructo y su medida. Entende-

mos por estudio confirmatorio no sólo a aquel estudio que se realiza en contextos experimentales o causales, sino a todo estudio cuyo objetivo sea el validar hipótesis teóricas emanadas de la red nomológica de un constructo. Por estudios exploratorios no entendemos estudios realizados en ausencia total de contexto teórico, sino más bien estudios en los que, por imposibilidad o dificultad, no han podido definirse hipótesis específicas sobre los datos o variables objeto de una investigación concreta.

2.3. *¿Hasta qué punto las medidas supuestas del constructo producen resultados que son predecibles a partir de hipótesis teóricas altamente aceptadas, concernientes al constructo?*

Para dar respuesta a esta cuestión se precisa la integración de evidencia procedente de múltiples tipos de investigación. El problema a resolver aquí es la determinación del cómo se comportan medidas supuestas del constructo en el curso de investigaciones científicas basadas en la teoría subyacente en el constructo. En principio, muchos pueden ser los procedimientos y técnicas utilizables en estos estudios, pero no olvidemos, que los procedimientos y las técnicas no son otra cosa que instrumentos al uso para la respuesta a cuestiones específicas más o menos complejas, por ello, la determinación de qué tipo de procedimiento o de qué técnica o técnicas deberán utilizarse en cada caso, dependerá del objetivo específico de cada estudio —en concordancia con las hipótesis que pretenda validar— y del tipo de datos seleccionados.

En general, para abordar esta cuestión, desde la perspectiva de la validación de constructo, pueden utilizarse estudios experimentales y/o ex-post-facto, en los cuales se puede determinar el comportamiento diferenciador de las medidas supuestas del constructo, así como la forma, magnitud y dirección de las relaciones que mantiene con otras variables afines o implicadas en la teoría del constructo.

Pensamos que los modelos confirmatorios (en contextos experimentales o no experimentales) son los que pueden ofrecer evidencias más valiosas para la clarificación de un constructo.

En este momento no vamos a desarrollar modelos que puedan utilizarse en la validación de constructos, dada la extensión que ello preciaría. No obstante, y de modo orientativo, presentamos la siguiente clasificación:

A) Estudios exploratorios:

- Validez de contenido.
- Análisis correlacionales y estudios de características técnicas de instrumentos de medida.
- Análisis factorial exploratorio.
- Análisis de conglomerados.
- Análisis de correspondencias.

## B) Estudios confirmatorios:

- Estudios experimentales.
- Análisis factorial como técnica de validación de constructo.
- Matriz multirasgo-multimétodo.
- Modelos estructurales causales.
- Análisis factorial confirmatorio.
- Análisis cluster confirmatorio.
- Estudios diferenciales.

### 3. Pasos en la validación de un constructo

Como ya se ha apuntado con anterioridad, el proceso de validación de un constructo en realidad no difiere en su esencia del proceso científico que se utiliza para desarrollar y contrastar teorías. Tampoco difiere dicho proceso del propio de la investigación científica (problema, hipótesis, variables, diseño...) aplicado a las ciencias sociales. En realidad, podríamos definir el proceso de validación de un constructo como un proceso *iterativo, interactivo y de integración de resultados*. Iterativo, porque requiere la repetición de un proceso esquemático básico para la confirmación de hipótesis concretas y específicas implicadas en la teoría sobre un constructo. Interactivo, porque exige estudios diversificados que permitan descubrir las interacciones (covariaciones) entre las variables esenciales implicadas en la red de relaciones establecida por el constructo. Y de integración de resultados, porque los distintos estudios realizados sobre la base de un constructo, será necesario interpretarlos de forma integrada, descubriendo *qué contradicciones* se manifiestan en los resultados de los distintos estudios y *qué coincidencias* muestran estos estudios diversificados.

A modo de resumen de los puntos anteriores, presentamos aquí, en cinco pasos (que podrían ser más o menos, dependiendo de diversos criterios utilizables en la estructuración del proceso de validación) nuestra síntesis sobre los pasos a seguir por el proceso-tipo de validación de un constructo.

#### 3.1. Explicitación de la teoría sobre el constructo

Este primer paso es similar a la definición del problema en el proceso de investigación científica. En realidad, no es posible admitir un constructo científico sino no se define su contexto teórico, es decir, si no se define la *red nomológica* (en la terminología de Cronbach y Meehl, 1967) mínima de asociaciones o proposiciones en torno al constructo. De hecho, sin esta red de relaciones o árbol estructural del constructo, en un razonable grado de explicitación, no es posible aceptar un constructo como científico (ver punto 2).

Es evidente, como afirman Cronbach y Meehl (1957), que si dicha red no hace contacto con las observaciones (es decir, si no hay definiciones operativas de variables), ni se explicita o manifiesta claramente, la validez de constructo no puede exigirse.

Ahora bien, no olvidemos tampoco, que los constructos deben entenderse como «*compendios inductivos*», esto es, la teoría que subyace a ellos no tiene por qué estar establecida en sus orígenes, es más, en las primeras etapas de desarrollo de un constructo y en ocasiones en las más avanzadas, sobre todo cuando la orientación es práctica, interviene poco o nada la teoría entendida en su sentido usual, por lo cual, *las primeras formulaciones de las relaciones de la red pueden ser meramente descriptivas*. Con la propia investigación y las continuas redefiniciones del constructo, se irán formulando las leyes de la red con mayor precisión. Por otra parte y como también afirman Cronbach y Meehl (1967), la vaguedad de las actuales leyes pedagógicas y/o psicológicas suponen también una restricción importante del esquema de la red. En la práctica, aún las más avanzadas leyes de las ciencias físicas tan solo se aproximan a este ideal y la pedagogía o las ciencias sociales en general, que trabajan con formulaciones «crudas» o semiexplícitas, se encuentran en bastantes ocasiones en la imposibilidad de establecer una red consistente.

En todo caso, el significado de un constructo no está determinado cuando la red es muy incompleta y/o tiene muchas hebras perdidas y/o algunos constructos están unidos sólo por finos hilos, dado que su «definición explícita» es vaga y por tanto no sabremos *qué es* determinado constructo; mientras tanto nuestra hipotética red de relaciones está sin validar.

En esta primera fase, los estudios exploratorios pueden ser una poderosa herramienta de investigación, dado que permitirán ir afinando en las posibilidades de definición de las variables no implicadas directamente en la red.

En todo caso, debe quedar claro, que los constructos no surgen de forma «mágica»; es decir, no surgen fuera del contexto teórico, sino que aparecen ligados a una teoría desde sus orígenes, aún cuando no pueda explicitarse su red de relaciones con precisión. Así, por ejemplo, el constructo de Inteligencia (Spearman, 1904) surgió por las consistentes relaciones entre variables empíricas que medían conductas «inteligentes»; el constructo de estilo cognitivo Dependencia-Independencia de Campo Perceptivo (Witkin, 1954 y 1962) surgió por las consistentes diferencias que Witkin encontró en distintas muestras de sujetos al aplicar el test de figuras enmascaradas de Gottschald y otros tests de percepción de la verticalidad (R.F.T. y C.I.S.I.)..., etc. [3].

### 3.2. *Formulación de hipótesis derivadas de la red teórica y de los estudios exploratorios sobre el constructo*

Este segundo paso —también segundo paso del proceso general de

investigación científica— es la consecuencia lógica del establecimiento de la red teórica de relaciones del constructo y de los primeros estudios exploratorios realizados sobre variables implicadas.

Por lo general, la formulación de dichas hipótesis exige la definición operativa de variables, bien propuestas como medidas del propio constructo o bien (si lo anterior no es posible), variables implícitas en la red de relaciones del constructo. No olvidemos, por otra parte, que los procesos de validación de constructo se proponen como objetivo fundamental validar las posibles medidas (directas o indirectas) del mismo.

Estas formulaciones de hipótesis deberán ser precisas, tanto si exigen posteriormente estudios experimentales, o estudios no experimentales, de tipo confirmatorio preferentemente. Es evidente, que dichos estudios sólo podrán *aceptar* o *rechazar* (en sentido amplio) las hipótesis que se propongan.

### 3.3. Puesta a prueba de hipótesis. Estudios confirmatorios

El siguiente paso de todo proceso de investigación científica —una vez definidos sus determinantes: problema, hipótesis y variables— es el diseño de un plan o estrategia de investigación cuyo objetivo es la puesta a prueba de las hipótesis definidas. Dicho diseño —en el amplio sentido que por ejemplo Kerlinger (1973) utiliza— exige la decisión sobre una serie de aspectos, tales como: muestreo, selección de instrumentos, control, elección de modelos estadísticos adecuados para el contraste de la hipótesis, análisis de los datos, marco interpretativo de resultados, etc.

Obviamente, en esta fase las técnicas confirmatorias que anteriormente hemos descrito (experimentales y no experimentales) serán la base del análisis de resultados, aunque no se excluye que algunas hipótesis sobre la significación de simples relaciones entre variables, puedan ser evaluados a través de técnicas correlacionales o de estudios de características técnicas de instrumentos de medida (técnicas exploratorias).

Obviamente, si las evidencias experimentales o correlacionales, muestran acuerdo con las hipótesis que se derivan de la red, el constructo va estableciéndose, pero ¿qué ocurre si las predicciones del investigador y sus datos son discordantes con las derivaciones de la red? Podemos interpretarlo de tres maneras:

- a) El test o los tests propuestos como medidas del constructo no miden la variable de constructo.
- b) La red teórica que generó las hipótesis es incorrecta.
- c) El diseño experimental o correlacional (de validación en términos más generales) fracasó, por no haber podido probar adecuadamente la/s hipótesis.

Lógicamente, en función de cada interpretación, la investigación posterior deberá ir en la línea de corregir el error encontrado. Es decir, si el test no mide el constructo, no puede ser propuesto como medida del

mismo, pues para ello es exigible un fuerte ajuste positivo entre las predicciones y los datos subsecuentes. Si la red es incorrecta, habrá que modificarla en el sentido oportuno y probar nuevas derivaciones. Y si el diseño o procedimiento de validación es incorrecta, parece razonable diseñar o ejecutar un nuevo estudio adecuado.

#### 3.4. *Evaluación de resultados. Integración de resultados*

De todo lo anterior, se deriva la necesidad de investigar secciones separadas de la red (independientes), pues cuanto más diversificadas estén las predicciones del sistema, mayor confianza podemos tener en que solamente pequeñas partes de la red total resultarán defectuosas, no afectando ellas la validez general del constructo establecido teóricamente. Todo ello indica —como afirman Cronbach y Meehl (1967), Mischel (1977)..., etc.— que la validez de constructo no puede expresarse, por lo general, en la forma de un sólo coeficiente simple y que será necesario realizar múltiples investigaciones para corroborar las distintas partes de la red total.

La fase de evaluación de resultados provenientes de distintos estudios realizados sobre distintas hipótesis derivadas de la red total, exige un esfuerzo lógico y teórico de integración, que debe realizarse a partir de la propia teoría sobre el constructo. En ocasiones, es curioso observar que estas tareas de integración de resultados no se realizan con la frecuencia y la minuciosidad que exige todo proceso de validación de teorías o de constructos y por lo cual es frecuente observar la repetición del mismo tipo de estudio (sobre la base de la misma hipótesis) por distintos autores o investigadores, diferenciándose únicamente las muestras de análisis o bien la inclusión o no de alguna variable empírica en el estudio.

Es evidente, pues, que todo proceso de validación de constructo exige la *acumulación de resultados* procedentes de distintos tipos de estudios, pero también es precisa la *integración* de dichos resultados que permitan *dotar al constructo teórico de cierta unidad*, la cual es necesaria para que dicho constructo vaya consolidándose. En esta tarea de integración de resultados, las técnicas implicadas en el Metaanálisis pueden ser de un valor inapreciable, dado que «permiten aplicar procedimientos formales para combinar los resultados de experimentos y estudios correlacionales», posibilitando así inferir generalizaciones sobre la consistencia de una red teórica subyacente a un constructo [4].

El resultado final de esta fase de evaluación o integración de resultados es la consolidación de distintas hipótesis emanadas de la red o bien la reformulación de la propia red teórica (siempre y cuando los resultados de las investigaciones sean consistentes).

### 3.5. *Formulación de nuevas hipótesis. Puesta a prueba a través de estudios confirmatorios. Evaluación de resultados. Reformulación de la red de relaciones del constructo*

En realidad este quinto paso no es un nuevo paso en el proceso de validación de constructo, sino que es la vuelta al segundo paso del proceso de validación. Esto es, el último paso era la evaluación e integración de resultados, pero como el resultado a su vez de dicho paso es la consolidación o reformulación de la propia red del constructo, será necesario volver a reiniciar el proceso de validación de constructo, pues «nunca» puede decirse que un constructo está validado en su totalidad.

La consecuencia, pues, del paso anterior, suele ser la formulación de nuevas hipótesis (bien reformulación de algunas ya contrastadas o bien formulación de nuevas hipótesis a partir de la evidencia encontrada en anteriores estudios y de la lógica modificación de la red teórica del constructo). Estas nuevas hipótesis deberán ser puestas a prueba a través del diseño de estudios confirmatorios. En estos nuevos procesos de puesta a prueba de hipótesis, pensamos que no se justifica la realización de estudios exploratorios, a no ser que no tengamos suficiente claridad conceptual para poder formular hipótesis de cierta precisión. Los resultados de dichos estudios confirmatorios deberán ser evaluados e integrados entre sí, con la consiguiente modificación o consolidación del constructo y, por tanto, con la consiguiente reformulación de la red de relaciones teóricas o red nomológica del constructo.

A modo de esquema, la tabla 2 pretende mostrar gráficamente el proceso cíclico implicado en la validación de un constructo.

## 4. *Conclusiones y discusión*

### 4.1. *El estudio de constructos en el ámbito pedagógico*

Los estudios de validación de constructo se han realizado preferentemente desde la psicología, siendo el objeto fundamental a validar, variables o dimensiones intelectuales, aptitudinales y de personalidad. La bibliografía sobre este tipo de estudios en relación a constructos psicológicos es muy extensa, aunque en ocasiones hay claras deficiencias en las tareas de integración de resultados, lo cual hace que abunden teorías distintas sobre un mismo constructo y que en ocasiones los mismos constructos no pueden definirse a partir de una única red nomológica consistente de proposiciones y relaciones entre variables. Ahora bien, la utilidad de todas estas teorías no deja de ser grande, pues permiten caracterizar, en ocasiones con gran precisión, a una determinada dimensión cognitiva o conductual.

En el campo educativo, las variables de rendimiento han sido, sin lugar a dudas, el tipo de variables más estudiado. Aunque el producto educativo —en su sentido más amplio— se presente como una realidad compleja y en cierto grado evasiva, por sus dificultades intrínsecas de

operacionalización, no deja de ser cierto que sin su consideración gran parte de la razón de ser de la investigación, en las distintas áreas pedagógicas, se perdería. Cada vez más, el rendimiento, como producto de la educación, viene a ser considerado en términos de estructuras y núcleos organizados de conducta facilitadores de la solución de problemas y nexos de unión entre distintas materias de enseñanza y no tanto

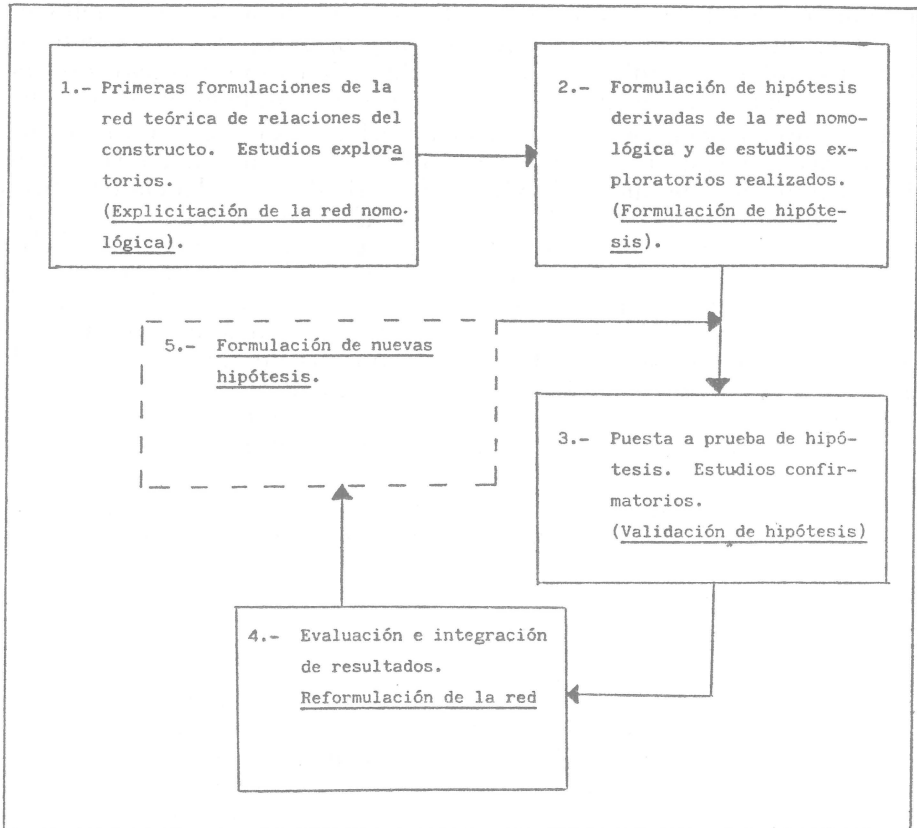


TABLA 2.—Pasos del proceso de validación de constructo

como la simple adquisición de datos y destrezas vinculadas a determinados contenidos (Orden, 1985c). Es decir, la reconsideración del estudio del rendimiento, como variable dependiente por excelencia, supone un cambio radical en su medida y por ello el tema de la validez de constructo se hace más grave y urgente. En realidad, se trata de validar variables o constructos que subyacen a variables representativas del producto educativo. Ello implica —como afirma De la Orden (1985c)— el hallazgo de evidencia que apoye la teoría que sirve de base a la variable,

tanto como a la medida misma. La validez de constructo expresará que las variables derivadas de una teoría del producto educativo son adecuadas. En este sentido, una prueba de rendimiento tendrá validez de constructo en la medida en que sus resultados permitan inferencias sólidas sobre la variable medida cuyas características comportamentales, coherentes con la teoría, han sido empíricamente contrastadas. El problema fundamental es que, en general, las variables utilizadas en la mayoría de las ocasiones y que representan al rendimiento, no han sido derivadas de una teoría, sino seleccionadas de forma ecléctica y pragmática, por lo que hablar de validez de constructo de dichas medidas no tiene demasiado sentido. Por otra parte, y desde una dimensión aplicada de la investigación sobre el rendimiento, no cabe duda de que hasta ahora la metodología general de investigación en pedagogía ha acudido prioritariamente a los modelos experimentales (entendidos en sentido amplio) provenientes de las ciencias biológicas y ello porque las relaciones entre las variables consideradas en esas ciencias presentaban cierto homomorfismo con las relaciones que se dan entre las variables de nuestro campo. Aún cuando no hayan sido bien utilizadas siempre, no puede negarse que las aportaciones de dicha metodología experimental al conocimiento científico de la educación han sido, no sólo útiles, sino fundamentales.

No obstante lo anterior, pensamos que la combinación de la metodología experimental con las metodologías exploratoria y confirmatoria en campos multivariados, dentro del marco de referencia de procesos de validación de constructo, no ha sido suficientemente explotada en el estudio de variables de producto educativo.

Haertel (1985) apunta una línea de investigación sobre resultados educativos que puede ser fecunda, considerando a éstos como constructos apoyados en teorías psicológicas que conciben los resultados de la educación en términos de procesos cognitivos o mecanismos de memoria que han sido desarrollados a partir de la instrucción (Snow, 1980). En esta perspectiva, De la Orden (1985c) afirma que la validación de constructo, especialmente para la interpretación del producto de la educación en base a medidas de referencia criterial, es una necesidad ineludible que puede abrir caminos prometedores hacia una conceptualización más adecuada de los resultados de la instrucción. En todo caso, tal necesidad pensamos que podría extenderse a otro tipo de constructos pedagógicos de gran importancia, tales como son variables o dimensiones actitudinales, de autoconcepto, de autosatisfacción y de intereses.

#### *4.2. Dificultades y posibilidades de los procesos de validación de constructo en el ámbito pedagógico*

A través de los capítulos anteriores, hemos tratado de poner de manifiesto la lógica de la validación de constructo, así como los aspectos que dicho tipo de validación pretende evaluar y el tipo de estudios o modelos de investigación que pueden utilizarse en dichos procesos de validación. También hemos apuntado en el apartado anterior la nece-

idad creciente de validación que exigen los constructos o dimensiones pedagógicas, consideradas como variables dependientes por excelencia. Ahora bien, no se nos escapan las dificultades inherentes a los procesos de validación de constructo. Tales dificultades son ciertamente las mismas dificultades que afectan a la investigación básica en cualquiera de las ciencias sociales, es decir, el ser un camino largo, lento y a veces oscuro que no siempre están dispuestos a afrontar los investigadores y estudiosos de nuestro campo. Un camino que exige la acumulación de datos y de resultados provenientes de múltiples fuentes y tipos de estudios y que exige tareas arduas de integración de resultados. Por otra parte, en nuestro país —también ocurre en otros países, pero en menor escala— hay falta de líneas claras de investigación que tengan como objetivos finales la validación de constructos, a lo que se suma la carencia (escandalosa en ocasiones) de presupuestos y equipos de trabajo mínimos (personales y materiales) para afrontar este tipo de tareas. A todo lo anterior se suman también dificultades de acceso a fuentes documentales o instrumentales, básicas para las tareas de recogida de información, aplicación de instrumentos y procesamiento de datos.

Sin embargo, las posibilidades de realización de este tipo de estudios son objetivamente más factibles hoy en día que hace unos cuarenta años, dado que tenemos a nuestro alcance diversos elementos facilitadores de este tipo de investigación; tales elementos son: *las modernas técnicas estadísticas multivariadas*, isomorfas con el tipo de tareas que exigen los procesos de validación de constructo; *la informática* y especialmente los paquetes de programas estadísticos para ordenador (SPSS, BMDP, OSIRIS, SAS, LISREL, SPADE, ADDAD, etc.), los cuales permiten el procesamiento de datos, en función de la técnica o modelo estadístico que precisamos, con gran celeridad y precisión; y *las modernas técnicas de integración de datos cuantitativos* procedentes de la investigación empírica, las denominadas técnicas de Metaanálisis.

**Dirección del autor:** José Manuel García Ramos, Departamento de Métodos de Investigación y Diagnóstico de Educación, Facultad de Filosofía y Ciencias de la Educación, Ciudad Universitaria, 28040 Madrid.

#### NOTAS

- [1] Tomado de Bartolomé (1983).
- [2] Como afirman diversos autores: Cronbach y Meehl (1967); Martínez Arias (1981); Bartolomé (1983); De la Orden (1985); ...; la validez de constructo no debe restringirse en su utilización a los rasgos de personalidad, sino que puede y debe ser utilizada en procesos de validación de otras variables y/o teorías de evaluación pedagógica, de teoría conductual, de «assessment», etc.
- [3] Ver Witkin (1954); Witkin y otros (1962); Fernández Ballesteros (1980); García Ramos (1984), etc.
- [4] Ver Glass (1976); Cooper y Rosenthal (1980); Hedges y Olkin (1981); Slavin (1984); Carlberg (1984); De la Orden Hoz (1985b), etc.

## BIBLIOGRAFIA

- APA (1954) *Normas sobre tests y manuales educativos y psicológicos* (Madrid, TEA) (1976).
- BALTES, P. B. y otros (1981) *Métodos de investigación en psicología evolutiva: enfoque del ciclo vital* (Madrid, Morata).
- BARTOLOMÉ, M. (1983) *Validez de los instrumentos empleados en los procesos de medida. Evaluación y Assessment*. (Universidad de Barcelona, Lección Magistral, inédita).
- BECHTOLDT, H. (1959) Validez de constructo: una crítica, en MEGARGEE, E. J. *Métrica de la personalidad*, 2 vols. (México, Trillas), 1971 (pp. 162-181).
- BUNGE, M. (1976) *La investigación científica* (Barcelona, Ariel).
- CAMPBELL, D. y FISKE, D. (1957) *Convergent and discriminant validation by the multitrait-multimethod matrix*, en MEHRENS y EBEL (1967), pp. 273-302.
- CARLBERG, C. G. (1984) Meta-analysis in education (A reply to Slavin), *Educational Researcher*, 13 (8), pp. 6-15.
- COOPER, H. N. y ROSENTHAL, R. (1980) Statistical versus traditional procedures for summarising research findings, *Psychological Bulletin*, 87, pp. 442-449.
- CRONBACH, L. J. y MEEHL, P. E. (1967) Construct validity in psychological tests, en MEHRENS, W. y EBEL, R. L. (eds.) *Principles of educational and psychological measurement. A book of selected readings* (Chicago, Rand McNally), pp. 243-270.
- CRONBACH, L. J. (1971) Tests validation, en THORNDIKE, R. L. (ed.) *Educational measurement* (Washington, American Council of Education), pp. 443-507.
- CUADRAS, C. (1981) *Métodos de análisis multivariante* (Barcelona, Eunibar).
- FISKE, P. (1971) *Measuring the concepts of personality* (Chicago, Aldine, Pub.).
- GARCÍA RAMOS, J. M. (1984) *Validación y medida del constructo dependencia-independencia de campo perceptivo* (Madrid, tesis doctoral, inédita, Universidad Complutense).
- (1985) Análisis de conglomerados, en ORDEN, A. DE LA *Investigación educativa* (Madrid, Diccionario Anaya), pp. 1-12.
- (1986) *Modelos exploratorios y confirmatorios en la investigación pedagógica* (Madrid, lección inédita, Universidad Complutense).
- GLASS, G. V. (1977) Integrating findings. The meta-analysis of research, *Review of Research in Education* 5, pp. 351-379.
- HAERTEL, E. (1985) Construct validity and criterion-referenced testing, *Review of Educational Research*, 55, 1, pp. 23-46.
- HEDGES, L. V. y OLKIN, I. (1980) Vot-counting methode in research syntesis, *Psychological Bulletin*, 88, pp. 249-269.
- JORESKOG, K. G. (1969) A general approach to confirmatory factor analysis, *Psychometrika*, 34, pp. 183-202.
- JORESKOG, K. G. y SORBOM, D. (1983) *LISREL V y VI* (Uppsala, Universidad Uppsala).
- KERLINGER, F. N. (1973) *Investigación del comportamiento. Técnicas y metodología* (México, Interamericana).
- MAGNUSON, D. (1966) *Teoría de tests* (México, Trillas).
- MARTÍNEZ ARIAS, R. (1981) Validez de constructo, en FERNÁNDEZ BALLESTEROS, R. y OTROS *Evaluación conductual* (Madrid, Pirámide).
- MARTÍNEZ ARIAS, R. (1985) Modelos de ecuaciones estructurales en la orientación escolar y profesional, en *La orientación escolar ante la reforma de las enseñanzas medias y el empleo juvenil* (Madrid, I Congreso de Orientación Escolar y Profesional) (pp. 572-591).
- MEGARGEE, E. (1971) *Métrica de la personalidad* (México, Trillas).
- MISCHER, W. (1977) *Personalidad y evaluación* (México, Trillas).
- NUNNALLY, J. C. (1978) *Psychometric theory* (New York, McGraw-Hill).

- ORDEN HOZ, A. DE LA (1985a) Análisis causal, en ORDEN HOZ, A. DE LA *Investigación educativa* (Madrid, Diccionario ANAYA).
- (1985b) *Integración cuantitativa de resultados*, en ORDEN HOZ, A. DE LA *Investigación educativa* (Madrid, Diccionario Anaya).
- (1985c) *Hacia una conceptualización del producto educativo* (Gijón, Ponencia del III Seminario de Modelos de Investigación Educativa).
- PÉREZ JUSTE, R. (1983) *Pedagogía Experimental (Adaptación)*, Unidad Didáctica (Madrid, Universidad Nacional de Educación a Distancia).
- POPHAM, W. J. (1980) *Problemas y técnicas de la evaluación educativa* (Madrid, Anaya 2).
- SEOANE, J. (1980) *Análisis factorial*, Unidad Didáctica IV de Psicometría (Madrid, Universidad Nacional de Educación a Distancia).
- SHAVELSON, R. J. y WEEB, N. M. (1981) Generalizability theory: 1973-1980, *British Journal of Mathematical and Statistical Psychology*, 34, pp. 133-166.
- SLAVIN, R. E. (1984) Meta-analysis in education: how has it been used?, *Educational Researcher. Psychological Bulletin*, 88, pp. 6-15.
- SNOW, R. E. (1980) Aptitude and achievement, en B. W. SCHRAEDER (ed.) *Measuring achievement: Progress over a decade. New directions for testing and measurement* (San Francisco, Jossey-Bass).
- TRAUB, R. E. y WOLFE, R. G. (1981) Latent trait theories and the assessment of educational achievement, *Review of Research in Education*, n. 9, AERA.
- WILDE, G. J. L. (1977) Trait description and measurement by personality questionnaire, en CASTELL y DREYER *Handbook of modern personality theory* (New York, Hensincher, Pb. 10).
- WITKIN, H. (1954) *Personality through perception* (West-Post. Harper).
- WITKIN, H. y GOODENOUGH, D. (1977) *Field dependence revisited* (Princeton. Educational testing service. Research Bulletin), pp. 77-196.

SUMARIO: El presente artículo analiza la problemática específica de los procesos de validación de constructos en el ámbito pedagógico. En la actualidad, parece claro que la validez psicométrica no resuelve, sino en parte, la problemática planteada por variables de construcción de alto nivel de abstracción, dado que sólo ofrece índices relativos, sobre la validez de un instrumento de medida, siendo esta información insuficiente para validar constructos más amplios como pueden ser por ejemplo, el mismo constructo de inteligencia, constructos de estilo cognitivo y de aprendizaje e incluso constructos referidos a la variable dependiente por excelencia del ámbito pedagógico: el producto de la educación, el rendimiento, en cualquiera de sus posibles conceptualizaciones y manifestaciones.

Este trabajo pretende esclarecer el sentido y los objetivos de lo que suele denominarse «validez de constructo», sin duda, la más importante desde el punto de vista científico. Por otra parte, se intenta también reflejar en él los pasos propios de cualquier proceso-tipo de validación de constructos, pasos que van desde la explicitación de la teoría-base del constructo, hasta la evaluación e integración de resultados, con la consiguiente formulación de nuevas hipótesis que deberán ser de nuevo puestas a prueba. En este sentido, es definido dicho proceso de validación de constructo, como iterativo, interactivo e integrativo.

**Descriptor:** Construct validity, Validity, Measurement, Test theory.