

Exigencias de calidad e instrumentos de evaluación. Comparación de las tradiciones llamadas cuantitativa y cualitativa

FELIPE MARTÍNEZ RIZO

Departamento de Educación / UAA

INTRODUCCIÓN

LAS CRÍTICAS DE LAS PRUEBAS OBJETIVAS Y LA POLÉMICA «CUANTI-CUALI»

En los medios académicos del campo de las ciencias sociales, y específicamente entre los partidarios de los enfoques llamados cualitativos, se escuchan consideraciones sobre las posibles limitaciones de los conceptos de validez, confiabilidad y similares, fundamentales en los acercamientos psicométricos. Estas reflexiones se enmarcan en el conjunto de las discusiones que, por demasiado tiempo, han enfrentado a los más acérrimos defensores de uno u otro enfoque, y que en el medio educativo han sido designadas con la expresión de «guerras paradigmáticas».¹

En otros lugares² se ha mostrado cómo tales polémicas, con variantes lógicas, se han repetido en diversos momentos a lo largo del siglo XX, pero que pueden rastrearse en el XIX y el XVIII

1 Gage, N. L., "The paradigms wars and their aftermath: A 'historical Sketch' of research on teaching since 1989", *Educational Researcher*, Vol. 18 (1989) N° 7, pp. 4-10.

2 Martínez Rizo, Felipe, "The controversy about quantification in social research. An extension of Gage's 'historical sketch' ". *Educational Researcher*,

e, incluso, en el XVII. A lo largo de ese tiempo ha habido también opiniones que han insistido en el carácter complementario, más que contrapuesto, de los enfoques «cuanti» y «cuali». Esta última es la tónica que, afortunadamente, parece estarse imponiendo en la última década del 1900, como muestran múltiples trabajos que coinciden en defender dicha complementariedad en la literatura más reciente, tanto en inglés³ como en español.⁴

En la misma dirección, y dado que no pocas de las críticas que se hacen a los instrumentos de tipo objetivo que se desarrollan para la evaluación de aprendizajes, no se limitan a señalar fallas precisas de algún instrumento en particular, sino que pretenden cuestionar radicalmente ese tipo de instrumentos, a partir de planteamientos generales de carácter epistemológico, resulta conveniente hacer algunas consideraciones sobre la compatibilidad y posible complementariedad de los puntos de vista denominados tradicionalmente cuantitativos, y los que se designa con la etiqueta de cualitativos, en el terreno que nos ocupa. Para ello se utilizará la obra de dos autores de reconocido prestigio en el campo cualitativo: Yvonna S. Lincoln y Elton G. Guba.⁵

LA CONFIABILIDAD Y LA VALIDEZ Y SUS EQUIVALENTES

Los conceptos de validez, confiabilidad y similares se desarrollaron, como criterios de calidad del trabajo científico, en un con-

Vol. 20 (1991) N° 9, pp. 9-12; "La polémica sobre la cuantificación en las ciencias del hombre". *Papers* (U. A. de Barcelona) N° 42 (1993) pp. 13-34; *El oficio del investigador educativo*. Aguascalientes, Universidad Autónoma de Aguascalientes, 1997, pp. 247-285.

3 Cfr. Newman, Isadore, y Carolyn R. Benz, *Qualitative-quantitative research methodology. Exploring the interactive continuum*. Carbondale & Edwardsville, Southern Illinois Press, 1998.

4 Cfr. Bericat, Eduardo, *La integración de los métodos cuantitativo y cualitativo en la investigación social. Significado y medida*. Barcelona, Ariel, 1998.

5 En particular en su obra *Naturalistic Inquiry*. Beverly Hills, Sage, 1985.

texto en el que se consideraba que la metodología por excelencia de la investigación se identificaba con el llamado «método experimental».

De acuerdo con Lincoln y Guba, los criterios de calidad tradicionales tratan de responder cuatro preguntas fundamentales para establecer la medida en que «podemos fiarnos» de una investigación o indagación cualquiera sobre algún aspecto de la realidad, su «*trustworthiness*». Esas preguntas plantean las cuestiones de la *verdad* de un hallazgo, o sea su correspondencia con la realidad (*truth value*); la *aplicabilidad* de un hallazgo en contextos diferentes a aquel en el que se encontró (*applicability*); la *consistencia* de los hallazgos, esto es, la posibilidad de replicarlos (*consistency*); y la *neutralidad*, esto es la medida en que los hallazgos no dependen de perspectivas o sesgos particulares de los investigadores (*neutrality*).⁶

Las respuestas que la metodología convencional ha dado a esas cuatro preguntas se expresan con los cuatro criterios tradicionales de calidad de un estudio que son la *validez interna*, la *validez externa*, la *confiabilidad* y la *objetividad*, cada una de las cuales tiene sus reglas y condiciones.

Según Lincoln y Guba, los acercamientos cualitativos (o mejor, en su terminología, los trabajos del *paradigma naturalista*) tratan de responder las mismas preguntas, pero lo hacen de maneras diferentes; estos dos autores proponen a su vez cuatro conceptos que, en su opinión, expresan las mismas cuatro dimensiones de la calidad de la investigación, en una forma que resulta más adecuada a la naturaleza de los trabajos de corte cualitativo -o del *paradigma naturalista*- evitando las concepciones estrechamente asociadas a los conceptos tradicionales anteriores.

Esos cuatro conceptos alternativos propuestos son los de *credibilidad*, *transferibilidad*, lo que podríamos llamar *confiabilidad*

6 *Op. cit.*, p. 290.

*alternativa (dependability) y confirmabilidad.*⁷

Los términos escogidos por Lincoln y Guba son muy similares a los de la tradición experimental, casi sinónimos en el lenguaje cotidiano y, a la vez, diferentes. Esto se aprecia mejor en la visión de conjunto del cuadro siguiente, y se nota especialmente en el caso de *reliability* y *dependability*, para los que no hay equivalentes castellanos distintos; por ello en este trabajo ambos se traducen por el término confiabilidad, con el matiz de “alternativa” en el segundo caso.

Cuadro 1		
<i>Las cuatro preguntas clave para saber si una investigación es de fiar (trustworthiness) en las dos tradiciones, según Lincoln y Guba.</i>		
EN GENERAL	TRADICIÓN EXPERIMENTAL	PARADIGMA NATURALISTA
VERDAD <i>(truth value)</i>	VALIDEZ INTERNA <i>(internal validity)</i>	CREDIBILIDAD <i>(credibility)</i>
APLICABILIDAD <i>(applicability)</i>	VALIDEZ EXTERNA <i>external validity)</i>	TRANSFERIBILIDAD <i>(transferability)</i>
CONSISTENCIA <i>(consistency)</i>	CONFIABILIDAD <i>(reliability)</i>	CONFIABILIDAD B <i>(dependability)</i>
NEUTRALIDAD <i>(neutrality)</i>	OBJETIVIDAD <i>(objectivity)</i>	CONFIRMABILIDAD <i>(confirmability)</i>

⁷ *Idem*, pp. 294-301.

En los párrafos siguientes se comparan uno a uno los cuatro pares de términos, apuntando de manera breve la forma particular en que el «paradigma experimental» (lo que aquí se designa con la expresión «la tradición cuantitativa») y el «paradigma naturalista» (la «tradición cualitativa») buscan atender cada dimensión de la calidad de las investigaciones respectivas.

LA PREGUNTA SOBRE LA VERDAD

La primera pregunta puede formularse de la siguiente manera: ¿son *verdaderos* los hallazgos de la investigación, corresponden a la realidad?

En la «tradición cuantitativa», o experimental, el término que se utiliza para designar la dimensión de la calidad a que se refiere esta pregunta es el de la *validez interna* de la investigación.

En la investigación de tipo experimental, la *validez interna* clásicamente se atiende cuidando las circunstancias que pueden amenazarla: la influencia de la historia; la de los procesos de maduración; la de las pruebas previas que pueden mejorar el desempeño de los sujetos en las posteriores; las deficiencias de los instrumentos; las tendencias de regresión estadística hacia los valores promedio; la selección diferenciada de los grupos experimentales y de control; la pérdida diferente de sujetos en los grupos (mortalidad experimental); y la interacción entre maduración y selección.⁸

En la «tradición cualitativa» o, si se prefiere, en el paradigma naturalista, esta dimensión de la calidad de una investigación podría designarse, según Lincoln y Guba, con el término *credibilidad*.

8 Cfr. Campbell, Donald y Julian C. Stanley. "Experimental and quasi-experimental design for research and teaching". En Gage, N. L. *Handbook of Research on Teaching*. Chicago, Rand McNally, 1963. Edición en español: Buenos Aires, Amorrortu.

En los estudios de tipo cualitativo, la *credibilidad* puede atenderse con la prolongación del estudio; la persistencia de las observaciones; la triangulación de fuentes, métodos y personas; la discusión de los hallazgos entre los investigadores (*peer debriefing*); el análisis de casos negativos; la doble verificación de registros con posterioridad a su recolección (*referential adequacy*); y los procedimientos de verificación de registros e interpretaciones mediante la retroalimentación de los sujetos mismos del estudio, tanto a lo largo del proceso de investigación como al término del mismo.⁹

LA PREGUNTA SOBRE LA APLICABILIDAD

La segunda pregunta se expresa así: ¿son *aplicables* los hallazgos del estudio a otros contextos?

La expresión que se utiliza en la «tradición cuantitativa» para atender este aspecto es la de *validez externa*, y las amenazas al respecto en esta tradición pueden sistematizarse como los efectos de selección, los del contexto del estudio (*setting*), los de la historia y los llamados efectos de constructo (*construct effects*), en el sentido de que precisamente las dimensiones estudiadas, definidas de la manera que lo hace cierto estudio, pueden ser características del grupo estudiado.¹⁰

En la «tradición cualitativa» esta dimensión puede tratarse utilizando el término *transferibilidad*, y la manera de cuidarla es la llamada “descripción gruesa” (*thick description*), o sea en la característica de los trabajos que consiste en que se detallen las características de los casos y contextos y todas las circunstancias

9 Lincoln, Yvonna S. Y Elton G. Guba, *op. cit.*, pp. 310-316.

10 Cfr. Lecompte, M. D. y J. P. Goetz, “Problems of reliability and validity in ethnographic research”, *Review of Educational Research*, Vol. 52 (1982) pp. 31-60, citados por Lincoln y Guba, *op. cit.*, p. 298.

relevantes, de manera que sea posible juzgar si un trabajo es demasiado específico o transferible.¹¹

LA PREGUNTA SOBRE LA CONSISTENCIA

La tercera pregunta es: ¿son *consistentes* los hallazgos del estudio, pueden replicarse?

El término clásico correspondiente en la «tradición cuantitativa» es *confiabilidad*, y se atiende con réplicas controladas, análisis de los resultados de partes de los instrumentos comparados con los de otras partes de los mismos y mecanismos similares.

El término propuesto por Lincoln y Guba para esta dimensión es el de *dependability* que, no habiendo al parecer un equivalente en castellano distinto al que se usa en la tradición experimental, optamos por traducir como *confiabilidad alternativa*, y puede atenderse en los trabajos de orientación cualitativa mediante la triangulación, ciertas formas de replicación (*stepwise replication*) y con la llamada *auditoría de investigación*, que se refiere a mecanismos sistemáticos de registro de los pasos que se van dando en el proceso y los hallazgos parciales, de manera que sea posible que otro investigador revise esos pasos y esos hallazgos parciales, gracias a los registros conservados de ellos, permitiendo una triangulación que arroje como resultado esta característica de *confiabilidad*, en una forma alternativa.¹²

LA PREGUNTA SOBRE LA NEUTRALIDAD

La última de nuestras cuatro preguntas se formula como sigue: ¿son *neutrales* los hallazgos, o dependen de las características o sesgos personales de sus autores? En la corriente cuantitativa

11 Lincoln, Yvonna S. y Elton G. Guba, *op. cit.*, p. 316.

12 *Op. cit.*, p. 316-318.

esta pregunta se atiende mediante el concepto de *objetividad*, y tradicionalmente se considera que se alcanza a través de la intersubjetividad, de la coincidencia de las subjetividades independientes de varias personas.

El término correspondiente propuesto por Lincoln y Guba es *confirmabilidad*, y en la tradición cualitativa puede atenderse de manera similar a la característica anterior, mediante la auditoría de la investigación, según la propuesta de Edward S. Halpern.¹³

CONCLUSIÓN

La epistemología contemporánea ha llevado a los estudiosos a abandonar, desde hace varias décadas, la ingenua creencia del positivismo decimonónico en la posibilidad de llegar a captar la realidad en forma plena y segura, que abarque totalmente su complejidad y que, a la vez, garantice la ausencia de error. Hoy sabemos que es imposible superar en forma absoluta las limitaciones del conocimiento humano; que es imposible también, en consecuencia, eliminar de manera absoluta la opacidad de la realidad.¹⁴

Lo anterior no debe entenderse en el sentido de adoptar una postura relativista, igualmente absoluta. Puede haber trabajos que logren aproximaciones muy adecuadas y consistentes a la realidad, en tanto que otros pueden quedar muy lejos de ese ideal. De

13 Sobre la cual puede verse la obra de Halpern, Edward S. y Th. A. Schwandt, *Linking Audit and Metaevaluation*, Applied Social Research Methods Series, Vol. II, Beverly Hills, Sage, 1988. Pueden verse comentarios a este trabajo de Lincoln, Yvonna S. y Elton G. Guba, *op. cit.*, pp. 316-318 y de Martínez Rizo, Felipe, *El oficio del investigador educativo*, *op. cit.*, pp. 217-231.

14 En relación con esta idea pueden verse los siguientes trabajos del autor, en los que se desarrolla con mayor amplitud: Martínez Rizo, Felipe, *El oficio del investigador educativo*, *op. cit.*, pp. 119-135; y "La metodología de la investigación y los límites del conocimiento humano", *Caleidoscopio*, número 1, enero-junio 1997, pp. 95-111.

hecho los hay, de unos y de otros, y no es demasiado difícil reconocerlos.

No podemos llegar directamente a la verdad, pero sí podemos aproximarnos cada vez más a ella, como también podemos alejarnos de tan anhelado destino. Y la manera de lograr aproximación en vez de distanciamiento es mediante estudios que apliquen de la manera más rigurosa los diversos mecanismos de control de la calidad de la investigación que la práctica de los mejores académicos de todos los enfoques ha ido desarrollando a lo largo del tiempo.

Los esquemas anteriormente presentados, que sintetizan la manera en que los enfoques cuantitativos y cualitativos responden las cuatro preguntas clave relativas a la calidad de la investigación, permiten apreciar un alto grado de coincidencia entre las dos concepciones de la investigación mencionadas, que a veces se visualizan como completamente opuestas, y muestran que es posible sistematizar los medios necesarios para asegurar la calidad de un estudio particular.

Los estándares de calidad de tests y pruebas psicológicas y educativas, como los de los procedimientos de evaluación de programas, que las instituciones más prestigiadas a nivel internacional han desarrollado son, precisamente, formas particulares de sistematizar ese tipo de mecanismos de control de calidad que la investigación ha desarrollado. Será su aplicación lo que haga que los instrumentos de diagnóstico permitan obtener resultados «dignos de fiar», sobre los cuales pueda apoyarse responsablemente la toma de decisiones y el establecimiento de políticas. ☼

