

APUNTES SOBRE METODOLOGÍA DE LA INVESTIGACIÓN

LA CONSULTA A EXPERTOS COMO ESTRATEGIA PARA LA RECOLECCIÓN DE EVIDENCIAS DE VALIDEZ BASADAS EN EL CONTENIDO

Arturo Barraza Macías
tbarraza@terra.com.mx

Coordinador del Programa de Investigación de la Universidad Pedagógica de Durango

La validez es uno de los temas cruciales en el proceso de construcción de un test, una escala, un inventario o un cuestionario, ya que se requiere comprobar la utilidad de la medida realizada, es decir, el significado de las puntuaciones obtenidas. Es precisamente la validez la que permitirá realizar las inferencias e interpretaciones correctas de las puntuaciones que se obtengan al aplicar un test y establecer la relación con el constructo/variable que se trata de medir.

Samuel Messick (1989), uno de los autores más influyentes en el estudio de la validez, define el concepto en los siguientes términos:

“La validez es un juicio evaluativo global del grado en el que la evidencia empírica y la lógica teórica apoyan la concepción y conveniencia de las inferencias y acciones que se realizan basándose en las puntuaciones que proporcionan los test u otros instrumentos de medida”. (p. 19)

Vale la pena destacar dos aspectos significativos de esta definición:

- No se valida el test o el instrumento, sino las inferencias e interpretaciones realizadas a partir de las puntuaciones que proporciona el test.
- La validez es un resumen de evidencias y consecuencias sobre la utilización del test.

Visión histórica

El concepto de validez ha ido modificándose a lo largo del tiempo. En un principio, antes de los años 50, la validez de un instrumento de medición, de carácter psicológico, educativo o sociológico, era evaluada por una diversidad de procedimientos.

El tipo de evidencia utilizado para demostrar la validez del test variaba con el propósito del test, la orientación teórica del test y el tipo de datos disponibles (Anastasi, 1986), en ese sentido se hablaba de validez aparente, validez intrínseca, por definición, validez lógica, validez factorial, etc. Este tipo de estudios de validez eran marcadamente pragmáticos, empíricos e incluso ateóricos. (González, 2006)

Debido a esta gran diversidad de concepciones, la American Psychological Association (APA), asociada con la Association for Applied Psychology (AAP), organizó un comité, en 1950, para establecer un código de ética que tuviera en cuenta aspectos científicos y aplicados. En 1954 se

publica el resultado de su trabajo en el documento titulado *Technical recommendations for psychological test and diagnostic techniques* (citado por Pardo, 2006), en donde aparecen cuatro categorías de validez:

- Contenido
- Predictiva
- Concurrente
- Constructo

Es en este documento donde aparece por primera vez la validez de constructo. En esta primera publicación, de lo que más tarde constituirían los estándares, se pretendía establecer algún orden en la práctica de la construcción de instrumentos de medición.

En los estándares de 1974, que la APA publica en colaboración con la American Educational Research Association (AERA) y el National Council on Measurement in Education (NCME), la validez predictiva y concurrente se subsumen en criterial o de criterio. Queda así configurada la estructura tripartita de la validez (contenido, criterio y constructo). Este momento es identificado como la primera etapa de los estándares y se denomina operacional (Elosua, 2003).

Esta concepción tripartita de la validez se encuentra presente en la mayor parte de los manuales de investigación (v.gr. Hernández, Fernández y Baptista, 2006). Sin embargo, dicha concepción ha sido cuestionada fuertemente en dos direcciones:

- Esta forma conceptualizar la validez es fragmentaria e incompleta, debido especialmente a que no tiene en cuenta la evidencia, tanto de las implicaciones valorativas del significado del puntaje como una base para la acción, como las consecuencias sociales del uso del puntaje (Pardo, 2006).
- Esta concepción acerca de la validez ha supuesto dos peligros para el constructor de instrumentos de medición: 1) pensar que se trata de tres tipos de validez distintos y 2) pensar que en un estudio de validez es imprescindible recoger evidencia sobre los tres aspectos de la misma (Gómez y Hidalgo, 2002).

Estas críticas han ocasionado que los psicómetras vuelvan sus ojos a otras alternativas, destacando entre ellas: 1) la validez de constructo como concepto incluyente, 2) la acumulación de evidencias de validez y 3) las consecuencias del uso de los test.

1.- Validez de constructo como concepto unitario

Hacia mediados de los años cincuenta, Cronbach y Meehl (en Gómez y Hidalgo, 2002) publican un trabajo sobre este tema, entendiendo que es uno de los aspectos más importantes, donde la comprobación de la validez de constructo implica la recogida de distintas evidencias, no todas de ellas cuantitativas, y la integración de la información recogida.

La idea que se empieza a fraguar es que el proceso de validación no difiere del proceso de construcción de teorías científicas y algunas de las estrategias para investigar la validez de constructo tienen el mismo fin. Estos autores mencionan como estrategias: estudiar diferencias entre grupos, correlaciones entre subtests, estudios sobre la estructura interna del test, estudios factoriales, correlaciones con criterios externos, estudios longitudinales, estudios experimentales y estudios sobre el análisis del contenido.

La aportación de Campbell y Fiske, (en González, 2006) se centra en la introducción de los conceptos de validez discriminante y convergente, y en la incorporación de la matriz multimétodo-multirrasgo para evaluarlas: Esta técnica permite examinar el grado en que el instrumento de medición arroja una mediada adecuada del constructo.

Messick (en Gómez y Hidalgo, 2002) defiende que el concepto de validez de constructo es un concepto más general que los de validez predictiva o concurrente, que son específicos de los criterios externos utilizados así como de los grupos evaluados, y que la validez de contenido no es una propiedad de las respuestas obtenidas sino del test construido.

Este autor concluye que toda la evidencia de validez debe ser de constructo. En los años 80 numerosos trabajos enfatizan la importancia de la validez de constructo, siendo la esencia misma de todo proceso de validación, además se apunta que la validez es una, y que no podemos hablar de distintos tipos de validez, sino que todo es validez de constructo.

Los Estándares de 1985 (APA, AERA y NCME, 1985), mencionan ya que los distintos tipos de validez (criterial, de contenido, factorial, discriminante, etc.) son distintas formas de expresión de la validez de constructo, y cualquiera de las primeras contribuye a expresar parte de la última. También recogen otra idea gestada durante estos años, a saber, lo validado no es el test mismo sino una interpretación de los datos obtenidos por un determinado procedimiento, por lo que la validez de las puntuaciones del test deben ser establecidas en cada uso que se haga del mismo.

Este momento, signado por el concepto unitario de validez y el papel central de la teoría en el proceso, es identificado por Elosua, (2003) como la segunda etapa de los estándares y es denominado estadio teórico.

2.- Acumulación de evidencia

La evolución que sufre la validez durante los años ochenta se refleja en la definición dada en la edición de 1985 de los Standards of educational and psychological testing (APA, AERA y NCME, 1985): La validez es la consideración más importante en la evaluación de un test.

El concepto se refiere a la adecuación, significado y utilidad de las inferencias específicas hechas con las puntuaciones de los tests. La validación de un test es el proceso de acumular evidencia para apoyar tales inferencias.

Una variedad de evidencias pueden obtenerse de las puntuaciones producidas por un test dado, y hay muchas formas de acumular evidencia para apoyar una inferencia específica. La validez, sin embargo, es un proceso unitario.

Aunque la evidencia puede ser acumulada de muchas formas, la validez se refiere siempre al grado en que esa evidencia apoya las inferencias que se hacen a partir de las puntuaciones. La acumulación de evidencia por distintas fuentes es una idea desarrollada de manera paralela al concepto unitario de validez y al papel central de la teoría en el proceso, por lo cual también es integrada a la segunda etapa de los estándares denominad estadio teórico (Elosua, 2003).

3.- Las consecuencias del uso de los instrumentos de medición

La posición de Messick (en Gómez y Hidalgo, 2002) en cuanto a la validez contempla un aspecto más referido a los valores sociales y las consecuencias éticas. Este autor afirma que la validez no es del instrumento de medición o de la observación, lo que se valida son las inferencias derivadas de las puntuaciones del instrumento o de otros indicadores, inferencias sobre el significado de las puntuaciones o la interpretación para propósitos aplicados y sobre las implicaciones para la acción, es decir, las consecuencias sociales y éticas.

El debate actual se centra en la cuestión de si la investigación de las posibles consecuencias de la administración y uso de los instrumentos de medición debe incluirse como una parte más del plan de validación de un instrumento, es decir, si la validación es un proceso científico o sociopolítico (Crocker, 1997). Este es precisamente el contenido del número especial de la revista *Educational Measurement: Issues and Practices* (vol. 16, n° 4, 1997).

Las posturas están encontradas y, aunque todos asumen o destacan la importancia de las consecuencias sociales del uso de los instrumentos de medición, disienten en si deben ser valoradas como parte de la validez del instrumento y uso del mismo o, por el contrario, deben ser valoradas por aquellos que desarrollan los instrumentos pero no incluidas en la validez del mismo. Esta discusión es el sello distintivo de la etapa actual de los estándares y es denominada contextual por Elosua (2003).

Más allá de las tres líneas alternativas de discusión esbozadas en los párrafos anteriores, puedo afirmar, basándome en la consulta realizada y coincidiendo con Gómez y Hidalgo (2002), que actualmente existe un consenso alrededor de los siguientes puntos:

- Lo que se valida no es el instrumento de medición sino las puntuaciones que arroja, y por lo tanto la pregunta que se trata de responder es ¿es válido el uso o la interpretación de las puntuaciones de este instrumento?
- La validez no se puede resumir en un sólo indicador o índice numérico, al igual que ocurría con la confiabilidad (coeficiente de confiabilidad, error de medida, función de información, etc.), sino que la validez de las puntuaciones de un instrumento se asegura mediante la acumulación de evidencia teórica, estadística, empírica y conceptual del uso de las puntuaciones
- Una puntuación puede ser válida para un uso y no para otro
- La validación es un proceso continuo y dinámico y
- La teoría juega un papel muy importante como guía tanto del desarrollo de un instrumento de medición como de su proceso de validación.

Tipos de evidencia

Aunque Messick (1989) aboga por un concepto unitario de validez, y esta concepción ha sido adoptada por la comunidad científica, como queda recogido en los últimos estándares publicados (AERA, APA, NCME, 1999), él también señaló que diferentes tipos de inferencias con los instrumentos de medición requieren distintos tipos de evidencia. Estos tipos de evidencia pueden obtenerse estudiando el contenido del instrumento en función de los contenidos del dominio de referencia, examinando las relaciones entre las respuestas a las tareas, ítems y/o partes del instrumento, estudiando las relaciones entre las puntuaciones del instrumento y otras medidas, investigando las diferencias a través de los grupos o sobre el tiempo, y estudiando las respuestas de los sujetos a tratamientos experimentales, entre otras aproximaciones.

La evidencia de la validez surge de múltiples fuentes, sin embargo, siguiendo a Messick (1995), las diferentes estrategias de validación se pueden agrupar en cinco tipos de evidencias (Pardo, 2006):

1) Evidencia basada en el contenido de un instrumento de medición: incluye evidencia de la relevancia, representatividad y calidad técnica del contenido de la prueba.

La relevancia de las preguntas se evidencia en el enfoque particular utilizado como marco conceptual de la prueba, en donde deben aparecer los criterios particulares para cada una de las variables empleadas. Esto es, que debe proporcionarse la relación entre la variable empleada (y en últimas el ítem utilizado) y la teoría.

La representatividad se observaría a través de la relación entre el énfasis de las variables en el instrumento y el énfasis de elementos correlativos en el marco conceptual.

La calidad técnica del contenido de la prueba, se evidenciaría a través del proceso de análisis de ítem y de prueba. Este análisis involucra aspectos cualitativos y cuantitativos al interior de algún marco de análisis particular.

2) Evidencia basada en el proceso de respuesta: argumentos que evidencien las consistencias entre las respuestas al instrumento (evidencia empírica) y los procesos que, desde la teoría, se plantea que asumen los evaluados en las tareas propuestas.

Los argumentos relacionados con este aspecto se encuentran, definitivamente, en el marco conceptual. Es evidente que un marco conceptual puede tomar diversas formas pero, sin importar desde qué perspectiva se elabore, debe contener un párrafo en donde se expliciten los procesos que seguirán los evaluados para llegar a las respuestas que se le plantean.

3) Evidencia basada en la estructura interna: se valora la relación entre los ítems de un instrumento y de ellos con los componentes de la prueba para determinar si todos ellos conforman el constructo establecido en el marco conceptual, ya sea este unidimensional o multidimensional.

Es claro que aquí se analizan los elementos del constructo en su más íntima estructura. Como consecuencia se tiene que el tipo de escala seleccionada para emitir los puntajes debe estar de acuerdo con los planteamientos del constructo.

En general, un instrumento de medición tiene como propósito fundamental la generación de puntuaciones (un solo resultado para un conjunto de ítems). El término puntaje se utiliza genéricamente en su sentido más amplio para indicar cualquier codificación de consistencias observadas o regularidades en la ejecución de una prueba, cuestionario, procedimiento de observación o cualquier otro instrumento de evaluación (tales como muestras de trabajos, portafolios o simulación de problemas reales).

Este uso subsume procedimientos tanto cuantitativos como cualitativos. Adicionalmente se debe tener en cuenta, al construir un instrumento, su dimensionalidad. Un instrumento puede tener cualquier número de dimensiones. No obstante, la mayoría de ellos se construyen con pocas dimensiones, o una sola.

4) Evidencia basada en otras variables: incluye evidencia convergente y discriminante a partir de comparaciones multimétodo-multirrasgo. Se examina el grado en el cual las propiedades de los

puntajes y las inferencias se generalizan a grupos poblacionales, de acuerdo con las necesidades del proceso de medición, de evaluación o de la investigación que se adelante.

5) Evidencia basada en las consecuencias de la medición: valora las implicaciones de las interpretaciones de los puntajes como una base para la acción futura; igualmente, las consecuencias presentes y futuras del uso del instrumento especialmente a partir de estudios de sesgos y equidad.

Estos cinco tipos de evidencias son organizados por Elosua (2003) en dos grandes rubros: fuentes de evidencia interna (ítems) y fuentes de evidencia externa (instrumento), como se observa en el siguiente cuadro:

Evidencia	Tipo
Interna (ítems)	Contenido
	Proceso de respuesta
	Estructura interna (dimensionalidad)
Externa (instrumento)	Relaciones con otras variables
	Consecuencias

Cuadro No. 1 Fuentes de evidencia (síntesis del cuadro presentado por Elosua, 2003)

Evidencia basada en el contenido

“El análisis del contenido aglutina dos tipos de estudios suplementarios. Unos encaminados a evaluar las relaciones entre el constructo y el contenido del test, y otros dirigidos a valorar los factores contextuales internos y externos que puedan añadir varianza no deseada” (Elosua, 2003).

El objetivo de este tipo de estudios es obtener evidencias para determinar si los ítems del instrumento representan adecuadamente el constructo que se pretende medir. “No olvidemos que los ítems son el “estímulo” necesario para obtener las muestras de conducta indicadoras del constructo. Este tipo de estudio examina si los ítems pueden hacer bien su papel” (González, 2006: 16).

Para lograr este objetivo, la literatura, en el campo de la metodología de la investigación, marca tradicionalmente cuatro estrategias: la fundamentación en una teoría, la consulta a expertos, el análisis de errores y la supervisión de los resultados totales (Anastasi y Urbina, 1998). De estas cuatro estrategias son las dos primeras las más recurrentes y reconocidas.

En el primer caso, la fundamentación teórica, la evidencia basada en el contenido la realiza el mismo investigador apoyándose en la teoría y los supuestos que guían la selección de los ítems (Briones, 2001); en ese sentido “el dominio de contenido de una variable, normalmente, es definido o establecido por la literatura (teorías y estudios antecedentes) (Hernández, et. al.; 2006: 279).

La forma en que esta estrategia se opera presenta ciertas ambigüedades e imprecisiones: algunos autores hacen referencia al análisis racional de los contenidos posibles (Muñiz, 2003), otros a la operacionalización de las variables (Castañeda, De la Torre, Morán y Lara, 2005), y otros la articulan a una estrategia global que integra: fundamentación teórica, consulta a expertos y análisis estadístico (Hernández, et. al. 2006).

El segundo caso, la estrategia de consulta a expertos, consiste en preguntar a uno o varios expertos si los reactivos evalúan lo que uno quiere que evalúen (Salkind, 1999). Esta estrategia es planteada como

complementaria o simultánea de la fundamentación teórica por algunos autores (Briones, 2001; Hernández, et. al. 2006; y Muñiz, 2003), mientras que otros plantean su uso solamente en caso de que el constructo o variable de estudio sea definido de manera aproximada o imprecisa, esto es, sin una buena fundamentación teórica (Gregory, 2001).

Particularmente, en el caso de esta estrategia, coincido plenamente con Gregory (2001), por dos razones: primero, por que es necesario recordar que el aspecto central de este tipo de evidencia es la relevancia, entendida ésta como la relación de la variable, y los ítems, con la teoría utilizada (Pardo, 2006), y segundo, por que de esta manera se supera la crítica que considera a este tipo de evidencia como algo subjetivo (González, 2006). Sin embargo el uso tan extendido de la consulta a expertos como fuente de evidencia de contenido, y los problemas estructurales que acarrea, hacen que en el presente artículo esboce una propuesta al respecto.

Un procedimiento para la consulta a expertos

De los diferentes autores que abordan la consulta a expertos (Briones, 2001; Hernández, et. al. 2006; Muñiz, 2003 y Salkind, 1999) solamente uno de ellos proporciona un procedimiento en específico que permite llegar a emitir un juicio cuantitativo sobre la validez de contenido (Gregory, 2001).

Esta forma de abordar la consulta a expertos se toma como base para elaborar un procedimiento que se apoya en dos supuestos básicos:

- A la validez se le puede caracterizar en un continuo que va de débil a aceptable y de aceptable a fuerte.
- La consulta a expertos se utiliza básicamente para establecer si los ítems del instrumento representan adecuadamente el constructo que se pretende medir.

Estos supuestos hacen que rechace, lo que he dado en llamar, el autoritarismo de los expertos que se manifiesta en dos vertientes: a) el hecho de evaluar en términos dicotómicos cada ítem del cuestionario: pertenece o no pertenece al constructo (en este sentido establezco una sana distancia de la propuesta de Gregory, 2001) y b) el derecho a una evaluación global del instrumento que implica desde aspectos de redacción, hasta la evaluación de la pertinencia, la coherencia o la calidad.

El procedimiento que propongo en este artículo es el siguiente:

- a) Primer paso: se elabora el instrumento con las indicaciones y la escala que se muestra en el siguiente ejemplo:

El presente cuestionario tiene como objetivo central reconocer las características del estrés que suele acompañar a los alumnos de educación media superior y superior durante la realización de sus estudios. Le pedimos su colaboración para establecer la validez de contenido, en caso de aceptar le sugerimos seguir las siguientes instrucciones:

- Lea cuidadosamente la especificación de cada una de las secciones de este cuestionario, ya que cada una representa diferentes dimensiones de estudio.
- A continuación indique que tan bien considera usted que el ítem o reactivo sea parte de la dimensión de estudio que se aborda en la sección
- Juzgue cada ítem o reactivo de manera individual comparando el contenido del ítem con el contenido de la dimensión de estudio.
- Para evaluar utilice la siguiente escala

- 0 el ítem **no pertenece** a la dimensión de estudio.
- 1 el ítem **probablemente no pertenece** a la dimensión de estudio.
- 2 el ítem **probablemente si pertenece** a la dimensión de estudio.
- 3 el ítem **si pertenece** a la dimensión de estudio.

Primera dimensión de estudio: estresores.

El estresor es el estímulo o situación, que es valorada como amenazante por la persona, y que desencadena en ella las reacciones que son conocidas habitualmente como estrés. En el siguiente listado se ofrecen una serie situaciones del ámbito escolar que, en determinado momento y en función de cada sujeto, pueden llegar a ser considerados como estresores. Evalúe de manera individual cada ítem para determinar si pertenece o no a la dimensión de estudio denominada “estresores”; señale con una X la opción que considere más adecuada.

	(0) No pertenece	(1) Probablemente no pertenece	(2) Probablemente si pertenece	(3) Si pertenece
La competencia con los compañeros del grupo				
Exceso de responsabilidad personal por cumplir con las obligaciones escolares				
Sobrecarga de tareas y trabajos escolares				
La personalidad y el carácter del profesor				
Las evaluaciones de los profesores (exámenes, ensayos, trabajos de investigación, etc.)				
El tipo de trabajo que te piden los profesores (consulta de temas, fichas de trabajo, ensayos, mapas conceptuales, etc.)				
No entender los temas que se abordan en la clase				
Participación en clase (responder a preguntas, exposiciones, etc.)				
Tiempo limitado para hacer el trabajo				

- b) Una vez elaborado el instrumento se le solicita su llenado a los expertos; a diferencia de Gregory (2001) considero que el mínimo de expertos a participar es de cinco, mientras que el máximo sería de 10, considerando en estricto sentido lo que es un experto. En este punto difiero con aquellos procesos que consideran un experto al profesionista del campo disciplinario donde se integra el constructo (v.gr. si el constructo es estilo de aprendizaje cualquier psicólogo de formación es un experto), en contraparte creo que un experto debe tener como mínimo estudios de maestría en el campo disciplinario donde se inserte su constructo y estar realizando investigación.

- c) Una vez llenados los instrumentos se observa la media de cada ítem y en caso de que la misma sea menor a 1.5 el ítem se elimina.
- d) Con los ítems restantes se obtiene una media general del instrumento o de cada una de las dimensiones que lo componen y el resultado se interpreta con la siguiente escala:
 - De 1.6 a 2.0 se considera que presenta una validez débil.
 - De 2.1 a 2.5 se considera que presenta una validez aceptable
 - De 2.6 en adelante se considera que presenta una validez fuerte.

A manera de cierre

El procedimiento que he presentado no evita que la consulta a expertos siga siendo considerada como una validez con alto grado de subjetividad, sin embargo, el hecho de precisar el procedimiento y establecer una forma de cuantificar y valorar el puntaje final obtenido permite evitar muchos de los problemas asociados en la práctica a la recolección de evidencia mediante este procedimiento.

Por otra parte reitero mi idea de que la consulta a experto solamente se debe de emplear cuando el constructo no cuenta con una buena fundamentación teórica y esto hace que sea definido de forma aproximada o imprecisa.

Referencias

- American Psychological Association, American Educational Research Association, and National Council on Measurement in Education (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- American Psychological Association, American Educational Research Association, and National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Anastasi, Anne (1986), *Evolving concepts for test validation*, en Annual Review of Psychology, No. 37, pp. 1-15.
- Anastasi Anne y Urbina Susana (1998), *Test psicológicos*, México, Prentice Hall.
- Briones Guillermo (2001), *Métodos y técnicas de investigación para las ciencias sociales*, México, Trillas.
- Castañeda Jiménez Juan, De la Torre Lozano María Olivia, Morán Rodríguez José Manuel y Lara Ramírez Luz Patricia, (2005), *Metodología de la Investigación*, México, Mc Graw Hill.
- Crocker, L. (1997), *Editorial: The great validity debate* en Educational Measurement: Issues and Practice, Vol. 16, No. 4.
- Elosua Oliden Paula (2003), *Sobre la validez de los test*, en la Revista Psicothema, Vol. 15, No. 2, pp.315-321.
- González Gómez Andrés (2006), *Introducción a la Teoría de la validez*, disponible en <http://www.ugr.es/~andredito/tema7.doc> (recuperado noviembre de 2006)
- Gómez Benito Juana y Hidalgo Montesino María Dolores (2002), *La validez de los test, escalas y cuestionarios*, en la Revista Electrónica “La sociología en los Escenarios”, No. 8.
- Gregory Robert J. (2001), *Evaluación psicológica. Historia, principios y aplicaciones*, México, El Manual Moderno.
- Hernández Sampieri Roberto, Fernández Collado Carlos y Baptista Lucio Pilar (2006), *Metodología de la Investigación*, México, Mc Graw Hill.

- Messick, S. (1989), "Validity", en *Educational Measurement*, de R.L. Linn (Ed.), New York: American Council on Education and Macmillan publishing company.
- Messick, S. (1995), *Standards of validity and the validity of standards in performance assessment*, en *Educational measurement: Issues and Practice*, No. 14, pp. 5-8.
- Muñiz José (2003), *Teoría Clásica de los Test*, Madrid, España, Ediciones Pirámide.
- Pardo Adames Carlos A. (2006), *Hablemos de validez*, Taller del Instituto Colombiano para el Fomento de la Educación Superior, disponible en <http://w3.icfes.gov.co:8080/seminariointernacional/ktmlite/files/uploads/CARLOS%20PARDO%20-%20ICFES.pdf>
- Salkind Neil J. (1999), *Métodos de Investigación*, México, Prentice Hall.

INVESTIGACIÓN EDUCATIVA