

LA VALIDEZ DE LOS DISEÑOS EXPERIMENTALES
SEGUN COOK Y CAMPBELL

Alfredo Estraño.
Jefe del Depto. de Post-Grado y Reciclaje.
I.U.P.E.R.A.E.L.

RESUMEN

El presente artículo discute el problema de la validez de los diseños de investigación a la luz de los recientes trabajos realizados por Thomas D. Cook y Donald T. Campbell. En el mismo se plantean cuatro tipos de validez: Validez de Conclusión estadística, validez interna, validez de constructo de causa-efecto y validez externa.

El objetivo de toda investigación es conocer la verdad, por lo que es necesario estar atento a todos aquellos factores que nos puedan alejar de ella. En este sentido, se ha introducido el término validez para referirse al grado o aproximación a la verdad. Asimismo, el término invalidéz se emplea para el caso contrario, es decir, para señalar el grado o aproximación a la falsedad.

El concepto de validez fué introducido en 1963 por Campbell y Stanley, quienes la dividieron en dos tipos: validez interna y validez externa (1). En 1979, Cook y Campbell introdujeron dos nuevos tipos de validez, con lo cual establecen cuatro tipos: validez de conclusión estadística, validez interna, validez de constructo de causa-efecto y validez externa (2).

En los siguientes párrafos se discutirán estos tipos de validez.

Validez de conclusión estadística: se refiere a la validez aproximada que permite afirmar que existe relación entre una o más variables independientes y una o más variables dependientes. En otras palabras, que las variables covarían.

Validez interna: se refiere a la validez aproximada que permite afirmar que existe una relación causal de una variable a otra, a partir de una manipulación o medición de las variables. Este tipo de validez nos responde a la pregunta: ¿existe relación causal entre las variables? En otras palabras, si hay covariación entre las variables.

Validez de constructo de causa-efecto: se refiere a la validez aproximada por medio de la cual se pueden hacer generalizaciones acerca de constructos de orden superior a partir de las operaciones de investigación.

El uso de este término se extiende a las variables independientes manipuladas así como a los rasgos de personalidad medidos (variables dependientes).

Dado que hay relación causal, ¿cuáles son los constructos de causa-efecto envueltos en dicha relación?

Validez externa: se refiere a la validez aproximada mediante la cual las conclusiones obtenidas son generalizables a toda la población, a cualquier situación y en cualquier época.

Dado que existe una relación causal entre dos constructos A y B, ¿se puede generalizar esta relación a toda la población, situaciones y épocas?

Validez de conclusión estadística

Aspectos a considerar al evaluar estadísticamente un experimento:

- 1.- Si el estudio es lo suficientemente sensible para permitir afirmaciones razonables acerca de covariación.

- 2.- En caso afirmativo, ¿hay alguna evidencia razonable desde la cual se pueda inferir que la supuesta causa y efecto covarían?
- 3.- Si esa evidencia existe, ¿con qué fuerza covarían las dos variables?

El primer aspecto se refiere al poder de la prueba estadística. En suma, la validez de conclusión estadística está íntimamente relacionada con la prueba de significación estadística. Aunque debe complementarse con la estimación de magnitudes de covariación, donde los intervalos de confianza desempeñan un importante papel descriptivo.

Factores que amenazan la validez de conclusión estadística.

- a) El bajo poder estadístico de la prueba.
- b) La violación de los supuestos de los tests estadísticos.
- c) La rata de error por experimento (número esperado de errores por experimento).
- d) La confiabilidad de las medidas.
- e) La confiabilidad en la implementación del tratamiento: variación de persona a persona o en una misma persona a través del tiempo.
- f) Los factores irrelevantes en la situación experimental que pueden aumentar la variación de error.
- g) La heterogeneidad de los respondientes debida al azar.

Validez Interna.

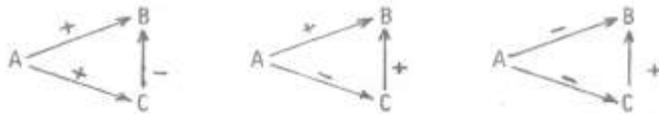
La validez de conclusión estadística nos ayuda a establecer si dos variables covarían; el problema que sí ue es determinar si esa realación es causal y la dirección de esa causalidad. (A causa B ó B causa A).

El establecimiento de la dirección de la causalidad se facilita cuando hay una secuencia temporal claramente identificable, por ejemplo, en estudios experimentales.

El establecimiento de la dirección de causalidad se dificulta cuando existe una 3ra. variable: A C B.

La consideración de una 3ra. variable es la esencia de la validez interna.

Situaciones:



Factores que disminuyen la validez interna:

- 1.- Historia: cuando el efecto se debe a un fenómeno que sucedió entre el pre-test y el post-test y no al tratamiento.
- 2.- Madurez: crecimiento del respondiente.
- 3.- Familiaridad con pruebas de medida.
- 4.- Instrumentación: cambio en el instrumento de medida entre el pre y el post-test.
- 5.- Regresión estadística: tendencia hacia la medida cuando los grupos experimentales son formados sobre la base de los puntajes del pre-test.
- 6.- Selección: diferencias entre los grupos control y experimental.
- 7.- Mortalidad: abandonos que afectan a la variable de pendiente.
- 8.- Interacciones con selección: ejemplos:
 Selección - maduración.
 Selección - historia.
 Selección - instrumentación.

- 9.- Ambigüedad en la dirección causal común en estudios correlacionales.
- 10.- Difusión o imitación de los tratamientos ("Efecto Hawthorne").
- 11.- Rechazo a los tratamientos que crean desigualdades.
- 12.- Mayor esfuerzo (rivalidad compensatoria) del grupo que recibe el tratamiento menos deseable (grupo control). Ejemplo: "efecto John Henry".
- 13.- Desmoralización del grupo que recibe el tratamiento menos deseable.

Papel de la aleatorización en la validez interna de diseños experimentales y cuasi-experimentales.

La aleatorización minimiza casi todos los factores que amenazan la validez interna de un diseño experimental. Cuando los sujetos son asignados al azar a los grupos, cada grupo está similarmente constituido, hablando en términos de promedio. No hay problemas de selección, de maduración ni de selección-maduración. No ocurre lo mismo con un diseño cuasi-experimental, donde el investigador tiene que detectar todas las amenazas y luego eliminarlas una a una. De aquí que la experimentación al azar facilite las inferencias causales.

Factores que la aleatorización no elimina:

- a) La imitación de los tratamientos.
- b) El deseo de un tratamiento igualitario.
- c) La rivalidad compensatoria.
- d) La desmoralización de los grupos, especialmente, el grupo control.
- e) La mortalidad diferencial, por lo general como consecuencia de los tratamientos.

Respecto a la mortalidad diferencial, es recomendable hacer estudios empíricos que analicen el efecto de retiros diferenciales que incluyan: 1) estudio de los sujetos que respondieron el post-test, 2) comparación con el pre-test de aquellos que respondieron el post-test.

Validez de constructo de causa-efecto.

Establecida la relación causal, interesa conocer cuáles son los constructos envueltos en esa relación causal. Este tipo de validez trata de "eliminar" la confusión que pueda existir en una relación causal:

$$\text{¿}A \xrightarrow{+} B \text{ ó } A \xrightarrow{+} Y \text{ ó } X \xrightarrow{+} B?$$

Ejemplos de problemas relacionados con validez de constructos de causa-efecto.

- 1.- El efecto placebo en experimentos médicos con drogas.
- 2.- El efecto "Hawthorne", otro fenómeno que produce incertidumbre acerca de cómo deben etiquetarse las operaciones de causa-efecto.

¿Cómo mejorar la validez de constructos de causa-efecto.

La delimitación y precisión de los constructos en juego debe comenzar desde la etapa de planificación y de prueba piloto de una investigación. Ello incluye:

- 1.- Definir claramente los constructos de manera que haya conformidad con el uso que el público da a las palabras.
- 2.- Analizar los datos considerando:
 - 2.1. ¿en qué medida las variables independientes alteran lo que pretenden alterar (manipulación y medida)?
 - 2.2. ¿en qué medida una variable independiente no varía con la medida de otros constructos diferen-

tes pero relacionados? Ejemplo: habilidad para comunicarse con los reportes de los sujetos y no con otros tales como poder.

2.3. ¿en qué medida las variables dependientes propuestas encajan en los factores que pretenden medir (inter-item correlación)? y,

2.4. ¿en qué medida las variables dependientes están determinadas por factores irrelevantes que las convierten en medidas aproximadas de lo que se intentó?

Este procedimiento depende de dos métodos:

- 1.- La prueba de la convergencia a través de diferentes medidas y manipulación de la misma variable.
- 2.- La prueba de la divergencia entre medidas y manipulación de variables relacionadas pero conceptualmente distintas.

Debe clarificarse que no existen definiciones comunes de todos los constructos.

Aunque esto no es necesariamente malo puesto que ello nos obliga a buscar constructos más confiables a través de sucesivos y exitosos tests. Hasta ahora, los investigadores han tratado de establecer validez de constructo de los efectos más que de las causas.

Factores que amenazan la validez de constructos de causa-efecto.

Los siguientes factores tienen que ver con poca representatividad del constructo y con excesiva representatividad del constructo mediante elementos irrelevantes:

- 1.- Explicación preoperacional inadecuada de los constructos.
- 2.- Sesgo debido a una sola operación del constructo; ejemplo: una variable independiente (causa) y una sola medida de variable dependiente (efecto). Es-

to puede sub-representar al constructo y contener irrelevancias.

- 3.- Sesgo por un sólo método; ejemplo: medición de actitudes sólo con test de lapiz y papel.
- 4.- Adivinanza de la (s) hipótesis planteada (s) que hace que el sujeto actúe como se espera. Solución: hacer que las hipótesis sean difíciles de adivinar.
- 5.- Temor a ser evaluado por expertos en ajustes de la personalidad o por expertos en descripción de las destrezas humanas.
- 6.- Expectativas del experimentador; recomendación: emplear experimentadores objetivos o usar personas diferentes a los que hicieron el tratamiento para las tareas de corrección.
- 7.- Constructos y niveles de constructos. Se refiere a la manipulación de varios niveles discretos de una variable independiente que es continua. Como consecuencia, se podría concluir que A no afecta B cuando de hecho A-nivel uno no afecta pero sí a nivel cuatro si el experimenta hubiese llegado a este nivel.
- 8.- Interacción de diferentes tratamientos. Esto ocurre cuando los sujetos son sometidos a más de un tratamiento.
- 9.- Interacción tratamiento y medición. ¿A qué situaciones de medición pueden generalizarse los resultados? Los post-tests repetidos pueden servir para verificar si se obtienen los mismos resultados.
- 10.- Generalización restringida a través de constructos. Es necesario explorar con otras personas como un tratamiento puede influenciar otros constructos diferentes a las iniciales.

Validez Externa.

Los autores usan este concepto para referirse a los

mismos factores que según Campbell y Stanley afectan la generalización de las conclusiones, excepto lo referente a generalizaciones a través de constructos de causa efecto, que lo incluyen como sesgo de una sola operación que fue incluido en la validez de conclusión de causa-efecto. Cook y Campbell usan el término validez externa para referirse a generalizaciones a través de personas, situaciones y épocas (de un mismo grupo o de diferentes grupos).

Es bueno hacer notar que las muestras accidentales de conveniencia no facilitan la inferencia a la población objeto de estudio ni tampoco permiten clarificar a qué población nos referimos.

Generalizar a la población en cuestiones diferentes a generalizar a otras poblaciones, por ejemplo, a varones y hembras.

Factores que amenazan la validez externa.

Someter a prueba estos factores no es más que efectuar tests de las interacciones estadísticas. Por ejemplo: interacción tratamiento educacional-clase social, para determinar si los mismos resultados se aplican a todas las clases sociales. A continuación se señalan los factores que amenazan la validez externa:

- 1.- Interacción selección y tratamiento. ¿A qué categorías o grupos de personas pueden generalizarse las relaciones de causa-efecto?
- 2.- Interacción situaciones experimentales y tratamiento. ¿Se obtendrán los mismos resultados si el experimento se realiza en otro ambiente diferente al original?
- 3.- Interacción historia-tratamiento. ¿A qué períodos del pasado y del futuro puede generalizarse una relación causal particular? Por ejemplo, durante un experimento puede suceder algo especial que no ocurre todos los días y que podría afectar los resultados.

Recomendaciones para incrementar la validez externa.

- 1.- Usar un muestreo representativo y al azar y luego asignar aleatoriamente los sujetos a los grupos ex per imentales, y/o,
- 2.- usar deliberadamente muestras heterogeneas que aseguren que todas las clases (soc., sexo, etc.) esten representadas en los grupos, y/o,
- 3.- usar uno o dos grupos (por ejemplo, escuelas) que se consideren típicos de la población.

Esta última técnica es muy útil en evaluación de proyectos donde se requiere poco poder de generalización. En este caso, el interés es explicar las clases de personas, situaciones o épocas a las que uno más desea generalizar y luego seleccionar al menos un ejemplo de cada clase que aparentemente sea similar a la clase más común.

Finalmente, cabe destacar lo siguiente:

- 1.- La validez externa y la de constructo se establecen o confirman a través de replicaciones del estudio.
- 2.- Al incrementar un tipo de validez se puede disminuir otro tipo. Por ejemplo, los experimentos al azar no son representativos de una situación normal.
- 3.- En investigación aplicada, es más importante la validez interna y la de constructo de los efectos.

REFERENCIAS

1.- Campbell, Donald T. y Stanley, Julian. Diseños Experimentales y Cuasi-Experimentales en la Investigación Social. Buenos Aires, Editorial A morrortu, 1970.

2.- Cook, Thomas D. y Campbell, Donald T. Quasi-Experimentation Design and Analysis Yssnes for Filled Settings. Chicago: Rand McNally College Publishing Co., 1979.

EDUCACION SUPERIOR

RESUMEN

La producción del ser de la familia
Educación superior es resultado de
se y con este trabajo, se la intentan
factor en instrumento de medición de la
habilidad de manejo y el conocimiento de
largo con el objetivo de poder clasificar
los estudiantes y adoptar las acciones
del caso.

Esta investigación y su presentación
fondo de 1980 en relación con
y la prueba de Añilub Académica con la que
tienen pocas semejanzas, aunque los
trabaja son más similares.

Hace mucho tiempo, que en los diversos
visado con el quehacer educativo, han venido
algunos afirmando sobre la mala preparación de
con los alumnos, en el manejo de su propia
más se multiplican en la educación superior.

En los últimos meses hemos leído o escuchado
los planteamientos muy concretos. En la gran
del, los conceptos colmatados tratan el tema
pasado: Emilia Añonny y Carlos Alberto Montaner.