

SESGOS EN LA EVALUACIÓN DE LA INVESTIGACIÓN PEDAGÓGICA Y PSICOLÓGICA: UNA SUCINTA REVISIÓN INTEGRATIVA

Antonio Fernández Cano
Universidad de Granada

1. INTRODUCCIÓN

Mucho se ha escrito y discutido sobre el diseño de investigaciones en Pedagogía y Psicología, pero escasamente sobre la evaluación de las mismas. El estudio del tópico, bien a nivel de proyecto o de estudio ya realizado, es de alta relevancia ya que puede cumplir múltiples funciones; entre otras:

— Estilar los mejores proyectos e investigaciones para, en consecuencia, aprender de ellos y recompensarlos procedentemente.

— Detectar y evitar plagios, fraudes, errores y violaciones ético-deontológicas.

— Aceptar los manuscritos de más alta calidad para su posterior publicación y/o inserción meta-analítica.

— Marca una correcta productividad de centros, departamentos e individuos.

— Señalar frentes «calientes» de investigación posibilitando el planteamiento de agendas.

Diversas modalidades se han venido utilizando para evaluar investigaciones. Las dos más usuales son la **revisión por pares**, o evaluación inmediata-formal o a priori, y el **impacto de citación**, evaluación mediata-formal o a posteriori. Aquí nos vamos a centrar en la evaluación de la investigación a partir de la revisión por pares.

2. CONCEPTUALIZACIÓN

La política de poder de los «colegios invisibles» y los propios paradigmas deben ser respetados si se quiere obtener el reconocimiento de la calidad de una investigación. Múltiples son los factores de adecuación; algunos de ellos, totalmente ajenos a una estructura de poder justa (entiéndase justa como pública) y/o ajenos a la naturaleza del propio paradigma. Estos factores diferenciales de sesgo se ponen de manifiesto, sobre todo, en el proceso de revisión de pares, no ya como una escasa fiabilidad (acuerdo/concordancia bajo entre valoradores) sino como un prejuicio, bastante más fiable, contra o a favor de factores de nula validez anexos a la investigación.

El uso de cualquier otro criterio distinto del de calidad metodológica introduce pues una amenaza potencial a la validez de la investigación. En el fondo, la existencia de sesgos en la evaluación de la investigación obedece a la tenaz influencia de la metateorías justificacionales (basadas en la autoridad) tan comunes en las ciencias sociales.

Existen abundantes estudios, comparativos e incluso experimentales, que testimonian la existencia de sesgos conscientes o inconscientes en el proceso de revisión por pares; aunque la mayoría de las indagaciones sobre sesgos se han realizado con metodología correlacional, lo cual ha suscitado serios problemas de interpretación.

3. HALLAZGOS GENERALES

3.1. Sesgo relativo a la afiliación institucional del autor

Se detecta un sesgo de favoritismo para autores de instituciones de alto prestigio incluso en revisión ciega (¡); o sea, los autores pertenecientes a instituciones con alta reputación reciben revisiones más favorables.

Incluso, también se detecta, corroborando una posición teórica propia del relativismo cognitivo, que la predisposición de los revisores está más sesgada a favor de autores que proceden de instituciones similares a las suyas, a un mismo grupo, área o asociación ya comparten un mayor conjunto de creencias sobre lo que es una buena investigación. El sesgo aparece como un **efecto halo**, en el que la afiliación institucional del autor está altamente correlacionada con la posibilidad de publicar (recibir evaluación positiva).

El controvertido estudio de Peters y Ceci (1982) fue aleccionador. Sustituyeron nombre e instituciones de 12 artículos ya publicados y volvieron a remitirlos a las mismas prominentes revistas en que aparecieron. Ocho de ellos fueron rechazados en base a «serias deficiencias metodológicas».

Estamos ante una de las manifestaciones del famoso «**Efecto S. Mateo**»: un curioso patrón de interacción selección x afiliación institucional.

La alternativa clásica de usar revisión ciega, aunque cuestionable, parece ser la única solución.

3.2. Sesgo relativo al revisor

Existe la creencia de un sesgo supercrítico por parte del revisor por aparecer «brillante pero cruel». Tal sesgo se manifiesta en hechos tales como: aceptar estudios en base a creencias o teorías asumidas por el revisor, divergentes valoraciones diferenciales (académicos vs. aplicados, investigadores vs. decisores, teóricos vs. prácticos, solicitantes vs. no solicitantes,...). En definitiva, los juicios sobre los que evalúa el revisor varían según la teoría que abrace (Lord et al. 1979) y según su rol profesional (Hall et al. 1988; Vockell y Asher, 1974; Ward et al. 1975).

Tal sesgo podría minimizarse aumentando el número de revisores que valoran un mismo estudio y/o asignando aleatoriamente los estudios a un mismo conjunto de revisores.

3.3. Sesgo pro estudios positivos

También denominado sesgo del prejuicio contra la hipótesis nula y que se manifiesta, véase Greenwald (1975), o Sterling (1959), en:

— Los autores son escasamente proclives a remitir sus trabajos para publicación si no obtienen resultados estadísticamente positivos).

— Las revistas publican mayoritariamente (alrededor del 90%) estudios con resultados positivos.

El obtener significación estadística (resultados positivos) es casi un prerrequisito para recibir calificadas evaluaciones y, en consecuencia, publicar y/o conseguir aceptabilidad. Se obvia que el avance científico se logra, a menudo, más poderosamente, mediante el rechazo de teorías: y la

estrategia capital para conseguir ésto es demostrar que las relaciones predictas (hipótesis) no se verifican; lo cual requiere, por tanto, aceptar la hipótesis nula.

El obtener significación estadística se ha convertido en un erróneo «sine qua nom» de respetabilidad y aceptación científica.

Tal sesgo se ha tratado de evitar utilizando dos aproximaciones:

- Valorando los informes sin la sección de resultados y con la sección de discusión modificada, y
- Valorando sólo el diseño y la importancia del estudio.

3.4. Ausencia de estudios de réplica

Los estudios de réplica, detectados por los revisores o declarados por los autores, son escasamente publicados, desvalorados; cuando la replicación en ciencias físicas es habitual y necesaria. Sin embargo, las réplicas indirectas (algunas variaciones, a veces, mínimas y/o nimias) o no confesadas son práctica abrumadora en ciencias sociales hasta el punto de haber dado paso a modalidades de investigación de síntesis. (Véase, al respecto, Darr, 1987).

3.5. Relativos al sexo del autor/a y/o revisor/a

Ponderar si la evaluación podría estar sesgada en función del sexo del autor y/o del revisor ha sido una constante, en la valoración de la producción investigadora, desde el estudio seminal de Goldberg (1968). El hallazgo de este estudio manifestaba que «hay una tendencia entre las mujeres a subvalorar el trabajo de profesionales de su mismo sexo».

El trabajo de Goldberg (1968) ha sido objeto de varios estudios de réplica con resultados contradictorios. La tendencia, con el tiempo, parece haber sido valorar equilibradamente a mujeres-expertas frente a hombres-expertos, manteniéndose el sesgo entre hombres y mujeres, no expertos, a favor de ellos.

En los últimos estudios (Lloyd, 1990), los resultados evidencian una desconfirmación total del postulado de Goldberg en el sentido de que las revisoras aceptaron más manuscritos de autoras que de autores y que los manuscritos de mujeres fueron más aceptados por revisoras que por revisores. Sin embargo, el propio Goldberg ya insinuaba que tal sesgo era más función del área de investigación. El sesgo sería más fuerte en áreas dominadas por varones que en áreas neutras o dominadas por mujeres.

Para controlar tal sesgo, se instituyeron políticas no sexistas, revisión ciega, se alentó la participación de mujeres a través de, por ejemplo, el aumento de editoras en consejos de edición de revistas.

3.6. Al sexo del autor en políticas editoriales

Las tasas de publicación, según sexo, eran tiempo atrás ostensiblemente divergentes y favorecían a varones. Pero tal sesgo, no sólo se ponía de manifiesto en cuanto autoría, sino progresivamente en cada paso hacia la pirámide de la eminencia: primeros autores, citas y editores.

Desde el estudio de arranque de Teghtsoonian (1974), el porcentaje de mujeres ha ido progresivamente en incremento aunque las divergencias aún subsistan (Gannon et al. 1992).

4. CONCLUSIONES

La existencia de sesgos en la evaluación de la investigación a través de la revisión por pares parece evidente. Las sospechas de conservadurismo, endogamia y pesimismo en el proceso científico han

sido una crítica constante. Armstrong (1982) llegó a ironizar, al respecto, emitiendo su sarcástica «fórmula de autor» al objeto de que cualquier investigador obtuviese una valoración inmediata y positiva de su trabajo; a saber:

- * No considerar un problema importante.
- * No retar las creencias existentes.
- * No obtener resultados sorprendentes.
- * No usar métodos sencillos.
- * No revelar todo lo que se sabe o aconteció sobre/en el estudio.
- * No escribir con claridad.

Mucho habrá que estudiar e investigar al respecto en nuestro país. Parece irónico que el método científico se haya aplicado tan escasamente a la evaluación del producto de la actividad científica, no ya para seleccionar lo mediocre de lo apto; sino para seleccionar entre lo apto, lo mejor.

El hecho de conocer la existencia de presuntos sesgos quizá pueda ayudarnos a una más válida y fiable evaluación de nuestra labor investigadora.

BIBLIOGRAFÍA

- AMSTRONG, J. S. (1982): Research on scientific journal: Implications for editors and authors. *Journal of Forecasting*, 1, 83-104.
- DARR, R. (1987): Another look at Meehl, Lakatos, and the scientific practices of psychologist. *American Psychologist*, 42, 145-151.
- GANNON, L., LUCHETTA, T., RHODES, K., PARDIE, L. y SEGRIST, D. (1992): Sex bias in psychological research. Progress or complacency? *American Psychologist*, 47 (3), 389-395.
- GOLDBERG, P. A. (1968): Are some women prejudiced against women? *Transaction*, 5 (5), 28-30.
- GREENWALD, A. (1975): Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, 81, 1-20.
- HALL, B. W., WARD, A. W. y COMER, C. B. (1988): Published educational research: An empirical study of its quality. *Journal of Educational Research*, 81 (3), 182-189.
- LLOYD, M. E. (1990): Gender factors in reviewer recommendations for manuscript publication. *Journal of Applied Behavior Analysis*, 23 (4), 539-543.
- LORD, C., ROSS, L. y LEPPER, M. (1979): Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37, 2.098-2.109.
- PETERS, D. P. y CECI, S. J. (1982): Peer-review practices of psychological journals: The fate of published articles, submitted again. *The Behavioral and Brain Sciences*, 5 (2), 187-195, 255.
- STERLING, T. D. (1959): Publication decisions and their possible effects on inferences drawn from tests of significance - or vice versa. *Journal of American Statistical Association*, 54, 30-34. También en *The significance test controversy* (1970), D. Morrison y R. Henkel (Eds.), (pp. 295-300). Chicago.: Aldine.
- TEGHTSOONIAN, M. (1974): Distribution by sex of authors and editors of psychological journals, 1970-72: Are there enough women editors? *American Psychologist*, 29, 262-269.
- VOCKELL, E. L. y ASHER, W. (1974): Perceptions of document quality and use by educational decision makers and researchers. *American Educational Research Journal*, 11 (3), 249-258.
- WARD, A. W.; HALL, B. H. y SCHRAMM, C. F. (1975): Evaluation of published educational research: A national survey. *American Educational Research Journal*, 12 (2), 109-128.