

Respuesta a las consideraciones realizadas por Carretero y colaboradores al trabajo titulado *Hacia una valoración integrada de las revistas españolas de Ciencias Sociales y Humanas. Las revistas de Psicología*

María Dolores Alcain y Adelaida Román

Centro de Información y Documentación Científica (CINDOC), Consejo Superior de Investigaciones Científicas (CSIC), Madrid

En esta nota se aportan aclaraciones a algunas de las críticas planteadas, especialmente aquellas referentes a aspectos metodológicos. Se incide en la explicación del contexto en el que la investigación se ha desarrollado y en su carácter no finalista. En particular, se profundiza en la explicación tanto de la fórmula empleada para el cálculo de los índices de valoración empleados, como en las ventajas que, para el propósito perseguido, tienen las opciones metodológicas hechas.

In answer to the considerations made by Carretero and collaborators to the titled work «Towards an integrated evaluation of the Spanish journals on Social and Human Sciences. The journals of Psychology». This note gives an answer to some of the critics that have been raised, especially those referring to methodological aspects. The explanation insists on the context in which the investigation has been developed and its non-finalist character. In particular, the explanation deepens on the calculation of the indices of evaluation and on the advantages that, for the intended purpose, have the methodological options that have been used.

En aras a la brevedad para no cansar a los lectores, restringiremos las respuestas/aclaraciones a aquellos elementos de la crítica del artículo que se publica en este número de la revista (páginas 669-675) titulado: «Evaluación de la calidad de las revistas científicas de Psicología publicadas en España: consideraciones al trabajo de Alcain y Román (2005)», firmado por Carretero-Dios y col., que creemos pueden tener interés, especialmente aquellos referentes a la metodología.

Con carácter general, se invita desde aquí a releer detenidamente hasta el final los documentos que se critican, con el objeto específico de no perder de vista el contexto y los objetivos de la investigación realizada y aún en curso.

Por si cupiera alguna duda, se reitera que la investigación en curso de la que el documento citado (Alcain, 2003) es solo una parte, está orientada a la creación de un modelo de valoración de las revistas científicas españolas de Humanidades y Ciencias Sociales desde un enfoque no finalista, sino investigador. Éste es el marco en el que se inscribe la valoración por parte del profesorado universitario e investigador del contenido de las revistas de las diferentes áreas del conocimiento, y la encuesta realizada como una herramienta para esa valoración. En este contexto, los objetivos de la encuesta fueron varios: a) obtener una valoración del contenido de las revistas que permitiera obtener un núcleo de re-

vistas de cada área entre las mejor valoradas por la comunidad científica concernida, con el objeto de tomarlas como revistas fuente para proceder posteriormente a un análisis de citas; b) poder contar con datos sobre la valoración que los especialistas de cada área hacen de las revistas para incorporarlos, junto con otros indicadores, en un futuro modelo integrado de valoración (ver tabla XIV trabajo Alcain y Román, 2005). En ningún caso, el objetivo perseguido ha sido obtener, a partir únicamente de los datos de valoración del profesorado, un ordenamiento de las revistas científicas publicadas en España según su calidad.

Esta línea de investigación lleva abierta en el CINDOC-CSIC (<http://www.cindoc.csic.es>) hace ya más de 20 años, desarrollada en paralelo a la creación de las bases de datos ISOC e ICYT (<http://bddoc.csic.es:8080/>) que desde 1975 recogen la producción científica publicada en revistas científicas españolas.

Las responsables de los trabajos que se critican no hacen sino continuar una trayectoria avalada por el rigor y la profesionalidad de numerosos colegas que, ayer como hoy, tienen como objetivo único contribuir con sus estudios a mejorar el conocimiento de la producción científica española, especialmente a través de las revistas científicas. Cuentan para ello de un observatorio privilegiado, el CINDOC, lejos de las luchas de intereses que tan frecuentemente se dan en las universidades. Ello da a su trabajo el valor añadido de poder desarrollarlo en un terreno neutral, alejado de toda confrontación.

Metodología del estudio y selección de la muestra

Tres son las principales críticas: a) falta información sobre el tamaño de la población y el tipo de muestreo realizado; b) consta-

tación de las diferencias existentes entre las diferentes áreas del conocimiento; c) no tratamiento de la no-respuesta.

En cuanto a la información sobre el tamaño de la muestra, se dice con claridad que se envió a todos los profesores en plantilla de las universidades públicas españolas, adscritos a las áreas del conocimiento relacionadas con la Psicología y Ciencias afines. El envío se hizo a partir de los datos facilitados por la DGU (Dirección General de Universidades del Ministerio de Educación y Ciencia) sobre el profesorado de todas las áreas del conocimiento reconocidas oficialmente. El hecho de relativizar la expresión utilizando los términos «la práctica totalidad» era una manera de no absolutizar el dato, ya que, en el tiempo que mediaba entre la entrega de la información por parte de la DGU y el envío de las encuestas y la recepción de las respuestas, siempre podría haber habido alguna variación del censo (jubilación, muerte, nuevo ingreso) y parecía más correcto, aunque fuera estadísticamente no significativo.

Efectivamente, las áreas del conocimiento son diferentes en tamaño y tienen diferente número de profesores adscritos, como se puede comprobar consultando la página de la DGU. Los porcentajes de respuesta a la encuesta también son diferentes. Esos son datos de la realidad, que, obviamente se reflejan en la encuesta. No se han visto ni motivos ni necesidad de introducir correcciones.

b) En cuanto al no tratamiento de las no-respuestas, interesa aquí poner de manifiesto que es habitual en estadística considerar que la muestra es homogénea salvo datos en contrario y que, por lo tanto, es una hipótesis razonable pensar que las no-respuestas seguirían la misma distribución que las respuestas.

Procedimiento para la valoración de las revistas

La crítica de más interés que se hace es que los «evaluadores», en este caso los profesores encuestados, son a la vez juez y parte, es decir, se considera que es una valoración «interna» y no «externa».

Siempre que se aborda una valoración de contenidos por parte de especialistas es una «valoración desde dentro de la especialidad». Tanto en el informe (Alcain, 2003) como en el artículo comentado (Alcain y Román, 2005) se expresan los riesgos de subjetivismo inherentes a este tipo de valoraciones, que, sin embargo, por el hecho de ser respondidas por colectivos importantes de profesores hacen que unos sesgos se corrijan o contrapesen por otros sesgos de signo diferente. Ésta es la razón por la que se afirma que los datos de valoración obtenidos por esta u otras encuestas similares *deben de ser utilizados conjuntamente con otros indicadores objetivos*. Sin embargo, a pesar de los posibles sesgos, más fuertes cuanto menor es la población que contesta, tienen la utilidad de permitir destacar en cada área las revistas que gozan de un reconocimiento claro.

Corrección de la encuesta y resultados del estudio

- Críticas al cálculo de las valoraciones en función del total de respuestas del área y no de las respuestas a cada revista dada.
- Críticas al procedimiento de cálculo del Iv A y del Iv A+B.

Ya se ha comentado antes que es proceder habitual en estadística considerar que, salvo sospecha o prueba en contrario, la muestra es homogénea. Es decir, la hipótesis es que las no contestacio-

nes se distribuirán siguiendo las mismas proporciones, las mismas tendencias que las expresadas por las respuestas.

En los comentarios hechos a nuestro trabajo en el punto 2, lo único que en realidad se afirma es eso, es decir, que nuestro universo muestral es homogéneo. Efectivamente, así es, salvo prueba en contrario. Por ejemplo: si se sospechara que la edad del profesorado puede introducir sesgos en las respuestas, deberíamos haber hecho un muestreo estratificado. Pero incluso así, seguiría siendo cierto que nosotros podríamos llevar el porcentaje de respuestas al 100% siempre que respetásemos la proporción de las edades en las distintas respuestas.

Evidentemente, éste es un ejemplo llevado al extremo y que no afecta al caso. Pero la única manera de comprobar que hay un sesgo introducido por la falta de respuestas (las no respuestas) es que los autores de las críticas hagan una afirmación al respecto y la sustenten con un estudio que pruebe esa hipótesis. La ciencia y las autoras criticadas se lo agradecerían. Les animamos desde aquí a hacerlo.

Con objeto de simplificar la respuesta a la crítica sobre la forma de cálculo de los índices de valoración, las notaciones empleadas se han simplificado al máximo.

El cálculo de los índices de valoración se sustenta en las siguientes definiciones:

$$Iv A = (A * V) / 100, \text{ donde}$$

A es igual al número de los que valoran como muy buena una revista determinada partido por el número total de respuestas del área. Corresponde a los valores de la columna «A» que se expresa en porcentajes.

Se denotará por MB (muy buenas) al numerador y RT (respuestas totales del área) al denominador, de modo que

$$A = MB / RT$$

V es igual al número de respuestas sobre una revista dada partido por el número total de respuestas del área. Esta «V» corresponde al encabezamiento de la columna «valoran».

Se denotará por RRD (respuestas para una revista dada) al numerador de V, mientras que el denominador ya ha sido denotado como RT, así pues,

$$V = (RRD) / RT$$

En consecuencia,

$$Iv A = (A * V) / 100 = (MB / RT) * (RRD / RT) * (1/100)$$

La fórmula Iv A es, en este caso, una definición que expresa un número (Iv A) en función de otros: «A» y «V».

Las definiciones de este tipo pueden ser inútiles, innecesarias y hasta «sinuosas», pero **nunca pueden ser erróneas** (no hay contradicción interna en la definición).

El número de respuestas recibidas siempre es mayor o igual que el número de respuestas que valoran una revista determinada. Por tanto, al dividir en nuestra definición de índice de valoración, por las respuestas totales del área en lugar de por las respuestas que valoran una revista determinada, siempre se obtiene un coeficiente de valoración menor que el obtenido por el procedimiento que ustedes estiman mejor.

Con ello se deducen dos cosas:

- 1) Que el Iv que se obtiene por cualquiera de los dos procedimientos es, en términos generales, del mismo tipo.
- 2) El índice que proponemos hace que RT (número de respuestas totales del área) influya haciendo disminuir el índice de valoración en función del número de profesores que responden a la encuesta pero no valoran una revista concreta.

Parece ilustrativo aportar un ejemplo que pone de manifiesto cómo el cálculo de los Iv, tanto por el procedimiento criticado como por el procedimiento propuesto por los autores de la crítica, no introducen variaciones de importancia en los «ranking» establecidos. Efectivamente, en el caso del Iv A+B propuesto por los auto-

res de la crítica, los valores son más altos, pero, sin embargo, no discriminan. El hecho de tener en cuenta el porcentaje de «valoran» referido al número que contestó la encuesta, hace que se penalice a las revistas que han sido valoradas por un número menor de profesores. El propósito principal para aplicar la fórmula ha sido obtener el núcleo de revistas mejor valoradas en cada área y utilizarlas como revistas fuente en el inicio del estudio de citas. En el futuro ese núcleo de revistas se irá renovando según las citas que obtengan las revistas y aumente su índice de impacto.

Las autoras de esta respuesta esperan con esto haber aclarado algunas de las dudas planteadas, y agradecen el interés que sus trabajos han despertado. Todo debate se resuelve siempre con más información y más luz sobre los temas debatidos, siempre que el objetivo que se persiga sea contribuir al conocimiento científico.

Metodología y CC de Comportamiento Contestadas= 68	Valoran	% valoran	Iv A	Iv A+B	Iv A propuesto	Iv A+B propuesto
Psicothema; Oviedo; 1989	65	95,59	53,42	80,13	58,46	87,69
Metodología de las Ciencias del Comportamiento; Murcia; 1999	63	92,65	51,77	73,57	60,32	85,71
Psicológica; València; 1980	51	75,00	28,68	46,32	50,98	82,35
Metodología de Encuestas. Revista de la Sociedad Internacional de Profesionales de la Investigación en Encuestas; Sevilla; 1999	44	64,71	9,52	33,30	22,73	79,55
REMA. Revista Electrónica de Metodología Aplicada; Oviedo; 1996	39	57,35	5,06	23,62	15,38	71,79
Revista de Psicología General y Aplicada; Madrid; 1946	44	64,71	2,85	19,98	6,82	47,73
Psiquis; Madrid; 1979	10	14,71	0,00	0,43	0,00	20,00
Revista de Psicología Social Aplicada; Alacant; 1991	10	14,71	0,00	0,43	0,00	20,00
Revista de Psicoanálisis; Madrid; 1995	4	5,88	0,00	0,00	0,00	0,00
Revista de Psicodidáctica; Vitoria-Gasteiz; 1996	1	1,47	0,00	0,00	0,00	0,00

Referencias

- Alcain, M. D. (2003), Índices de impacto de las revistas españolas de ciencias sociales (biblioteconomía y documentación, psicología, psiquiatría y urbanística y ordenación del territorio) a partir del análisis de las revistas mejor valoradas por los pares (Ref. EA2003-0020). <http://www.cindoc.csic.es/info/inforev2.html>.
- Alcain, M. D. y Román, A. (2005). Hacia una valoración integrada de las revistas españolas de Ciencias Sociales y Humanas: las revistas de Psicología. *Psicothema*, 17(2), pp. 179-189.
- Carretero et al. (2005). Evaluación de la calidad de las revistas científicas de Psicología publicadas en España: consideraciones al trabajo de Alcain y Román (2005), *Psicothema*, 17(4), pp. 669-675.