

INDICADORES DE PRODUCTIVIDAD CIENTÍFICA

Implicaciones para la evaluación de la psicología española

D. Páez y J.F. Salgado

Darío Páez Rovira es Catedrático de Psicología Social en la Universidad del País Vasco. Jesús Fernando Salgado Velo es Catedrático de Psicología Social en la Universidad de Santiago de Compostela.

La evaluación de la actividad científica y de la productividad de los investigadores es una cuestión de interés desde muchas perspectivas distintas:

la financiación de la investigación, la recompensa de la actividad de investigación, la formulación de políticas de investigación y la toma de decisiones relacionadas con tales políticas, la planificación estratégica de la actividad universitaria, la negociación salarial en los casos en que ésta se hace directamente entre el investigador y la organización contratante, la dotación de plazas de profesorado universitario y su promoción así como la concesión de becas, entre otras. Por estas razones, la preocupación por sistemas "justos" de evaluación de la actividad investigadora ha sido y es un preocupación tanto para los propios implicados en la investigación como para aquellos que tienen que tomar decisiones al respecto. En España, tres entidades del Ministerio de Ciencia e Innovación (antes del Ministerio de Educación y Ciencia) son las encargadas, con diferentes propósitos, de evaluar la actividad científica. La calidad de los proyectos investigación es evaluada por la Agencia Nacional de Evaluación y Prospectiva (ANEP) (ver Gordillo, González Marques y Muñiz, 2004). La actividad investigadora del profesorado universitario funcionario la evalúa la Comisión Nacional Evaluadora de la Actividad Investigadora (CNEAI) y los criterios de evaluación se describen en la convocatoria anual en el Boletín Oficial del Estado (BOE). Por su parte la Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA) realiza, entre otros, la evaluación de la calidad de los programas de doctorado así como la evaluación del profesorado cara a la contratación

o a la acreditación para el acceso a los cuerpos docentes universitarios (ver ANECA, 2006).

Así pues, de modo general, dos aspectos de la actividad investigadora son tenidos en cuenta por estas agencias y comisiones: la cantidad y la calidad de lo producido. No siempre los indicadores de estos dos aspectos han confluído, dando lugar a numerosos debates al respecto y posiciones encontradas. Los criterios fundamentales utilizados en España son, además de la cantidad de artículos publicados (criterio de producción) en revistas indexadas en el *Institute for Scientific Information* (ISI), el número de citas que el trabajo de investigación ha recibido y el índice de impacto de la revista en la que se ha publicado en trabajo de investigación (criterio de calidad). Un reciente Informe de la ANEP especificaba así los criterios de evaluación de la I+D en Ciencias Sociales (Economía, Geografía, Sociología, Ciencia Política, Educación y Psicología): “Una publicación de investigación aporta un resultado novedoso en un campo del saber. Por ello, una publicación de investigación tiene que estar al alcance de la comunidad científica, tiene que haber sido difundida por un medio que incluya la evaluación por pares y los resultados de la investigación deben de ser susceptibles de réplica y verificación. Además, para ser considerada relevante (en el sentido de hacer avanzar el conocimiento), la publicación ha de tener impacto. [...] Para los trabajos difundidos en publicaciones periódicas, el listado de referencia sería el elaborado por el ISI para cada una de las áreas. [...] Los criterios de valoración han de incluir, fundamentalmente, el índice de impacto de la revista y las citas que el artículo haya recibido, todo ello normalizado, teniendo en cuenta la productividad media, a nivel mundial, de las diferentes áreas y especialidades científicas. En este sentido, se considera urgente que, por parte de la administración, se realicen estudios de productividad para las diferentes áreas, de forma que se disponga para la evaluación de una norma basada en la evidencia” (Cabreres, Carmona, Fuentes, Gómez, Gutiérrez, Herrero, Jordana, Olcina, Páez y Sánchez, 2007, páginas 6-8). Por consiguiente, este artículo busca presentar las posibilidades y limitaciones de los indicadores bibliométricos a los que el texto anterior hace mención, así como de otros indicadores (e.g. los tramos o sexenios) para evaluar el rendimiento o productividad científica. La evaluación de la actividad investigadora incluye tanto los “insumos” para la producción (fondos de investigación, subvenciones, becas de formación, etc.) como los resultados de esta producción de conocimiento, esencialmente las publicaciones científicas. La relación costes/resultados es un aspecto interesante y poco desarrollado de la evaluación. Nos concentraremos en este artículo en las publicaciones científicas como indicador de rendimiento, ya que hay información y reflexión suficiente sobre ello. Nos parece importante que las opiniones sobre estos temas se fundamenten en evidencias y no en juicios a priori. Los datos que presentamos, resumiendo estudios parciales, permitirán

situar la evaluación en un ámbito realista, superando una serie de prejuicios existentes (e.g. la situación española es particular, es escandaloso que la producción sea asimétrica, los estudios bibliométricos subestiman la producción española). Además nos permitirán sugerir criterios normativos, aunque éstos se basen en estudios parciales y limitados, ya que en la actualidad no hay estudios administrativos oficiales. Empezaremos analizando las características de la producción científica y su relación con la docencia, para cuestionar una serie de creencias aparentemente dominantes en nuestro entorno. A continuación describiremos las características de la producción científica en la Psicología española. Posteriormente analizaremos los tramos o sexenios, un criterio de evaluación que nos parece claramente inadecuado aunque se haya planteado frecuentemente como tal. Seguidamente, expondremos las potencialidades y limitaciones de los indicadores basados en la producción de artículos y citas. Finalmente, en un último apartado describiremos criterios normativos basados en los estudios reseñados.

Las características reales de la producción científica. Posición y tendencias de la psicología española

España ha pasado del rango 32 en 1963 al 9 en producción científica mundial entre los años 60 y el fin del siglo XX. Desde la aprobación de la Ley de la Ciencia en 1986 al siglo XXI el número de investigadores ha pasado de 24,500 a 101,000 – 37% en la privada. En el 2005, en la Enseñanza Superior (E.S.) había 54.028 investigadores, un 4,7% más que en el 2004. Sin embargo, la tasa de crecimiento del 2005 respecto al 2004 fue ligeramente inferior (4,9% en 2004) y tres puntos menos que en 2003 (Fundación Conocimiento y Desarrollo, 2007, de ahora en adelante Fundación CyD).

Actualmente, la economía española es la novena-décima en importancia sobre el PNB mundial. La producción científica se corresponde con este peso en lo referente a la cantidad de publicaciones, ya que en la actualidad es el noveno país productor de ciencia – pasando del 0,7 en 1981, al 1,1 en 1986-88, al 2,0% en 1995-1997 y al 2,85% en 1997-2001 (King, 2004). Un estudio utilizando la base Scopus encontró que España se situaba en el rango décimo de la producción mundial de Psicología, detrás de EEUU, Reino Unido, Canadá, Alemania, Australia, Holanda, Francia, Italia y Japón. Además, la posición en el ranking en Psicología era ligeramente inferior a la posición de España en el ranking de producción científica total (De Moya, 2008). En el año 2002 la participación de España en la producción científica total era del 2,95 mientras que la participación de Psicología y Psiquiatría era de 1,76, es decir, la producción tenía un impacto inferior al medio (Aguirre, 2003). Entre los años 1980 y 2000, se duplicó la inversión en I+D y en educación, redundando en una duplicación del número de profesores universitarios (de

los que aproximadamente un tercio hacen investigación) y una cuadruplicación del número de investigadores. En consecuencia, entre 1986-88 y 95-97 se ha duplicado la producción de artículos científicos (King, 2004). La producción científica de las Universidades comprende el 60,4% de la producción científica española en revistas internacionales, utilizando solo el 32% del presupuesto (Fundación CyD, 2005; 2006). Por el contrario, la producción universitaria solo representa el 10% de las patentes, indicador de innovación tecnológica (Fundación CyD, 2007). La producción de artículos científicos ha aumentado regularmente y se ha orientado progresivamente a revistas internacionales. Las universidades con mayor producción total fueron las mismas que las más productivas en términos relativos (Fundación CyD, 2005; 2006). En síntesis, si los investigadores se han cuadruplicado, la producción se ha duplicado. Se está llegando a una cierta estabilización, como indican los análisis de la producción global y los de producción sectorial en Psicología (Musi-Lechuga et al, 2005). Es decir, las mismas posiciones y tendencias al crecimiento de la producción científica se dan en Psicología, si bien de forma menos marcada.

Las características reales de la producción científica y su relación con la docencia

En nuestro contexto han sido muy frecuentes afirmaciones como “publicar es una pérdida de tiempo porque nuestra producción se cita poco” o “son muy pocos los que publican” o “potenciar la investigación conlleva un efecto negativo en la docencia”. Sin embargo, una revisión de los estudios empíricos existentes no apoyan estas creencias. En primer lugar, se observa que el bajo impacto de la producción es general. Igualmente ocurre con la distribución asimétrica de la producción e impacto, que es asimétrica en EEUU y otros países (véase el artículo de Joy, 2006, reproducido parcialmente en este monográfico). Por otro lado, según los estudios internacionales, no hay relación negativa entre investigación y docencia. La limitada evidencia disponible de nuestro entorno va en el mismo sentido.

Con respecto a la escasa citación, según algunas estimaciones generales, el 90% de los artículos no son nunca citados el primer año. Incluso en las revistas de mayor impacto, la mayoría de las citaciones vienen de un 15% de los artículos publicados (Lee, Cassano y Vicente, 2005). A finales del siglo XX el 55% de los artículos de las revistas ISI no habían sido citados en los primeros cinco años, y que la moda citación era cero (Hamilton, 1991). A este respecto, el ISI en un informe de 1999 reflejaba que el porcentaje de artículos no citados a medio plazo oscilaba entre el 44,5% de Ciencias de la Computación y el 13% de Inmunología, siendo del 25% el porcentaje de artículos no citados en Psicología y Psiquiatría (ISI, 1999).

Finalmente, también se ha constatado que el 10% de los científicos más destacados producen alrededor de 50% de las publicaciones científicas. Un estudio ya antiguo encontró que el 10% de los autores en Psicología publicaban el 47% de los textos, mientras que el 50% de los autores publicaban solo el 15% de éstos (Deniss, 1954 citado en Lee, Cassano, y Vicente, 2005). El artículo de Joy (2006) confirma que 120 profesores concentran el 47% de las publicaciones de una muestra de alrededor de 1200. En el caso español, un 5% o menos de los autores más prolíficos producen el 30% de los artículos publicados por titulares y catedráticos de psicología (Musi-Lechuga et al, 2005). Esta concentración de producción en una minoría también se da en el impacto reflejado en el número de citas. Una minoría de científicos recibe una gran parte de las citas en todos los campos científicos, incluida la Psicología. Por ejemplo, un 10% de todos los artículos de Astrofísica y un 5,2% de los de Ciencias de la Computación obtienen el 50% de todas las citas. En Psicología y Psiquiatría ese porcentaje es del 7,1% (ISI, 1999).

También se ha criticado la alta tasa de errores de identificación de autores en los estudios bibliométricos locales realizados, atribuyéndose la misma, generalmente, a distorsiones voluntarias. El hecho es que en los estudios bibliométricos se da un error de identificación de citas, con una media del 7% pudiendo llegar al 30% (Avital y Collopy, 2001; van Raan, 2005).

En cuanto a la importancia de la investigación sobre la docencia y su efecto negativo sobre esta última, cabe señalar que en la universidad española la investigación es menos prioritaria de lo que se percibe. En el 2004 el 44% del profesorado universitario no había solicitado un sexenio o no se lo habían concedido y en el área de Ciencias Sociales el porcentaje era del 41% (Fundación CyD, 2005; 2006). Es decir, solo seis profesores funcionarios sobre diez intentan ver reconocida su actividad de investigación. Respecto al hecho de que la investigación menoscaba la docencia, los tres meta-análisis que han examinado la relación entre productividad (generalmente operacionalizada como la cantidad de artículos publicados) y la eficacia docente (operacionalizada mediante las encuestas de opinión del alumnado sobre la calidad docente del profesor) han encontrado correlaciones positivas pequeñas (alrededor de $r=0,10$). Una revisión de 30 de los estudios incluidos en estos meta-análisis mostró que en un 66% se daba una asociación de cero y en un 37% una asociación positiva (Feldman, 1987; Allen, 1996; Hattie y Marsh, 1996).

Como limitaciones señalar que estos estudios se han llevado a cabo en universidades sobre todo anglosajonas y la operacionalización tanto de la productividad como de la eficacia docente ha sido limitada. Pese a ellas, a partir de estos resultados se ha concluido que no hay una relación negativa, sino que la relación entre investigación y docencia es ligeramente positiva. Dicho de otra manera, la investigación no afecta nega-

tivamente a la docencia, y es probable que haya buenos investigadores que sean buenos docentes, aunque ser buen investigador no implica ser buen docente. Singularmente en las universidades anglosajonas la creencia dominante es que la investigación y la docencia están positivamente asociadas, mientras que en España se piensa lo contrario. Una estimación de los docentes que eran buenos en docencia e investigación (por encima de la media en evaluación de alumnos y en producción de artículos) evidenció que un 22% eran buenos o superiores a la media en ambos aspectos (Fairweather, 2002). Aunque no hay estudios locales, un análisis de los datos agregados parciales del área de Psicología en las evaluaciones de una Agencia de Evaluación Autonómica, encontró una correlación positiva y significativa entre docencia e investigación [$r(40)=.64, p<.01$]. Las altas puntuaciones en investigación, esencialmente publicaciones y becas competitivas, se asociaron a mejores puntuaciones en docencia. Éstas incluían además de evaluación de alumnos, indicadores como número de tesis, cursos de doctorado, cursos de verano, y elaboración de material docente en la forma de libros, lo que puede explicar en parte que estas actividades se asocien con investigación. Los análisis de esta pequeña muestra evidenciaron que la investigación no obstaculiza la participación en la gestión. Las puntuaciones de gestión e investigación no correlacionaban significativamente [$r(40)=.13$] (Páez, 2007).

Concluyendo, la situación española es similar a la general: unos pocos autores producen la mayor parte de los textos, unos pocos textos reciben la mayor parte de las citas y muchos son poco o nada citados. Por lo tanto, los autores productivos reciben más citaciones. Además, en la estimación de producción de artículos y citas se cometen errores importantes y la relación entre investigación no es de oposición con la docencia. Más bien es de independencia o, en todo caso, de una ligera relación positiva.

Los Tramos o Sexenios de Investigación.

En el Informe de CyD 2005 se indicaba que el 34% de los profesores no habían solicitado sexenios de investigación, complementos económicos otorgados según las publicaciones indexadas en ISI Web of Knowledge (WoK) de Thomson Reuters. De los sexenios demandados, se otorgaron un 69%, variando entre el 56% en economía y empresariales y el 87% en biología celular y molecular. La proporción de personas con sexenio sobre el total de profesores era del 47,85. La media general de catedráticos con sexenios era del 73% y la de Titulares de Universidad (TU) era del 58% (Fundación C y D, 2005). En el 2004 el promedio de sexenios por profesor era de 1,2 (Fundación CyD, 2007).

Se ha argumentado que los tramos de investigación que una persona tiene son un indicador de calidad, otorgándole en la evaluación de la

productividad un peso casi igual al de artículos publicados en revistas del *Journal Citation Reports* (JCR) de WoK (Buela-Casal, Bermúdez, Sierra, Quevedo-Blasco y Castro, 2009). Sin embargo, un estudio sobre los catedráticos de Psicología Social encontró que el número de sexenios de investigación correlacionaba más fuertemente con la antigüedad como catedrático y con la antigüedad como doctor que con los índices de productividad científica como el índice h , el total de citas o la cita máxima (Salgado y Páez, 2007). En el mismo sentido Carcía-Pérez (2009) encontró que en el área de Metodología no había diferencia entre catedráticos y titulares en el índice h . Según Musi-Lechiga et al (2005), la producción media de artículos ISI de catedráticos y titulares (con mayor producción sobre el total) era similar en la mayoría de áreas de Psicología. Recordemos que los primeros tienen más sexenios que los segundos por definición. A partir de esta evidencia limitada podemos concluir que número de sexenios muestra ser un indicador de valor reducido. De hecho, los resultados sugieren que la obtención de sexenios estaría más basada en la posición desempeñada y en la antigüedad que en la productividad científica. Por ende pensamos que, en la actualidad, los sexenios no son un indicador útil de evaluación de la calidad de la productividad científica.

Las bases de datos de citaciones

Las revistas incluidas en la base de datos del WoK, *Web of Science* (WoS), tienen la ventaja de que basan la selección de artículos en el sistema de revisión de pares y, en general, son las revistas de mayor impacto mundial en la ciencia. Ahora bien, la muestra de revistas de WoS tiene una serie de limitaciones: se da un sesgo lingüístico a favor del inglés, predominan las áreas de ciencias naturales y de la vida, y las áreas de ciencias sociales y de humanidades están sub-representadas. Una comparación entre la lista de 1986 de las revistas científicas de la UNESCO con las revistas integradas en WoS mostró que había 2,5 más revistas en la primera que en la segunda. El 60% de las revistas del WoS eran de EEUU frente al 17% de la lista de la UNESCO (Schoepflin, 1990). En la actualidad la WoS incluye alrededor de 8700 revistas. Cubre solo alrededor del 10% del total de revistas científicas, que se estimaban en alrededor de 100 mil a fines del siglo XX. Ahora bien, cubre un cuarto de las cerca de 24 mil revistas que utilizan revisión por pares (Beauvois y Pansu, 2008). Por otro lado, en 2005 el 98,7 % de las revistas de WoS eran de habla inglesa, representando alrededor del 0,20 del total las revistas en otros idiomas (francés o chino). Recientemente se ha ampliado el número de revistas que publican en otros idiomas o que provienen de países con un idioma diferente al inglés. En cuanto a las áreas de ciencias humanas y sociales, indicar que solo son de esas áreas un 20% de las revistas de la WoS.

Sin embargo, pese a todas estas limitaciones, y confirmando que el inglés es el latín científico de nuestro tiempo, hay que señalar que de nueve de cada diez citas realizadas en las publicaciones de Psicología en español son respecto a revistas en inglés o internacionales. Además los profesores e investigadores que publican mayoritariamente en publicaciones españolas citan muy poco las revistas en castellano (Alcain y Román, 2005).

Otra cuestión criticada de estas bases es que no integran libros y monografías. Un estudio concluyó que en el caso de las ciencias sociales el 40% de las publicaciones se hacía en libros y monografías, frente al 15% de las ciencias naturales, diferencias que se encuentran incluso en los países anglosajones (Pestaña, Gómez, Fernández, Zulueta, Méndez, Koenig y Bookstein, 1995). Ahora bien, hay que indicar que un procedimiento del sistema de búsqueda de WoS, *Cited Reference Search*, permite estimar las citas realizadas en las revistas de WoK a todo tipo de texto (incluyendo tesis, libros, monografías y actas).

Por otro lado, la base Scopus, lanzada por Elsevier en el 2004 como una alternativa europea a la WoS, incluye una base más amplia de revistas (cerca de 13 mil) así como 125 series de libros, 700 actas de congresos científicos y 275 millones de páginas Web. En esta base un 60% de los textos provienen de fuera de EEUU (Jacso, 2008). Es decir, hay alternativas en ambas bases que permiten integrar la producción de libros y monografías.

Sin embargo, tanto WoS como Scopus tienen como limitaciones esenciales el ser de pago y tener una cobertura parcial. Si bien es verdad que, actualmente, WoK está disponible para todos los investigadores públicos españoles. Como base gratuita alternativa está Google Scholar y el sistema de búsqueda Publish or Perish (ver www.harzig.com). Las limitaciones centrales de los sistemas gratuitos son la base indiferenciada de citas y las dificultades para identificar claramente a los autores, revistas y artículos.

Estas tres bases y sistemas entregan indicadores de citas, de índice *h* y de índice *g*. Los estudios limitados existentes sugieren que la WoS proporciona más fácilmente datos fiables, que las bases se recubren parcialmente (alrededor del 40%) y que en general se obtienen mayores citas con Scopus que con WoS y aún más con Google Scholar (Kermarrec, Faou, Merlet, Robert y Segoufin, 2007). Conjugando estos sistemas, si los autores hacen públicos sus CV's para identificar claramente todos sus textos, se pueden obtener fácilmente estimadores comparables.

La calidad de las publicaciones. El factor de impacto de una revista (*Journal Impact Factor* o JIF)

La calidad de la revista en la que se publica se considera un indicador de evaluación. Por ejemplo, se requiere un mínimo de dos publicaciones indexadas en WoK para obtener un tramo o sexenio de investigación en Psicología. En evaluaciones como las realizadas recientemente por agencias de evaluación autonómicas estilo ANECA, reflejando la importancia de la calidad de la revista, se otorgaba un peso de 100 a publicaciones indexadas internacionales, de 50 a publicaciones indexadas nacionales (Psicothema y otras más en la actualidad) y de 25 a publicaciones no indexadas (Páez, 2007). El Factor de Impacto de la Revista o *Journal of Impact Factor* (JIF) es la relación entre las citaciones recibidas durante un periodo de dos años y el número de artículos publicados durante el mismo periodo. Por ejemplo, el JIF para Science en 1995 era de 21,9, ya que ese año se habían dado 24979 citas de artículos publicados en 1993, más 20684 citas de artículos publicados en 1994 y los artículos publicados esos dos años ascendieron a 2084. A partir del valor del JIF se puede concluir que las revistas con un impacto como Science son escasas. En concreto en 1994, un 6.7% tenía un JIF > 3, un 27% tenía $3 < \text{JIF} < 1$ y el resto tenía un JIF < 1. Es más, el 25% tenía un JIF inferior a 0,29.

Por otro lado, los datos disponibles reflejan que el JIF varía mucho según las áreas. Los factores de impacto de bio-medicina y de ciencias de la vida son mucho mayores que las de ciencias humanas y sociales. Esto se debe a que hay grandes diferencias en el rango de citas e impacto entre las disciplinas científicas. Según el ISI (1999), el total de citas necesarias para estar en el 1% de los artículos más citados oscilaba entre 285 citas para en Biología Molecular y Genética y 32 en Educación. En Psicología y Psiquiatría, 100 era la media de citas necesarias para estar en el 1% de artículos más citados. En 1999, en el área de Psicología, 53 revistas tenían un JIF entre 0,175 y 7,79. En 2006, 60 revistas tienen un impacto entre 0,12 y 9,78 (Beauvois y Pansu, 2008).

También se ha cuestionado la idoneidad de los dos años como periodo de tiempo a considerar respecto a las áreas de ciencias sociales y humanas. Las citaciones tienden a crecer hasta los 3-5 años en ciencias naturales y hasta los 5-7 años en ciencias sociales. En el Reino Unido, para el periodo 1981-1998, la media de citaciones de artículos en Psicología era de 8,7 frente a un 5,7 en Economía. Usando un intervalo de cinco años, la media de citaciones bajaba respectivamente a 2,3 y 1,5 (Katz, 1999). En la actualidad se dispone de un JIF para cinco años. En concreto, tres-cinco años se han planteado como ventana razonable para áreas como las ciencias sociales y humanas, ya que aún en ellas la posibilidad que un texto nada o poco citado en este periodo sea después muy citado es muy baja (Glanzel, 2008).

Un análisis de las 61 revistas de Psicología según el JCR, respecto al año 2009 y utilizando la ventana de cinco años, evidencia que un 38% tenía un JIF < 2, un 23% tenía un JIF entre 2-3, un 15% tenía un JIF entre 3-4 y un 13% tenía un JIF entre 4-5. Únicamente tres revistas tenían impactos superiores a 10: *Annual Review of Psychology* (17,26), *Psychological Bulletin* (15,23) y *Psychological Review* (10,60); las tres son publicaciones que incluyen revisiones de amplio espectro. Por otro lado, respecto a áreas concretas como por ejemplo Psicología Social y Psicología Clínica los resultados también son extremos. En Psicología Social, de las 47 revistas indexadas, tres tenían el JIF más alto: *Advances in Experimental Social Psychology*, publicación de artículos de revisión, (6,7), *Personality and Social Psychology Review* (4,76) y *Journal of Personality and Social Psychology* (4,5). El resto se distribuía del siguiente modo: 5 revistas tenían un impacto entre 2-3, 18 revistas entre 1-2. En general, un 45% tenían impacto inferior a 1 y dos tercios tenían un impacto inferior a tres. La revista en la que más publican los psicólogos sociales españoles (después de *Psicothema*), *European Journal of Social Psychology*, tenía un impacto de 1,57 situándose en el rango 13. En cuanto a las 87 revistas de Psicología Clínica, los resultados evidencian que, salvo *Journal of Clinical Psychiatry* (JIF=5,06), ninguna revista superaba un impacto de 5. Solo seis revistas tenían un impacto entre 4-5, teniendo el prestigioso *Journal of Abnormal Psychology* un impacto de 4,68. Catorce revistas tenían un impacto entre 2-3 y cuarenta tenían un impacto entre 1-2. Un 29% tenían impacto inferior a 1 y dos tercios inferior a dos. La revista internacional en la que más publicaban los investigadores de Personalidad, Evaluación y Tratamientos Psicológicos (*Petra*), *Personality and Individual Differences*, tiene un impacto de 1,40 y en la revista en la que más publicaban los de Básica y los de Metodología, *Perceptual and Motor Skills*, tiene un JIF de 0,27. De esta visión general se puede concluir que plantear criterios de evaluación en Psicología tales como tomar en consideración solo las revistas con impacto superior a 2, como se hace en Medicina o en algunos países, es inaplicable o improcedente (Thibergien y Beauvois, 2008). También es razonable esperar que los investigadores publiquen en revistas de impacto entre menos de 1-2 y que aquellos investigadores que publiquen en revistas de impacto superiores a 2 se sitúen por encima de la media.

Limitaciones del JIF

En primer lugar se ha argumentado que este indicador es sensible a la realización de auto-citaciones. Las revistas pueden "inflar" su impacto pidiendo a los autores de nuevos artículos que citen a textos antiguos de ella, al margen de su relación con el contenido. Por ejemplo, las buenas revistas, como el *Journal of Personality and Social Psychology* o el

Journal of Health and Social Behavior tienen tasas de auto-citaciones del 20-35% (Johnson y Wolinsky, 1990).

También la publicación por entregas, o en salami, puede inflar el impacto ya que en cada una se incluyen solo aspectos parciales de un estudio y la apreciación global exige citar todas las entregas. Esta táctica se puede apreciar sobre todo en el caso de grupos y autores consolidados, que entregan poco a poco una serie de estudios complementarios.

Por otro lado aunque la calidad de la revistas donde se publica es un indicador del esfuerzo académico, no lo es tanto de su impacto. La relación entre el JIF de la revista y las citaciones de los artículos publicados en ellas no es alta. Un análisis exhaustivo de las referencias del WoS encontró una tasa de referencias sin relación con el tema del artículo del 7%. Los sistemas de búsqueda proponen una tasa de Auto-citación o relación entre citaciones publicadas en la revista y el total de citaciones que esta obtiene. Sin embargo, se ha concluido que la tasa de auto-citación de las revistas no es una limitación central, ya que puede indicar la consistencia de las líneas de investigación (Johnson y Wolinsky, 1990; Glänzel, 2008). En síntesis, publicar en revistas de impacto inferior a uno indica un esfuerzo meritorio y en publicaciones de impacto superior a 1-2 un esfuerzo mayor. Sin embargo, como indicador tiene un valor limitado, ya que el impacto de la revista no es automáticamente el impacto de la producción del artículo de la persona. Es por esto que se han desarrollado indicadores como los siguientes.

Diferentes tipos de Indicadores bibliométricos sobre la cantidad y la calidad de la producción científica

Los indicadores de citaciones de la producción científica más comunes son (Salgado y Páez, 2007):

a) Número total de trabajos o artículos publicados: cuya ventaja reside en que mide la producción realizada pero cuya desventaja es que no mide la importancia ni el impacto de los trabajos.

b) Número total de citaciones: cuya ventaja es que mide el impacto total y cuya desventaja es que es difícil de lograr. Este indicador puede estar inflado por unos pocos "grandes impactos" que pueden no ser representativos de la persona, si es coautor de tales trabajos con muchos otros. También pueden dar un peso exagerado a los artículos de revisión o de meta-análisis, altamente citados, frente a contribuciones originales de investigación. Se ha estimado que un artículo de revisión tiene 5 veces más impacto que un artículo normal en una misma revista (Beauvois y Pansu, 2008). Ahora bien la influencia de artículos de síntesis y meta-análisis en parte las citaciones reflejan la utilidad científica del texto, es decir, estos textos tienen un impacto no solo por un artefacto

estadístico, sino porque en general reflejan el estado de la cuestión sobre un tema y atraen reconocimiento científico por ello.

Estos dos indicadores de cantidad y calidad se han criticado fuertemente ya que la cantidad no se asocia a la calidad. Sin embargo, los estudios muestran que la investigación de alta calidad proviene de las personas que publican más prolíficamente (Simonton, 2007). Confirmando esto, en varios estudios con muestras de psicólogos, la relación entre el total de artículos, el total de citaciones e índices más sofisticados como el índice *h*, que evalúan la calidad del impacto, es lineal, significativa y fuerte (superior a .70) (García-Pérez, 2009; Dorta-Contreras, Arencibi-Jorge, Marti-Lahera y Araujo-Ruiz, 2008). También se ha cuestionado su carácter sustantivo. Se ha planteado que considerar las citaciones como el sistema de recompensas de la ciencia, como planteaba el sociólogo de la ciencia Merton, es inadecuado e ingenuo (Beauvois y Pansu, 2008). Las citaciones, como las publicaciones, serían indicadores de visibilidad científica, más relacionados con redes de sociabilidad, colegios invisibles y capacidades de comunicación intercultural, que reales indicadores de productividad (Iñiguez, Martínez, Muñoz, Peñaranda, Sahún y Alvarado, 2008). Sin negar que estas afirmaciones tienen un núcleo de verdad, los estudios han mostrado que indicadores basados en citas tenían capacidad predictora de la obtención de reconocimiento científico de todo tipo, incluyendo los Nobeles, es decir, que tenían una validez como indicador predictivo de rendimiento científico (Simonton, 2003). Aunque en la actualidad no se considera ingenuamente a las citas como un reflejo "puro" de reconocimiento, hay consenso en que las citas son un indicador de recepción, uso o consumo de la información inherente a un artículo científico. Si un texto recibe varias citaciones al año, varios años después de su publicación, es muy probable que su contenido se integre en la disciplina científica en cuestión. Si un texto nunca es citado varios años después de su publicación, es muy probable que su contenido no se integre al cuerpo del conocimiento científico. Las citas constituyen un indicador de la recepción del texto, de su impacto público en el conocimiento, que es la forma esencial de evaluar la actividad científica (Cabrales et al, 2007). Como contra-ejemplo por excelencia de muchos críticos de los criterios cuantitativos se han mencionado casos como los de Mead o Wittgenstein, que publicaron poco y son reconocidos como grandes pensadores. Ante esto, podemos mencionar que lo que confirma que Wittgenstein es un gran filósofo o que Mead sea un precursor de la Psicología social, no es que haya publicado poco, sino que lo poco publicado y las notas divulgadas por sus discípulos, situados además en instituciones anglosajonas de gran prestigio, han sido ampliamente citados y discutidos. Es decir, que sus escritos han tenido un impacto público, evaluable por citaciones.

Con respecto a la distribución de la producción de artículos en nuestro contexto, los datos agregados de los años 2003-2004 indican que la

Media de artículos del ISI por profesor español universitario al año es de 0,5 por año (Fundación C y D, 2006). Una evaluación de la productividad durante el quinquenio 2000-2004 encontró que el promedio de las universidades más productivas españolas, como la Universidad de Barcelona (puntuación 0,92 sobre 100), era de 4,5 artículos del ISI por profesor a tiempo completo durante esos años. Es decir, alrededor de un artículo por año. En el caso de la Universidad Complutense, que se situaba en la media en productividad (puntuación 0,49 sobre 100), el promedio era de 2,4. Es decir, de 0,5 artículos por año, lo que coincide con la media antes descrita. La UPV que se situaba en la puntuación 0,32 sobre 100, tenía un promedio de 1,4 (Grupo Scimago, 2007). Con estos indicadores, y debido a que la distribución es asimétrica ya que la mediana es inferior a la media, podemos concluir que dos artículos en seis años sitúan a la persona por encima del 50% de los profesores españoles en general. De hecho es el criterio que se plantea para obtener sexenios en Psicología: dos publicaciones del ISI sobre un total de cinco. Podemos suponer que tres publicaciones en cinco años indican una alta productividad relativa.

Utilizando PsycInfo como base, se ha estimado que 6 publicaciones por sexenio es la mediana de los profesores de Psicología de las Universidades de Investigación más productivas de los EEUU. La mediana para Universidades de producción media, a las que se pueden asemejar por condiciones las españolas, era de 3-4 artículos por sexenio (Joy, 2006). El estudio de Musi-Lechuga et al. (2005) que analizó la producción de artículos en revistas del ISI de titulares y catedráticos de Psicología hasta el año 2004 era muy diferente por área y entre universidades. Según la media por área y la media para las cinco universidades con mayor producción por área, las personas con más de 16 y de 26 artículos respectivamente en Psicobiología, con más de 6 y 8 en Metodología, con más de 6 y 7 en Petra, con más de 5 y 7 artículos en Básica, con más de 3 y 4 en Social y con más de 2 en Evolutiva, se sitúan por encima de la media de los profesores y departamentos universitarios. Aunque no se especificaba los años medios de carrera académica, éstos se pueden estimar entre 20 y 25 para los catedráticos. Con el cálculo más exigente, usando el mínimo de años de carrera académica (20) y la estimación más alta de publicaciones (26 para Psicobiología) se puede proyectar que un artículo por año es un indicador de excelencia productiva local. Además las publicaciones se hacían en revistas indexadas nacionales mayormente (Psicothema) por lo que los autores que publiquen mayoritariamente en revistas ISI internacionales se sitúan por encima de la media.

Otros dos indicadores de calidad o de impacto que se ha utilizado son los siguientes:

- Citas por trabajo, es decir el ratio entre total de citas y cantidad de artículos, cuya ventaja es que permite comparar a autores de edad y

carreras diferentes, aunque recompensa una productividad débil y penaliza una productividad fuerte.

Un estudio sobre la producción 2002-2006, basado en Scopus, indica que la media de citaciones por artículo en Psicología es de 6,06 en España (rango 22 en 25 países más productivos), frente a 10,4-10,8 en EEUU, Reino Unido, Cañada, Suecia y Holanda, a 9,1 Italia, 8,6 Alemania y 5,95 Japón. Es decir, un artículo con más de 6 citaciones está por encima de la media local y con más de 10 se sitúa en la producción de mayor calidad mundial (De Moya, 2008).

- Número de publicaciones significativas: número de textos citados más de Y veces. Este indicador no tiene los inconvenientes de los indicadores precedentes, aunque depende de la base y arbitrariedad con la que se define el número Y . Por ejemplo, si utilizamos el criterio de más de 100 citaciones, estaríamos exigiendo valorar solo los artículos del 1% más citado a nivel mundial (usando la base WoS). Exigiendo citaciones por encima de 6 exigiríamos que fuera superior a la media española y por encima de 10 planteamos que sea superior a la media de los países más productivos (usando la base Scopus).

Ante las limitaciones de los indicadores anteriores se han planteado indicadores más sofisticados, que normalizan en relación a una línea base y temporal las citaciones. Algunos ejemplos son los siguientes:

- Índice h de Hirsch: número de veces que un texto ha sido citado más que su rango cronológico en la producción científica de una persona (véase Salgado y Páez, 2008).

- El índice a : o promedio de citaciones de los artículos que conforman el núcleo del índice h .

- El índice g : número de artículos en los que la suma de citaciones es al menos g al cuadrado. Un índice g de 10 indica que el autor ha escrito 10 textos cuya suma de citaciones es al menos de 100 (Egghe, 2006).

El índice h difiere según la base de datos utilizados y las áreas. El índice h sumario de la WoS es inferior al calculado por Publish or Perish. Ahora bien, la correlación entre las h calculadas con diferentes sistemas es alta en general. Utilizar un sistema mantiene los rangos y diferencias relativas. Como dijimos previamente, las publicaciones difieren según las instituciones y las áreas, lo que hace necesario calcular un índice h por área. Por ejemplo, la media de publicación es más baja en Educación y más alta en Psicobiología en el estudio de Salgado y Páez (2007) que analizó este índice en los 10 catedráticos más productivos por área, además de todos los del área de Social. En el caso del área de Metodología de las Ciencias del Comportamiento el valor de la h oscila entre 3 y 7, en Personalidad, Evaluación y Tratamientos Psicológicos entre 7 y 12, entre 6 y 15 en Psicología Básica, entre 4 y 6 en Psicología Evolutiva y de la Educación, entre 8 y 23 en Psicobiología y entre 5 y 10 en Psicología Social. Así pues, de los veinte investigadores con índice h más alto

de la Psicología española, nueve pertenecen al área de Psicobiología, cinco pertenecen a Psicología Básica, tres a Personalidad y tres a Psicología Social (Salgado y Páez, 2007). Es decir, el impacto del área de Psicología social, desde este punto de vista, sería un tercio del impacto del área de Psicobiología, la mitad del de Psicología Básica y parejo al de Psicología de la Personalidad. Utilizando simultáneamente la WoS y el sistema local INRECS como bases de datos, aplicando como criterio el rango mínimo y la media de los 10 investigadores más productivos, podemos concluir que índices h superiores a 8 y 16 respectivamente en Psicobiología, a 6 y 9 en Básica, a 7 y 8 en Petra, a 5 y 7 en Social, a 3 y 5 en Metodología, y a 4 y 5 respectivamente en Evolutiva, indican que la persona se sitúa en una posición de alta productividad relativa en España.

Criticas a los Indicadores de Impacto más complejos

Las críticas que se han planteado a estos índices, en particular al índice h , son las siguientes:

Los índices pueden aumentar con el paso del tiempo aunque el investigador no esté activo.

El índice h correlaciona con la edad. Ahora bien, en general la relación entre edad e índice h es positiva ya que a mayor edad, mayor carrera académica y probabilidad de acumular publicaciones y citas. Sin embargo, en el caso español y para muestras de catedráticos de Metodología y Psicología Social la relación era negativa, lo que se puede explicar porque la orientación a la publicación era menor en los años finales de la Dictadura y de la Transición. En cambio en la muestra total de Metodología, la relación entre edad e índice era cercana a cero (García-Pérez, 2009; Salgado y Páez, 2007). Es decir, en nuestro contexto y con los datos disponibles estas dos críticas no son válidas.

Estos índices sub-estiman a los autores que publican libros (aunque con el sistema de búsqueda *Cited Reference Search* de WoS se incluyen las citas a libros, lo que corrige este sesgo).

Los índices no toman en cuenta la cantidad de autores ni el orden de autoría (aunque perfectamente se pueden establecer índices que dividan por el número de autores y ponderen dando peso a la posición en la autoría). En otras palabras, esto incita al uso de sistemas de búsqueda que incluyan citas a libros (como se hizo en Salgado y Páez, 2007), así como a ponderar por el número de autores (y de orden si este tiene sentido) más que a cuestionar el índice.

Los índices ignoran el total de citas: sin embargo, la relación entre el índice h y el total de citas era de $r=.83$ para una muestra de Psicología Social y de $r=.88$ para una muestra de Metodología.

Los índices no toman en cuenta la inflación del impacto por las auto-citas (aunque éstas se pueden eliminar con los sistemas de bús-

queda). El promedio de auto-citaciones en Psicología, según un estudio utilizando Scopus (De Moya, 2008), es bastante alto en algunos países con mayor impacto científico: 54% en EEUU, superior al 30% en Reino Unido y Alemania, del 30% en España, aunque solo del 20% en los países Escandinavos, que tienen excelente impacto. Glänzel (2008) revisando la evidencia bibliométrica concluye que en general las auto-citaciones no tienen un impacto real. García-Pérez (2009) comparó los índices h directo y el corregido eliminando autocitaciones. El promedio directo era de 2,06 (SD=2,04) y el corregido era de 1,67. Es decir, obviamente era menor, aunque la disminución era solo un 0,20 de la desviación típica. Un 72% de los investigadores mantenía el mismo índice y un 25% disminuía un punto (solo 2 personas perdían tres puntos en su índice h y 10 perdían dos). Podemos concluir que la autocitación en general no afecta a los índices bibliométricos, aunque es importante una revisión individualizada para evaluar personas concretas, ya que al eliminarla una minoría ve disminuido de forma importante la evaluación de su productividad. Una auto-citación superior al 30% situaría a la persona por encima de la media. También es importante examinar si la auto-citación indica seguir una línea consistente de investigación y no una actividad de inflación de las citas.

Estos índices sumarios no toman en cuenta los artículos más importantes de un autor o los muy citados. Efectivamente, la relación entre el mayor número de citas recibidas por una persona y su índice h era menor a la que tiene este índice con otros indicadores como el total de citas o artículos. La relación era solo de $r=.61$ para los catedráticos de Psicología Social y de $r=.41$ para titulares y catedráticos de Metodología (Salgado y Páez, 2007; García-Pérez, 2009).

Pese a estas críticas, el índice h ha mostrado amplia validez concurrente y predictiva. Además de la evidencia revisada y encontrada en nuestro artículo anterior (Salgado y Páez, 2007), estudios recientes han encontrado que el índice h de postulantes aceptados en una institución puntera de investigación biomédica era más alta que el de los rechazados, que los índices h de grupos de 147 grupos de investigación de Químicas correlacionaban con otros indicadores bibliométricos y con la evaluación por pares de la productividad científica (Bornmann y Daniel, 2005, 2008), los índices h nacionales reproducen el rango de otros indicadores bibliométricos de la producción en Psicología (De Moya, 2008) y los índices h individuales correlacionan fuertemente con otros indicadores bibliométricos en muestras de psicólogos españoles y cubanos (Dorta et al, 2008; García-Pérez, 2009). Como limitación contrastada empíricamente, un estudio con físicos encontró que la media y el máximo número de citas eran indicadores fiables y exactos del rendimiento académico, mientras que el índice h era menos exacto como indicador de la productividad científica (Bornmann y Daniel, 2008).

Concluyendo, el índice *h* es bastante convergente con otros indicadores, el impacto de las autocitaciones es limitado, aunque se debe tomar en cuenta en algunos casos. En cambio, la mayor cantidad de citas debería ser un aspecto a tomar en cuenta ya que complementa a índices como el índice *h*. Por otro lado, tomar en cuenta la media de citas también parece necesario.

Indicadores para la evaluación en nuestro contexto

A partir de las medias, medianas y rangos de los estudios antes revisados, que son limitados pero los únicos disponibles hasta que la Administración realice estudios normativos con la población directa implicada en la evaluación, podemos concluir unos parámetros razonables para evaluar positivamente a los investigadores locales. Éstos son los que detallamos a continuación (ver Tabla 1):

Con respecto al factor de Impacto: personas que publiquen en revistas ISI, en particular internacionales, ya muestran un esfuerzo de productividad. Es esperable que estas personas publiquen en revistas de impacto inferior a 1 y entre 1 y 2.

Con respecto a los artículos publicados: dos artículos en revistas internacionales indexadas ISI, en seis años, son un resultado satisfactorio y 3-4 publicaciones internacionales ISI, en 6 años, indican una alta productividad relativa. En relación con la producción total, personas con más de 16 y 26 artículos respectivamente en Psicobiología, con más de 6 y 8 en Metodología, con más de 6 y 7 en Petra, con más de 5 y 7 artículos en Básica, con más de 3 y 4 en Social y de 2 en Evolutiva, se sitúan por encima de la media. Evidentemente esta producción total será diferente según la duración de la carrera científica, la que debe ser tomada en cuenta. Seis artículos por sexenio es una norma que indica de todos modos una excelente productividad, equivalente a la de los grupos de investigación más productivos.

Con respecto a las citas: un artículo con más de 6 citas está por encima de la media local y arriba de 10 se sitúa en la producción de mayor calidad mundial. Un porcentaje de artículos con citas superiores a la media de alrededor del 50% se puede considerar otro criterio normativo razonable.

Con respecto a indicadores más sofisticados: índices *h* superiores a 6-7 en general y de 8 en Psicobiología, a 6 en Básica, a 7 en Petra, a 5 en Social, a 4 en Evolutiva, y a 3 en Metodología, indican una alta productividad relativa. Una auto-citación de alrededor del 30% situaría a la persona evaluada en la media. Auto-citaciones elevadas deberían ser contrastadas si reflejan consistencia en la investigación o auto-inflación deliberada.

Tomar en cuenta los artículos más citados agrega una información complementaria importante a los anteriores criterios.

Tabla 1
Criterios normativos de evaluación de la producción científica
según la evidencia empírica

Evaluación	Factor de Impacto	Artículos publicados	Nº de Citaciones	Índice h
Buena	1 ó menos	2 en internacional ISI	6	6-7*
Muy buena	1 y 2	3-4 "	10	8 ó más

* Con relación a las áreas, índices h de 8 en Psicobiología, de 6 en Básica, de 7 en Petra, de 5 en Social, de 4 en Evolutiva y de 3 en Metodología.

Con este trabajo esperamos haber situado en una base empírica, aunque limitada, los criterios evaluativos de la producción científica de la Psicología española. Además creemos que los tres últimos criterios presentados tienen un peso superior a los anteriores, aunque sería un error limitar a un solo indicador la evaluación de la productividad científica.

Nota: La investigación incluida este trabajo ha sido financiada parcialmente por el Ministerio de Educación y Ciencia con cargo al proyecto de investigación SEB1098-2005. Dirección de contacto: Darío Páez, Departamento de Psicología Social y Metodología, Avda. De Tolosa, 70, 20018 Donostia y Jesús F. Salgado, Departamento de Psicología Social, Básica y Metodología de la Universidad de Santiago de Compostela, 15782 Santiago de Compostela. E-mail: dario.paez@ehu.es y psjesal@usc.es.

Referencias

- Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA) (2006): Manual de procedimiento para la emisión del informe conducente a la mención de calidad a los estudios de doctorado. [Accesible via http://www.aneca.es/modal_eval/docs/pdc_05_manual.pdf]
- Aguirre,A.(2003): La Investigación española crece hasta el 3% de la producción mundial. ABC, 28/06/2003.
- Alcain,M.D.-Román,A.(2005): Hacia una valoración integrada de las revistas españolas de Ciencias Sociales y Humanas: las revistas de Psicología. *Psicotema*, 17, 179-189.
- Allen,M.(1996): Research Productivity and Positive Teaching Evaluation: an examination of relationship using meta-analysis. *Journal of the Association for Communication of Administration*, 2, 77-96.
- Avital,M.-Collopy,F.(2001): Assesing Research Perfomance: Implications for selection and motivation. *Sprouts: Working Papers on Information Environments, Systems and Organizations*, 1, 40-61. [Accesible via <http://sprouts.case.edu/2001/010303.pdf>.]

- Beauvois, J.L.-Pansu, P.(2008): Facteur d' Impact et Mondialisation Culturelle. *Psychologie Française*, 53, 211-22.
- Bornmann, L.-Daniel, H.D.(2005): Does the h-index for ranking scientists really work? *Scientometrics*, 63, 391-392.
- Bornmann, L.-Daniel, H.D.(2008): What do we know about the h-index? *Journal of the American Society for Information Science and Technology*, 58, 1381-1385.
- Buela-Casal, G.-Bermudez, M.-Sierra, J.C.-Quevedo-Blasco, R.-Castro, A.(2009): Ranking de 2008 en productividad en investigación de las universidades españolas. *Psicothema*, 21, 309-316.
- Cabrales, A.-Carmona, S.-Fuentes, L.-Gomez, J.-Gutierrez, R.-Herrero, C.-Jordana, J.-Olcina, J.-Páez, D.-Sánchez, V.(2007): Informe sobre Criterios de Evaluación de la I+D en Ciencias Sociales. Tercer Grupo de Trabajo ANEP, Enero-Octubre, 2007. Madrid: ANEP.
- De Moya, F.(Director)(2008): Indicadores Bibliométricos de la Actividad Científica Española 2002-2006. Grupo de Investigación Scimago. Universidad de Granada. Madrid: Fundación Española para la Ciencia y la Tecnología.
- Dorta-Contreras, A.J.-Arencibi-Jorge, R.-Marti-Lahera.-Araujo-Ruiz, J.A.(2008): Productividad y visibilidad de los neurocientíficos cubanos. *Revista de Neurología*, 47, 355-360.
- Egghe, L.(2006): An improvement of the H-index the G-index. *ISSI Newsletter*, 2, 1, 8-9. Society for Information Technology, in press.
- Fairweather, J.S.(2002): The Mythologies of Faculty Productivity. *The Journal of Higher Education*, 73, 26-48.
- Feldman, K.A.(1987): Research productivity and scholarly accomplishment of college teachers as related to their instructional effectiveness: A review and exploration. *Research in Higher Education*, 26, 227-298.
- Fundación Conocimiento y Desarrollo (2005): La universidad en España. [Accesible via <http://www.fundacioncyd.org/Publicaciones/InformeCyD.asp>]
- Fundación Conocimiento y Desarrollo (2006): La universidad en España. [Accesible via <http://www.fundacioncyd.org/Publicaciones/InformeCyD.asp>]
- Fundación Conocimiento y Desarrollo (2007): La universidad en España. [Accesible via <http://www.fundacioncyd.org/Publicaciones/InformeCyD.asp>]
- García-Perez, J.(2009): The H index in Methodology of Behavioral Science. Manuscrito.
- Glänzel, W.(2008): Seven Myths in Bibliometrics. En H. Krestchmer y F Havemman (Eds.). *Proceedings of Fourth International Conference on Webometrics, Informetrics and Scientometrics 2008*, Berlin. [Accesible via <http://creativecommons.org/licenses/by/2.0>]
- Hamilton, D.P.(1991): Research papers: Who's uncited now?. *Science*, 251, 25.
- Hattie, J.-Marsh, H.W.(1996): The relationship between research and teaching: A meta-analysis. *Review of Educational Research*, 66, 507-542.
- Íñiguez, L.-Martínez, L.M.-Muñoz, J.M.-Peñaranda, M.C.-Sahagún, M.A.-Alvarado, J.G.(2008): The mapping of Spanish Social Psychology through its conferences: a bibliometric perspective. *The Spanish Journal of Psychology*, 11 (1), 137-158.
- ISI(1999): Citations reveal concentrated influences: some fields have it, but what does it mean, *Science Watch*, January/February, 1-2
- Jacso, P.(2008): The plausibility of computing the h-index of scholarly productivity and impact using reference-enhanced data bases. *Online Information Review*, 36, 266-283 [Accesible via <http://www.emeraldinsight.com/1468-4527.htm>]
- Johnson, R.J.-Wolinsky, F.D.(1990): The Legacy of Stress Research: the course and impact of this journal. *Journal of Health and Social Behavior*, 31, 217-239.

- Joy,S.(2006): What I Should be doing, and Where are they doing it? Scholarly Productivity of Academic Psychologists. *Perspectives in Psychological Science*, 1, 346- 364.
- Katz,J.S.(1999): *Bibliometric Indicators and Social Science*. ESRC SPRU, University of Sussex, UK.
- Kermarrec,A.M.-Faou,E.-Merlet,J.P.-Robert,P.-Segoiufin,L.(2007): Que mesurent les indicateurs bibliométriques?. Document de la Comission d'Analyse de l'Institut National de Recherche en Informatique et Automatique. Paris: France.
- King,D.A.(2004): The scientific impact of nations. *Nature*, 430, 311-316.
- Lee,J.D.-Cassano,A.-Vicente,K.J.(2005): Bibliometric análisis of Human Factors (1970-200). *Human Factors*, 47, 753-766.
- Musi-Lechuga,B.-Olivas-Avila,J.A.-Portillo-Reyes,V.-Villalobos-Galvís,F. (2005): Producción de los profesores funcionarios de Psicología en España en artículos de revistas con factor de impacto de la Web of Science. *Psicothema*, 17, 539-548.
- Páez,D.(2007): *La Evaluación de la Calidad de la Investigación en la Universidad Española*. Comunicación al Seminario Evaluación de la Enseñanza Superior, Facultad de Psicología, San Sebastián, 5 de Diciembre del 2007.
- Pestaña,A.-Gómez,I-Fernández,M.T.-Zulueta,M.A.-Méndez,A.-Koenig,M.E.D.-Bookstein,A.(1995): Scientific evaluation of RyD activities in medium size institutions: a case study on the Spanish CSIC. En M. Koenig y A. Bookstein (Eds.) *The Proceedings of the Fifth International Conference on the International Society for Scientometrics and Informatics*, N°5, River Forest IL, USA, pp425-434.
- Salgado,J.-Páez,D.(2007): La productividad científica y el índice h de Hirsch de la Psicología Social española. *Psicothema*, 19, 179-189.
- Scimago(2007): La productividad ISI de las Universidades Españolas (2000-2004): El Profesional de la Información, 16, 354-358.
- Schoepflin,U.(1990): Problems of representativity in the Social Science Citations Index. In P. Weingart, R. Schringer y M. Winterhager (Eds.). *Representations of Science and Technology*, Bielefeld, Germany: DSWO Press, pp. 177-188.
- Simonton,D.K.(2003): Scientific creativity as constrained stochastic behavior: the integration of Product, Person, and Process Perspective. *Psychological Bulletin*, 129, 475-494.
- Simonton,D.K.(2007): Psychology's limits as a scientific discipline: A personal view. *Applied and Preventive Psychology*, 12, 35-36.
- Thibergien,G.-Beauvois,J.L.(2008): Domination et Impérialisme en psychologie. *Psychologie Francaise*, 53, 135-155.
- Van Raan, A.F.J.(2005): Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62, 133-143.
- Van Raan,A.F.J.(2006): Comparison of the H-Index with standards bibliometrics indicators and with peers judgments for 147 chemistry research groups. [Accessible via <http://arxiv.org/ftp/physics/papers/0511/0511206.pdf>].