

Idoneidad de los indicadores de calidad de la producción científica y de la investigación¹

Appropriateness of quality indicators of the scientific and research production

Ana M. GONZÁLEZ RAMOS, Teresa GONZÁLEZ DE LA FE¹
Fernando FERNÁNDEZ PALACÍN y Manuel MUÑOZ MÁRQUEZ²

¹Dpto. de Sociología. Universidad de La Laguna

²Dpto. de Estadística e Investigación Operativa. Univ. de Cádiz

Recibido: 17.04.06

Aprobado: 03.05.06

RESUMEN

Cada vez somos más conscientes de que la producción del conocimiento científico ha de ser medida y evaluada para conocer la importancia de las aportaciones realizadas y la calidad de los resultados obtenidos por los investigadores. Solo el curso del tiempo ha dado lugar a un acuerdo implícito sobre esta cuestión, con ideas generales que afectan a toda la comunidad científica, pero consensuadas de forma particular por cada especialidad. En este trabajo abordaremos estas cuestiones, lo que nos llevará a plantear las dificultades de obtener un buen indicador de calidad, atendiendo sólo a un grupo limitado de medidas. Esto es especialmente evidente cuando consideramos la distancia existente entre áreas de conocimiento, que tienen culturas de investigación y, por tanto, de evaluación de la calidad de su trabajo, muy dispares. Por eso plantearemos que la medición de la calidad científica debe abordarse desde un punto de vista integral y considerando su naturaleza multifactorial.

PALABRAS CLAVE: Indicadores de calidad, metodología, evaluación científica, productividad, factor de impacto.

ABSTRACT

There is increasing awareness about the necessity of measuring and assess the production of scientific knowledge to understand the importance of the contributions and the quality of scientific results. Only the course of time has given rise to an implicit consensus about this issue, with general ideas that affect

¹ Esta investigación ha sido posible gracias a la financiación de la Consejería de Educación y Ciencia de la Junta de Andalucía a través del proyecto TELOYDISREN (FQM-270).

to all the scientific communities, but reached by agreement within each speciality. In this work we address these questions, which will lead us to expose the difficulties of obtaining an efficient indicator of quality by paying attention only to a narrow group of measures. This is especially evident when we examine the distance between knowledge areas, with different research cultures and hence, different methods of assessing the quality of their work. Thus, we will argue that the assessment of the scientific quality must be carried out from an integral point of view by taking into account its multifactorial nature.

KEY WORDS: Quality indicators, methodology, scientific assessment, productivity, impact factor.

SUMARIO

1. Introducción. 2. Diversidad y naturaleza de las medidas de evaluación. 3. Utilidad y modelos de evaluación de la calidad científica. 4. Poniendo a prueba los métodos de evaluación. 5. Conclusión. 6. Bibliografía.

1. INTRODUCCIÓN

La necesidad de crear mecanismos de control de la calidad de la producción científica es indiscutible, tanto por sus ventajas comparativas como por su capacidad de discriminación entre las contribuciones llevadas a cabo por los investigadores. En la actualidad, este asunto ha cobrado una relevancia especial puesto que los cambios normativos, producidos en la esfera de la ciencia y la tecnología, lo sitúan en el punto de mira de los agentes institucionales y de los propios investigadores. Para todos ellos es evidente la utilidad de poseer un mecanismo objetivo de reconocimiento del trabajo científico de los investigadores, en un doble sentido: el de evaluación (es decir, ser capaces de realizar una comparación discriminante) y de valoración (mediante la cual serían capaces de otorgarse recompensas). Ahora bien, no existe un acuerdo unánime sobre los factores que determinan la calidad de las aportaciones científicas y mucho menos, sobre cómo determinar su excelencia. El término calidad es un concepto variable que suele adjudicarse teniendo en cuenta varios criterios diferentes que suelen aplicarse de manera independiente.

Así pues, puede considerarse que la excelencia de un autor reside en que escribe muchos artículos o en que consigue abarcar cuantiosos fondos, sin indicar siempre a la misma persona. En definitiva, su excelencia depende del factor que apliquemos para evaluar su trabajo. Por si fuera poco, esos criterios varían de una especialidad a otra, de modo que en matemáticas no se tiene en cuenta el mismo conjunto de factores que en filología. La lista de elementos que se utiliza para evaluar la calidad del trabajo científico es tan amplia como las opiniones que las sustentan. Ello dificulta la determinación de un único resultado en la evaluación porque depende de la aplicación de cada medida. No hay unanimidad, por ejemplo, sobre el valor de la edición de libros frente a la publicación de artículos científicos, todo depende de la opinión que se reserve cada rama del conocimiento sobre su importancia, lo que guiará el comportamiento de los investigadores. En las ciencias sociales y humanas publicarlos constituyen un gran logro, sin embargo, en las ciencias naturales los libros suelen tener un carácter didáctico o divulgativo, mientras que los hallazgos científicos originales se publican en las revistas especializadas

(Becher, 1989; González Ramos, 2004). La diversidad de comportamientos no sólo se debe a las pautas culturales de cada especialidad, también depende de las decisiones individuales, coyunturales, que tome el grupo de investigación, por eso es posible encontrar distintos comportamientos dentro de la misma rama científica (González Ramos, 2004). También hallaremos resultados de evaluación muy diferentes si aplicamos criterios distintos en la determinación de la calidad y el prestigio de una institución en base, por ejemplo, de su popularidad o de su nivel de producción científica medida a través del número de proyectos obtenidos o, quizá, alternativamente, a través del número de publicaciones conseguido por el conjunto de su plantilla.

La aplicación de estos criterios de medición suele utilizarse de manera independiente en base a las bases lógicas que el evaluador haya preferido para definir el concepto de excelencia y calidad. ¿Pero qué ocurriría si aplicáramos sobre el mismo universo todos estos criterios de evaluación? ¿Coincidirían los resultados en señalar inequívocamente a aquellos agentes más valiosos del sistema o por el contrario, cada uno apuntaría a distintos elementos del contexto? En este artículo hemos realizado la comprobación empírica y el resultado es que los distintos criterios identifican a distintos agentes como los más destacados de su entorno (o, en el otro extremo de la escala, los menos destacados). Es posible que los primeros y los últimos puestos recaigan sobre un grupo más o menos amplio de instituciones o individuos, aunque variable en algunos casos, pero desde luego, la mayoría de ellos se sitúa en distintas posiciones de la escala. De este modo, los resultados nos llevan a concluir que la existencia de distintos indicadores de medición proporciona una gran ambigüedad sobre nuestro objetivo: la determinación de la calidad científica e investigadora.

A lo largo de este trabajo repasaremos, en primer lugar, distintos modelos de evaluación del conocimiento, lo que nos permitirá estar en disposición de mostrar la fundamentación sobre la que se sostiene su construcción. A continuación, aplicaremos esos indicadores al caso de la Universidad de Cádiz, mostrando así la robustez de las medidas y también las debilidades que se ponen de manifiesto al compararse entre sí. Finalizaremos nuestro trabajo con una valoración de todos esos aspectos desde un punto de

vista aplicado, es decir, con la intención de ampliar los límites del debate relativos a esta cuestión, desembocando en la construcción de una medida más eficaz de la calidad científica.

2. DIVERSIDAD Y NATURALEZA DE LAS MEDIDAS DE EVALUACIÓN

El conjunto de instrumentos destinados a evaluar la calidad de la actividad científica se ha extendido a lo largo del tiempo bajo la influencia de varios factores determinantes. Uno de ellos es el perfil del investigador, el interés que anima a los gestores y el que persigue los investigadores es muy diferente, por tanto, las medidas empleadas para medir la calidad científica del trabajo realizado reflejarán esas distintas motivaciones. No podemos ignorar tampoco, que las medidas se aplican a diversos contextos, tratando de captar tanto la calidad de unidades agregadas como de los sujetos individuales, lo que supone realizar abordajes muy diferentes sobre el mismo asunto. A pesar de que, en algunos casos, los instrumentos de medida se utilicen indistintamente, su valor no puede ser el mismo. Pero quizá, los aspectos más significativos en relación a la distinta construcción de los indicadores empleados para medir la calidad científica sean aquellos relacionados con decisiones metodológicas. En este sentido, encontramos herramientas de medición de carácter objetivas y otras de carácter subjetivas que, por supuesto, nos ofrecerán resultados de muy distinta índole.

Los indicadores cuantitativos muestran factores contables y estadísticos de las instituciones o de la producción científica realizada por sus miembros. Fueron usados, en primer lugar, para evaluar las universidades estadounidenses, pero pronto pasaron a ser una herramienta básica de evaluación de cualquier centro de producción del conocimiento de cualquier país occidental. En todo momento estuvieron auspiciados por las directrices elaboradas por los organismos internacionales en cumplimiento de una misión evaluadora y contable como los manuales de Frascati, Oslo y Canberra. Las encuestas, en cambio, tratan de obtener una valoración del sistema de

conocimiento a partir de la opinión expresada por los agentes implicados en él, por eso son instrumentos de naturaleza subjetiva. Con ellos se indaga acerca de la calidad de la actividad investigadora o del estado de la ciencia y la tecnología en general a partir de la opinión aportada por los propios actores o de la población en general (De Miguel *et al*, 1999; ANECA, 2004). Estas herramientas de evaluación se sostienen sobre la base de la valoración que hacen los propios actores, sobrentendiendo que todos los agentes están en disposición de tener un juicio categórico sobre estos temas. Sin embargo, en nuestra opinión, es preciso no confundir la disposición de dar una opinión con la capacidad de ofrecerla².

Independientemente de la idoneidad de estas medidas y de su mayor o menor precisión a la hora de lograr su cometido, la mayor dificultad que presentan todas ellas se deriva de la ausencia de un debate teórico profundo sobre lo que significa el concepto de calidad científica. La construcción del sistema de evaluación debe fundamentarse en todo lo que conocemos acerca de este asunto, tanto en términos teóricos como en términos empíricos, es decir, a partir de las deducciones que pueden extraerse del resultado de la aplicación de estos indicadores sobre contextos reales. Hasta ahora, se aceptan unos criterios u otros, más por la costumbre o por una toma de postura ideológica que por haber llevado a cabo una reflexión profunda sobre la idoneidad de la utilización de esa medida. Por eso, es necesario someter a debate estas medidas y ponerlas a prueba y así conocer con qué precisión reflejan el concepto que tratan de designar. El hecho de que ya se hayan ideado diferentes procedimientos para la medición de la producción científica nos proporciona mayores ventajas, estamos en mejor disposición para comparar todos esos indicadores con los datos disponibles, y acercarnos, de esta manera, a la verdadera naturaleza del concepto.

En este trabajo pondremos a prueba algunas de esas medidas, evaluando su extensión significativa y su eficiencia al indicar la realidad que trata de medir. Para ello realizaremos varias comprobaciones aplicando diversos indicadores cuantitativos sobre el sistema de producción

² No queremos menospreciar la valoración de la opinión pública pero sí puntualizar que a veces, podemos encontrarnos con personas poco dispuestas, interesadas o poco informadas sobre los aspectos sobre los que preguntamos. El empleo de encuestas en otros casos muy valiosos, en nuestra opinión serían menos adecuadas para tratar estos temas.

científica de la Universidad de Cádiz. Pondremos a prueba sólo medidas objetivas por que están más estandarizadas y, sobre todo, por que en los últimos tiempos, han sido aceptadas por gran parte de la comunidad científica como instrumentos de evaluación poderosos. Hemos elegido el estudio de casos para llevar a cabo este trabajo porque facilita la demostración de las distintas vertientes que conforman las dimensiones del concepto calidad de la producción científica, tanto para esclarecer su contenido y significado como para evaluar su fiabilidad y reconocer las ventajas e inconvenientes que conllevan su uso. Con todo ello, esperamos contribuir a ampliar la fundamentación teórica y empírica relacionada con esta cuestión.

3. UTILIDAD Y MODELOS DE EVALUACIÓN DE LA CALIDAD CIENTÍFICA

Con el transcurso del tiempo, todos reconocemos cuándo un modelo de conocimiento ha sido superado por otro y cuándo sus programas de investigación se muestran anticuados ante los retos que plantean los desafíos centrales de esa disciplina. Tradicionalmente, la sociología de la ciencia se ha ocupado de cuáles son los momentos de transición y qué factores hacen que unos programas científicos se sitúen por encima de otros, superando explicaciones alternativas que compiten entre sí. Sin embargo, resulta más complicado identificar esos momentos de transición, en el mismo momento en que se está produciendo. Los vaivenes históricos y el desarrollo de técnicas de investigación cada vez más sofisticadas impiden una identificación inequívoca de la relevancia de las aportaciones científicas en el mismo momento en que se está produciendo su descubrimiento. En este caso, la cuestión se dilucida a través del juicio de sus contemporáneos, que acuerdan la relevancia y trascendencia de esas aportaciones innovadoras. La calidad del trabajo realizado suele ser el criterio al que se acude para determinar la relevancia de ese esfuerzo, y los especialistas de esa área del conocimiento son los encargados de descubrir la valía del esfuerzo de sus colegas de profesión. Por tanto, el juicio de los pares es el primer mecanismo que pone en marcha los sistemas de evaluación del trabajo científico entre colegas. Sin embargo, este sistema es tan necesario como imperfecto porque se deriva de la

cultura de cada disciplina e incluso de razones oportunistas que se ciernen sobre ese momento histórico. Muchos descubrimientos han «padecido» esas consecuencias, pasando inadvertidas o castigándose con el ostracismo y el rechazo en el momento de su descubrimiento, para ser de la máxima importancia en el futuro. El sistema de recompensas basado en la calidad del trabajo producido tiene, como nos recuerda Merton (1977: 419-438) un valor instrumental desde el punto de vista individual y social, a través del cual se estructura el sistema de reconocimiento que potencia la intensidad del trabajo científico. Sin embargo, también es un sistema imperfecto puesto que recompensa de manera desigual a los investigadores, pues son evaluados no en función de su valía sino en cuanto a la acumulación de recompensas y recursos previos (Merton, 1977: 447).

Así pues, los modelos de evaluación de la calidad del trabajo científico cumplen distintas funciones que afectan a otros tantos niveles del sistema, individuales y colectivos. Por una parte, establece un mecanismo de reconocimiento honorífico para aquellos actores sociales que producen un descubrimiento o un trabajo de calidad. Pero también se convierte en el estímulo por su valor instrumental, indispensable para el avance de la ciencia pues identifica a los programas científicos en relación a su relevancia para un área del conocimiento concreto. Y, desde luego, permite establecer un criterio de decisión acerca de la rigurosidad del trabajo realizado, lo que nos remite directamente a la calidad del procedimiento empleado y a la significación de los resultados alcanzados desde un punto de vista técnico. Este es seguramente, el sentido más valioso pues refleja más fielmente el esfuerzo realizado por los científicos o, mejor aún, la importancia de los logros científicos alcanzados. Sin embargo, también es el más difícil de obtener a juzgar por el número, prácticamente nulo, de intentos que se ha llevado a cabo.

La evaluación era al principio, más que nada, un asunto de intelectuales, interesados por las condiciones del ambiente cultural y científico de la sociedad (Ortega, 1930). Pero, como ya dijimos anteriormente, pronto las instituciones universitarias comenzaron a preocuparse por la evaluación de sus instituciones, por los resultados alcanzados y la detección de problemas sobre los cuales habría que habilitar procedi-

mientos de mejora. Esta actividad será desarrollada más tarde, por otras instancias políticas interesadas en el establecimiento de indicadores básicos para demostrar la situación de los distintos ámbitos del sistema social, como la salud, el empleo y también, por tanto, el propio sistema educativo, de la ciencia y la tecnología. Los organismos oficiales son los encargados de cumplir esta función de recogida de información con carácter comparativo, sobre los sistemas sociales. Esta perspectiva evaluativa proporciona un enfoque macrodimensional y está relacionada con el diseño de políticas públicas de gestión y de promoción de la actividad científica e investigadora. Desde los años sesenta del siglo veinte, estos modelos de evaluación han seguido las directrices marcadas por los manuales de Frascati, Oslo y Canberra, que se han ido sucediendo en diferentes versiones, hasta la actualidad (Godin, 2003).

Más tarde, serían otras instituciones y agencias privadas las que entrarían en este espacio, ofreciendo diversos modos de medición de la calidad del sistema. Sus instrumentos proporcionaban, al mismo tiempo, una dimensión comparativa y un método para establecer un sistema de recompensas para diferenciar a aquellos autores que desempeñaban una actividad científica más importante. Aparece una nueva área de trabajo, la bibliometría, que prometía solucionar de una vez el problema de la evaluación del trabajo científico con la construcción de una medida objetiva basado en el recuento del número de artículos escritos por una persona o el prestigio de las revistas donde aparecían los artículos y la cantidad de citas recibidas. El método presentaba la ventaja de que producía instrumentos de medición estandarizados, fácilmente aplicable tanto a instituciones como a autores individuales, quienes hasta el momento solo podía ser recompensados y valorados mediante el reconocimiento cualitativo de los expertos de esa disciplina o la concesión de los premios ofrecidos a la labor global de una carrera excelente (los premios Nobel, por ejemplo) o la consecución de un logro significativo para su campo del conocimiento (como los premios de la física y las matemáticas que se hicieron tan populares en las Academias de las Ciencias del siglo XIX).

A todas estas medidas les une una característica común, la necesidad de expresar «resultados» de la actividad producida por el sistema. En efecto, si medimos los elementos integrantes

del sistema, tales como personas, infraestructuras o clientes es para conocer el resultado de su actividad, es decir, la cantidad de productos y servicios generados por las unidades presentes en el sistema. Los sistemas de evaluación contables permiten acceder fácilmente a la medición de los resultados del sistema, por ejemplo, el número de artículos, el dinero recaudado mediante actividades de investigación, la cantidad de tesis leídas, etcétera. Estos modelos de evaluación han adoptado una perspectiva contable de los logros alcanzados por una institución o conjunto de investigadores concretos. De este modo, las unidades medidas se convierten en elementos objetivos, cuantitativos y fácilmente mensurables, incluso susceptibles de convertirse en unidades económicas. Porque, en efecto, este tipo de evaluaciones están impregnadas de una filosofía de corte económico, que pretenden estimar el esfuerzo realizado por los agentes de producción del conocimiento científico. La operación se convierte en una contabilización de los recursos y de los resultados que se extraen de su aplicación directa sobre su contexto y, en definitiva, que comprueba la eficiencia del sistema. Los indicadores construidos para tal fin determinan la influencia y el peso relativo de diversos factores de salida del sistema (*output*), como los recursos materiales y humanos de que dispone o la combinación de esos elementos entre sí (Bellavista *et al*, 1997).

Dentro de este conjunto de medidas podemos mencionar el número de proyectos y patentes conseguidos por los investigadores, el número de artículos publicados y la visibilidad de sus trabajos para plasmada a través de la cantidad de citas recibidas y el factor de impacto de las revistas especializadas donde se publican sus trabajos. Institucionalmente, estos son los elementos que han triunfado sobre otros posibles indicadores a la hora de establecer la calidad de la producción científica de un equipo de investigación, un investigador o de otros niveles agregados de la observación, como el conjunto de la actividad de la institución del conocimiento o la entidad territorial a la que hacemos referencia. Son los elementos que debemos reseñar cuando elaboramos una petición de financiación en una convocatoria pública o requerimos a cualquier institución una evaluación de nuestros méritos profesionales. Todos estos elementos parecen ser *a priori*, por su objetividad y fácil aplicación, los mecanismos idóneos para establecer un

sistema de evaluación que permita comparar los méritos logrados por los investigadores.

Aunque estos criterios hayan sido aceptados, por lo general, también han recibido muchas críticas (seguramente porque nadie ha ofrecido una explicación convincente de que la cantidad suponga calidad). Se les ha objetado la desigualdad de oportunidades de la que parten los investigadores por la mera pertenencia a un área del conocimiento u otro, o a una institución más prestigiosa y con más recursos e infraestructuras que otra e, incluso, a las características personales de los investigadores como la edad y la procedencia geográfica, factores que favorecen el efecto Mateo³. Sin embargo, las que han despertado mayor número de críticas son las relacionadas con el número de citas y el índice de impacto de las revistas científicas. Las objeciones presentadas se centran en el origen de esas mediciones, que tienen como base la fuente de datos del Science Citation Index (SCI), del Institute for Scientific Information, a quienes se recrimina el establecimiento de indicadores poco fiables (Rousseau, 2002: 429; Maltrás, 1996: 113-151), por que contienen un triple sesgo, disciplinar, geográfico y lingüístico, que recae sobre el propio proceso de recopilación de datos. Los artículos y revistas que conforman el catálogo del SCI pertenecen a ciertas áreas de conocimiento, especialmente centradas en las ciencias naturales y, en algunas especialidades frente a otras. Predominan las revistas escritas en inglés sobre cualquier otro idioma y las editadas en países del área anglófona por encima de las editadas en otros lugares. En resumen, la base de datos cuenta con lagunas importantes en cuanto a la recogida de datos que afecta especialmente a ciertos autores, que podrían publicar en revistas de otras nacionalidades, sobre ciertas ramas o especialidades concretas no recogidas tan exhaustivamente en esta base de datos y/o que fueran escritas en un idioma distinto al inglés (King, 2004; Jiménez-Contreras *et al*, 2003; Rousseau, 2002; Maltrás, 1996).

Otras críticas recibidas no se refieren a cuestiones técnicas sino a otros aspectos al que hicimos referencia anteriormente. Esto es, al hecho de que no existen pruebas acerca de que la mera

alusión a una obra o, lo que es lo mismo, al número de citas recibidas, sea un buen indicador de la calidad de la obra, ni siquiera si éstas se producen en una cantidad muy elevada (King, 2004). Las citas pueden ser de muchos tipos, pueden proceder de comentarios críticos, de autocitas (es decir, referencias que el autor hace sobre su propia obra) o ser el resultado de «citas extrañas» (Rubio, 2001). Algunas de estas prácticas responden a otras funciones distintas al reconocimiento, que es la razón en la que se fundamenta la construcción de esta medida y sobre la que se sostiene su equiparación con el concepto de calidad de la obra científica. En resumen, aunque sin duda es un factor que nos aproxima a su medición, nada asegura que la cantidad de citas recibidas implique de manera indiscutible su calidad (Buela-Casal, 2002). Sin embargo, la utilización de dichos indicadores está tan extendida que, aunque muchos reconocen sus defectos, nadie prescinde de utilizarlos, aún con reservas. Las ventajas parecen compensar a los inconvenientes de su empleo, éstas estriban en la capacidad de síntesis del índice y su fácil identificación con el concepto que trata de medir, lo que ha favorecido su extensión y popularización entre los científicos de las ramas científico-técnicas. Con el paso del tiempo, este comportamiento ha servido de modelo para el resto de las ciencias porque, desde el punto de vista de la gestión, este indicador es muy útil para trabajar en políticas de promoción del conocimiento científico.

Otro de los indicadores empleados para reconocer el trabajo y la excelencia de los investigadores más relevantes dentro de su campo son los galardones, tema del que se ocupó ampliamente Merton (1973), pero que, por su propia naturaleza, se aleja de los propósitos que inspiran a los modelos de evaluación de la ciencia. Los premios destacan el trabajo excepcional de un círculo muy exclusivo de investigadores o realizado sobre un momento concreto de descubrimiento científico. Por eso mismo, son poco operativos para clasificar el trabajo de todo un colectivo o para desarrollar una evaluación periódica del trabajo de los investigadores. Esta reflexión nos permite adelantar que no sólo es

³ El efecto Mateo fue definido por Merton (1977) para poner de manifiesto las desigualdades que se producen en el ámbito de la ciencia, donde los científicos en mejor posición obtendrían más fácilmente más recursos para lograr sus propósitos. Por el contrario, los más jóvenes o con menos recursos tendrían más difícil acumular méritos y recursos para desarrollar su tarea.

necesario encontrar métodos que permitan destacar a los profesionales valiosos, sino que sean sensibles en la distinción entre distintos niveles del sistema susceptibles de ser comparados entre sí. Es decir, que clasifique el trabajo realizado por los investigadores, tomados por separado o colectivamente y, por supuesto, que sea capaz de mostrar esos aspectos a pesar de los diferentes «resultados» que se muestran en cada área del conocimiento y que se deben no a un menor rendimiento sino al establecimiento de pautas culturales distintas que modifican la actividad científica en las distintas especialidades. En definitiva, los indicadores que necesitamos para evaluar la producción del conocimiento científico deben denotar los diferentes aspectos que conlleva el trabajo de investigación, construyendo una escala que permita graduar el esfuerzo realizado por los investigadores.

4. PONIENDO A PRUEBA LOS MÉTODOS DE EVALUACIÓN

4.1. METODOLOGÍA DEL TRABAJO EMPÍRICO

La mejor manera de comprobar la validez de un instrumento de medición es sometiendo a prueba. En nuestro caso, esta operación requiere que devolvamos esta herramienta de medida a la realidad, para conocer el grado en que refleja el fenómeno que se quiere medir y su verdadera capacidad explicativa. Los trabajos emprendidos para probar las herramientas de medida son poco habituales en algunas áreas de las ciencias sociales, quizá porque la replicación de estudios es mucho más costosa que el sometimiento a las pruebas realizadas en un laboratorio propio de las ciencias experimentales. Sin embargo, esta es la idea inspiradora de nuestro trabajo, donde sometemos a prueba a los indicadores de calidad científica con la intención de evaluar el esfuerzo y el reconocimiento o, al menos, la visibilidad de los investigadores entre sus colegas de especialidad. Esto es, en razón del número de publicaciones, de citas recibidas y de la publicación en revistas de alto factor de impacto. Por eso hemos escogido un contexto bien delimitado, que nos sirva como campo experimental, apropiado para comprobar la eficacia de los indicadores que queremos validar.

La utilización de un estudio de casos responde, precisamente, a la necesidad de escoger un

espacio restringido sobre el que poder profundizar en el tema objeto de estudio, que funcionará como un modelo representativo de la realidad. El contexto elegido es el sistema de producción el conocimiento de la provincia de Cádiz, básicamente estructurada alrededor de su Universidad (González Ramos, 2004). Hemos aplicado los distintos índices empleados más habitualmente para evaluar la calidad de la producción científica para conocer los resultados que se desprenden de su medición. Nuestro propósito es conocer la consistencia de dichas medidas, comprobando la potencia explicativa y la fiabilidad de esos indicadores. Partimos de la base de que la contrastación empírica nos permitirá establecer, en primer lugar, la calidad de los datos que se extrae de esos instrumentos de análisis, de modo que estaremos en disposición de comprobar la fiabilidad de las herramientas de medición. En segundo lugar, debemos precisar que nuestro propósito no es conocer el estado de la investigación en Cádiz, sino que utilizamos este espacio como campo de pruebas, para comprobar la validez y fiabilidad de las herramientas de análisis que suelen emplearse como indicadores de calidad del conocimiento científico.

Una vez identificados los objetivos y la población objeto de estudio determinaremos cuáles son las fuentes de datos y los instrumentos de medición utilizadas, así como el procedimiento llevado a cabo para explorar nuestras líneas de trabajo. En este estudio, hemos contado con diversas fuentes de información procedentes de las bases de datos oficiales que recogen la producción científica de los investigadores andaluces. De ellas extraeremos los indicadores objetivos que someteremos a prueba. Por una parte, el Plan Andaluz de Investigación (PAI) recoge toda la información relativa a los grupos de investigación que operan en la Comunidad Autónoma Andaluza. Su fiabilidad está asegurada pues a partir de ella, los grupos consiguen financiación y el reconocimiento dentro del sistema de ciencia y tecnología regional. Además de otros datos de menor importancia para nuestros actuales propósitos, podemos hallar el número de proyectos y el tipo de participación mediante el cual los investigadores difunden sus resultados científicos, es decir, el número de comunicaciones, artículos, libros, capítulos de libros e incluso si fueron de ámbito nacional o internacional. En segundo lugar, hemos realizado una búsqueda en el Science

Citation Index (SCI) para conocer todos los artículos producidos por autores gaditanos registrados en las revistas catalogadas en esa base de datos. De ellos hemos recopilado el mayor número de datos posible, como por ejemplo, el tipo de documento, la red de investigadores con que colaboraron los investigadores en el artículo, el idioma y la revista donde fue escrito y, por supuesto, el número de citas recibidas en cada artículo y el factor de impacto de las revistas donde fueron publicados.

Con toda esta información estamos preparados para realizar una validación de los instrumentos de medida tal y como suele utilizarse para conocer la calidad científica de un sistema del conocimiento (o, de un investigador). De este modo contrastaremos la calidad de las fuentes y conoceremos el modo en que una y otra reflejan la calidad de la producción científica. Sólo nos queda añadir que la utilización de una doble vía de obtención de datos es fundamental para lograr profundizar en los aspectos metodológicos en los que se basa nuestro análisis.

4.2. EVIDENCIAS EMPÍRICAS Y ANÁLISIS DE LA REALIDAD SOCIAL

La producción científica ha experimentado un crecimiento exponencial en la mayoría de los países desarrollados debido a la fuerte competencia surgida entre las instituciones para obtener más financiación y mayores cotas de prestigio (Whitley, 2003:1023). España también ha evolucionado en las últimas décadas de la misma manera e incluso con mayor intensidad si cabe, como lo demuestran las líneas ascendentes que reflejan el incremento del número de artículos publicados por científicos españoles desde la década de los ochenta y los noventa (Moya Aneón *et al*, 2004; Jiménez-Contreras *et al*, 2003). Jiménez-Contreras, *et al*, (2003) aseguran que el principal factor que ha contribuido a la promoción de este cambio cuantitativo ha sido el proceso de normativización de la ciencia en España. Según estos autores, la creación de la Comisión Nacional de Evaluación de la Actividad Investigadora (CNEAI) estimuló positivamente la competencia creativa entre los investigadores, impulsando los índices de publicación. La influencia de los cambios normativos en la generación de nuevas pautas de comportamiento en la esfera científica es un tema clásico en la

literatura de la sociología de la ciencia y del conocimiento (Merton, 1977). Algunos autores, como Fernández Esquinas (2002) llaman nuestra atención sobre el hecho de que actualmente, estamos asistiendo a un momento crucial en la transformación del mundo de la ciencia debido a la reestructuración del mercado laboral de los profesionales de la investigación. En este trabajo también comprobaremos cómo los imperativos normativos y estructurales del contexto científico español están afectando incluso a las culturas tradicionales de las distintas especialidades científicas y a los resultados medibles de su producción. Un ejemplo de ello es que, si bien es verdad que los científicos muestran un comportamiento desigual en cuanto a la manera en que difunden los resultados de sus trabajos, prefiriendo por ejemplo, unos canales a otros, esas pautas están convergiendo en todas las áreas hacia la publicación de un mayor número de artículos (en vez de la utilización de otros recursos de difusión) y también empiezan a preferir revistas científicas especializadas de ámbito internacional a pesar de que antes no tuvieran tanta relevancia en algunas ramas del conocimiento.

Ninguno de estos procesos es ajeno a la actividad de los investigadores gaditanos que muestran un comportamiento muy similar al resto de los científicos españoles. El crecimiento de su producción científica se acelera a partir de la década de los noventa, ajustándose cada vez más a los cánones preferidos por los científicos pertenecientes a las áreas de ciencias naturales, las ramas técnicas y experimentales. Los datos procedentes del PAI reflejan que los investigadores prefieren participar en congresos, seminarios y otros tipos de eventos científicos a otro tipo de contribuciones escritas, tal y como sucede en general en el mundo de la ciencia donde abunda en primer lugar, la literatura gris base de la discusión previa de las futuras publicaciones científicas en revistas reconocidas en el área. En segundo lugar, eligen las publicaciones en revistas científicas, los capítulos de libros y, en último lugar, los libros completos. El análisis general de esta cuestión nos permite deducir que la publicación de artículos es cada vez más un recurso prestigioso para los científicos en general, impregnando incluso a ramas del conocimiento, como las humanidades, que hasta ahora lo habían utilizado en menor medida. El examen a lo largo del tiempo, nos indica que la partici-

pación a través de artículos está siendo más frecuente que en el pasado y la publicación de libros, que antes había sido el objetivo principal de ciertas áreas, está siendo combinado con este otro recurso, antes nada habitual.

Los datos procedentes del SCI muestran que los autores gaditanos han publicado casi dos mil artículos en veinte años, caracterizados por estar publicados en revistas europeas y, especialmente, en inglés. Esta información nos proporciona una imagen del modelo de ciencia en el que se inspiran los investigadores, en relación, por ejemplo, a su interés por dar a conocer sus hallazgos y al auditorio al que se dirige principalmente la difusión de sus trabajos. Algunos de los datos más significativos son que el 59,9% publica en revistas europeas, el 26,1% en revistas estadounidenses, el 13% en revistas nacionales y el porcentaje restante, en revistas procedentes de otros lugares del mundo. El peso de estas cifras no descansa únicamente en la elección realizada por los investigadores, pues hay que tener en cuenta el sesgo de infravaloración geográfica contenido en el registro de las revistas catalogadas en el SCI. Por tanto, debemos concluir que si bien estos datos reflejan las pautas de conducta de los investigadores, también confirman la menor representación de las revistas europeas y españolas en este catálogo de revistas.

A través del SCI también podemos verificar la existencia de otro de los sesgos, el relacionado con la falta de representación de las materias pertenecientes a algunas grandes áreas del conocimiento, como las ciencias sociales y humanas.

Los departamentos con más contribuciones científicas son los de química, física, medicina, biología y ciencias del mar lo que demuestra la gran actividad de los científicos pertenecientes a estas ramas del conocimiento dentro de nuestra institución de referencia, verificado mediante las dos fuente de datos. Según el SCI a estas especialidades, le seguirían en orden de importancia y el número de publicaciones, las ramas técnicas y las de informática, navales e ingeniería industrial. Sin embargo, las áreas de humanidades, filología, historia o filosofía, algunas de ellas muy productivas según el PAI, apenas aparecen en el SCI a pesar de haber escrutado las bases de datos centradas en estas ramas del conocimiento (Social Sciences Citation Index, Arts & Humanities Citation Index). Debemos achacarlo, en primer lugar, a la escasa presencia de estas áreas de conocimiento en el SCI, de modo que se confirma las críticas vertidas sobre esta base de datos. Pero de ningún modo podemos olvidar que también se debe a la preferencia de los científicos de estas especialidades, por publicar capítulos, libros y artículos publicados en revistas nacionales, recursos no incluidos en dicha base de datos, pero sí en la base de datos del PAI y que verifica nuestra hipótesis.

La divergencia entre estas dos fuentes de información alternativas es tan importante que hemos elaborado dos indicadores para conocer cuál es el grado de divergencia existente entre ellas. Para su elaboración hemos tenido en cuenta el número de investigadores activos en cada área del conocimiento.

Ilustración 1. Comparación de las tasas de productividad científica

	Tasa productividad PAI	Tasa productividad SCI
Agroalimentación	6,1	4,86
Ciencias de la Salud	8,39	3,73
Ciencias de la Vida	4,83	3,35
Física, Química y Matemáticas	7,25	7,19
Humanidades	8,12	1,83
Recursos Naturales	8,82	4,81
Economía, Sociales y Jurídicas	5,44	1,33
Tecnología de la Producción	5,92	5,16
Tecnología de la Información	5,67	1,4

Fuente: *Elaboración propia a partir del PAI y del SCI.*

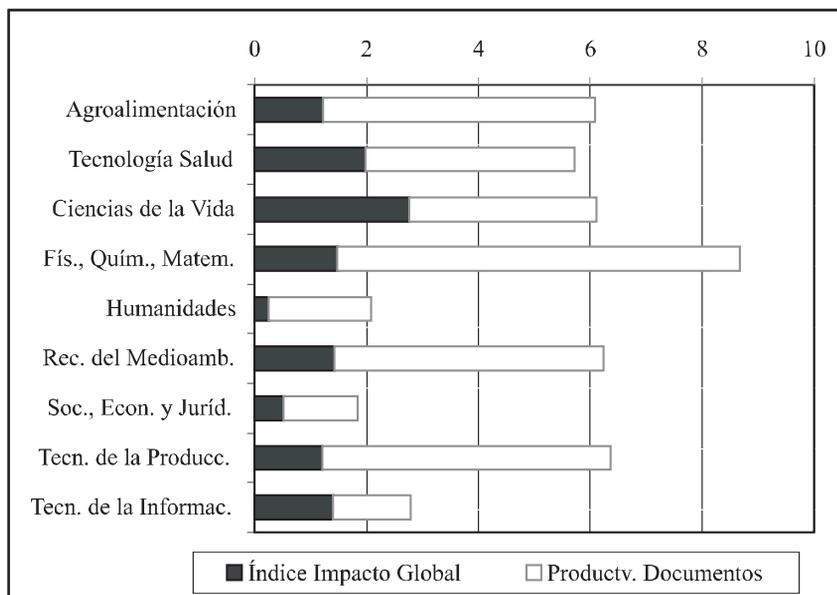
En el cuadro anterior podemos observar que únicamente las áreas de física, química y matemáticas y la de tecnologías de la producción (aunque en menor medida) presentan tasas de producción científica similares en cualquiera de las dos bases de datos. Esto quiere decir que prácticamente la totalidad de las contribuciones realizadas en estas áreas del conocimiento son publicadas en revistas inscritas en el SCI. La lejanía entre las tasas de producción científica según una base de datos y otra, en el resto de las áreas de conocimiento proceden del hecho de que publican en revistas que no están incluidas en el SCI. Ahora bien, ¿a qué se debe este comportamiento? Es posible que los primeros estén más interesados en publicar solamente en revistas de reconocido prestigio dentro de su área o, lo que es lo mismo, catalogadas en el SCI. O, bien, que sus canales de publicación se sitúan fuera de esa base de datos, lo cual nos remitiría de nuevo al sesgo de representación de ciertas áreas de conocimiento y ciertas especialidades en el catálogo de revistas del SCI. Esta tesis estaría apoyada por el hecho de que las ramas de humanidades y ciencias económicas, sociales y jurídicas son las que presentan una brecha mayor entre sendas tasas de productividad. Sin embargo, no podemos menospreciar la argumentación referida a la existencia de pautas culturales distintas entre las áreas de conocimiento, que demuestran una política de publicaciones diferente. Unas especialidades parecen más inclinadas que otras a publicar en otros idiomas, con colaboradores que no sean sus compañeros de universidad, a otorgar más responsabilidad a los becarios en su labor científica o a estar más interesados en editar artículos en revistas que poseen un mayor factor de impacto (González Ramos, 2004), por tanto, también es plausible que las ramas del conocimiento tengan una postura distinta frente al SCI, dotándolo de una mayor o menor significación simbólica.

El número de citas recibidas por un autor contribuye a aumentar su visibilidad entre colegas de disciplina. Respecto a los investigadores gaditanos, el 64,8% de sus documentos han sido citados al menos una vez durante estos veinte años, lo que supone una cifra bastante positiva para esta institución, pues se estima que al menos la mitad de los documentos producidos no son citados en ninguna ocasión (Maltrás, 1996). Además la mitad de esas citas proviene de autores extranjeros, un tercio son autocitas,

un 6% proviene de otros investigadores nacionales, casi un 3% procede de redes compuestas por autores nacionales y extranjeros y el resto, de procedencia anónima (normalmente incluidas en editoriales de las revistas). Esta elevada tasa de citas procedente de autores extranjeros es un dato muy positivo para los investigadores gaditanos porque supone que su trabajo es leído y considerado internacionalmente. Desde esa perspectiva ¿quién no querría utilizar los canales más extendidos de las revistas extranjeras, que suponen una mayor difusión de sus ideas y más visibilidad entre sus compañeros de disciplina? Muy probablemente este sea uno de los elementos más atractivos de la base de datos del SCI: que de una manera rápida proporciona un indicador de la importancia relativa del autor que estamos evaluando (o de esa institución o esa área geográfica).

Hasta el momento hemos encontrado razones de muy distinta índole, incluso problemas metodológicos de las fuentes de información; pasemos ahora a otros aspectos que nos ayudarán a desentrañar la trama de nuestro problema. Aspectos que no pueden asociarse a estos problemas técnicos porque se realizan sobre la misma base de datos, el SCI, y no puede haber contradicciones entre sus resultados. En nuestro estudio nos hemos preguntado si la tasa de producción de documentos, es decir, el número de artículos totales publicados, y la tasa de publicación de artículos en revistas con mayor índice de impacto coincidirían. Es decir, si el mismo conjunto de unidades observadas serían las que obtendrían las mayores tasas de publicación de artículos y publicación en revistas con mayor índice de impacto. Sin embargo, hemos encontrado que esto no es así. Tanto un análisis pormenorizado de los autores individualmente, como un análisis global por equipos de investigación y por áreas de conocimiento, permite comprobar que es muy difícil que coincidan ambas tasas en el mismo grupo. Cuando un índice señalaba hacia un sentido, el otro señalaba hacia otro, de modo que fue complicado encontrar que los dos índices se concentraran en el mismo conjunto de individuos o áreas. En el siguiente gráfico podemos ver el resultado de la aplicación de estos dos indicadores según su pertenencia a las grandes ramas del conocimiento.

Ilustración 2. Comparación entre la tasa de productividad y el índice de impacto global de las revistas donde fueron publicadas



Fuente: *Elaboración propia a partir del Science Citation Index*

La tasa de productividad es más elevada que el índice del impacto global de las revistas en casi todos los casos. Esto supone que, por ejemplo, en la rama de física, química y matemáticas, donde esa diferencia es la mayor, los investigadores a pesar de contar con numerosos artículos publicados, no lo hacen en aquellas revistas que poseen mayor índice de impacto. El área de tecnología de la información y el de ciencias de la salud, aunque ésta en menor medida, presenta cierto equilibrio entre ambos indicadores, lo que significa que su producción científica está más o menos equiparada con la elección de revistas, que dentro de su área poseen un mayor índice de impacto. En la rama de ciencias de la vida el fenómeno es inverso, es decir, el índice de impacto global supera la tasa de producción de documentos, lo que supone que el esfuerzo de los científicos se centra sobre todo en publicar en revistas que posean niveles de alto impacto. El examen desagregado por grupos de investigación e incluso por individuos indica lo mismo: aquellos grupos y científicos con más publicaciones no mantienen esa misma preponderancia cuando se examina el factor de impacto de las revistas donde publican sus contribuciones científicas y viceversa. Por

ejemplo, los autores más prolíficos, es decir con mayor tasa de publicaciones, no son los mismos que publican en revistas con un elevado factor de impacto. Ocasionalmente, es posible encontrar un cierto grupo, que en mayor o menor medida, concentren buenos resultados respecto a los dos indicadores, pero no podemos decir que esto sea un hecho generalizable. Por tanto, cuando decimos que los mejores investigadores son éstos o aquellos, en ningún caso podemos decir que lo son porque publiquen mucho y, al mismo tiempo, lo hagan en revistas con un elevado factor de impacto.

Parece que la concentración de méritos no favorece siempre, en el mismo sentido, a todos los investigadores ni a las mismas ramas científicas. Esto está relacionado con la diferencia entre especialidades científicas que establece distintas pautas de comportamiento, expectativas y posibilidades reales de los investigadores y de sus unidades de investigación a la hora de conseguir logros científicos dentro de su campo del conocimiento. Cada cultura científica influye en la toma de decisiones relativa al hecho de publicar mucho o de hacerlo en revistas prestigiosas, es decir, influye en la orientación del trabajo científico de los investigadores. Afectaría

incluso al cálculo de los límites de competitividad que los investigadores harían sobre su trabajo dentro del área: Siendo improbable conseguir a la vez, publicar una gran cantidad de artículos y hacerlo en revistas con un elevado nivel de impacto es necesario decidir *a priori* qué posibilidades de competitividad posee su trabajo y actuar en consecuencia. O bien, con un estilo de trabajo más proclive a la proliferación de resultados o, bien, de selección de las revistas donde se va a publicar los hallazgos científicos.

5. CONCLUSIÓN

Los indicadores que miden el reconocimiento del trabajo científico, tanto si se hace a través del número de artículos o mediante el número de citas y el factor de impacto acumulado por las revistas donde publican sus resultados, presentan comportamientos muy diferentes dependiendo de las áreas de conocimiento y no siempre apuntan al mismo colectivo cuando se atiende a su «excelencia». También hemos advertido una discrepancia al aplicar distintas fuentes de datos que miden los índices de productividad científica. Por tanto, debemos admitir, que ninguno de esos indicadores, utilizados por separado, es suficientemente válido, que aporte una medida certera y precisa de la calidad científica. Ello se debe a diversas razones. Algunas de carácter estrictamente metodológicas, pues hemos comprobado que a veces son debidas a desajustes en la manera de medir el fenómeno y que, por tanto, es un problema de fiabilidad de las fuentes. Otras pertenecen al plano de la socialización y la cultura de los científicos, que influyen en los valores y en las pautas de comportamiento de los investigadores. Así, las elecciones realizadas por los investigadores a la hora de difundir sus resultados científicos al resto de la comunidad científica, responden a una serie de consideraciones construidas a partir de la cultura predominante en esa especialidad del conocimiento (González Ramos, 2004; Campbell, 2003; Bellavista *et al*, 1998). En tercer lugar, es importante destacar que algunas medidas no son capaces de discriminar los resultados con suficiente precisión, debido a la insuficiencia de los recursos existentes para recopilar los datos de los que procede el análisis. Así, algunas ramas del conocimiento tienen métodos de recogida de datos más

eficientes que otras. Sin embargo se usa el mismo instrumento de medida en todas las áreas de conocimiento. El problema radica en que, aunque la construcción de la medida sea objetiva, ésta no puede ser válida para todos los casos si no es sensible a las diferencias estructurales de la comunidad científica que da lugar a una práctica científica diferente. Es preciso crear un indicador que sea válido, pero también que sea capaz de apreciar esas diferencias, de otro modo, estaríamos midiendo las horas con un reloj de arena. La medida ha de permitirnos contrarrestar el resultado del esfuerzo realizado por cada institución o área de conocimiento para conseguir desempeñar su trabajo, minimizando así el indeseable efecto Mateo en la ciencia.

A lo largo de estas páginas hemos comprobado las ventajas de utilizar un método objetivo de medición de la calidad de la actividad investigadora, pero también los problemas que conlleva utilizarlos sin tener en cuenta su trasfondo. Hemos demostrado que su utilización independiente supone ignorar que sólo estamos obteniendo un reflejo de la imagen general que debería obtenerse al aplicar el instrumento de medición. El problema radica en que tratamos de sintetizar en una sola medida un fenómeno complejo, lleno de matices y que se refiere a múltiples niveles de comparación. La actividad científica se establece en un sistema de relaciones de cooperación y de competencia, de acumulación de méritos de muy distinto tipo y, por si fuera poco, dotado de un sistema de recompensas que no funciona homogéneamente. El panorama se hace más confuso cuando se emplean varios instrumentos de medición, que ofrecen una imagen distinta entre sí. Olvidamos que la validez de un buen indicador se sostiene sobre su contingencia con la realidad y la fiabilidad de su medición (es decir, en que la repetición de la medida ofrezca idénticos resultados), pero esto no sucede así porque no se está teniendo en consideración la complejidad de la medida.

El indicador debe estar construido sólidamente, es decir, su fundamentación debe basarse en una buena acotación de los elementos intrínsecos al fenómeno que se trata de medir, siendo capaz de mostrar las matices que reside en el concepto. Nadie diría que el nivel de salud de una población se mide por el número de médicos por habitantes, este es sólo uno de los indicadores que debe componer el sistema de indicadores que nos ayudará a realizar una medida correcta sobre la

realidad. Se ha de construir un indicador tan complejo como lo sea el fenómeno analizado, por lo que no vemos otra manera de conseguirlo sino estableciendo un índice sintético, construido a partir de los elementos significativos del contenido temático al que se refiere el concepto de calidad de la producción científica. El instrumento ha de ser potente y sensible a la vez, de modo que sea capaz de explicar el fenómeno desde un punto de vista comparativo global, pero también que refleje las diferencias más sutiles halladas entre las distintas especialidades del conocimiento o niveles de sistema del conocimiento. Debe servir tanto para medir el nivel de producción científica que se desprende de toda una institución como para valorar el esfuerzo realizado por los equipos de investigación.

Las dimensiones en base a las cuales se ha de construir el instrumento de medida, han de abastecerse de los canales de información existentes pero, por eso mismo, debemos cuidar el estado de la materia prima. Las fuentes de datos deben ser de calidad, tan extensas en unas áreas del conocimiento como en otras y deben abarcar un área geográfica adecuada, de otro modo, sólo conseguiremos potenciar el trabajo que se realiza en ciertas partes del sistema de ciencia. En ese extremo, consideramos que lo más oportuno es que las instituciones de gestión del conocimiento organizadas a partir de sistemas regionales del conocimiento sean las encargadas de diseñar ese sistema de recogida de datos y de vigilar la producción de la calidad de los datos recopilados. En cuanto a los conceptos que han de integrarlo merece una reflexión más detenida pero podemos adelantar que debe centrarse en el nivel de proyectos conseguidos, el número de contribuciones científicas realizadas, número de

citadas recibidas que respondan a la necesidad de reconocer el mérito de los autores más innovadores u originales dentro de su área y, como no, a la importancia y la rigurosidad de las revistas científicas, y las editoriales donde se publican sus trabajos. Aún habría que añadir, en nuestra opinión, otro aspecto que hasta ahora no ha sido incluido en ningún estudio de evaluación de la calidad de la producción científica y que, sin embargo, está directamente relacionado con la calidad del trabajo realizado. Nos referimos a las condiciones en que se realiza el trabajo científico y el uso correcto de los métodos de investigación desarrollado por los investigadores para realizar su propia investigación. Esta dimensión debería ser capaz de medir si los esfuerzos de los investigadores se han dirigido a minimizar los efectos de los sesgos del diseño empleado y han hecho todo lo posible por extraer el mejor partido a sus datos, asegurándose de que ha tenido en cuenta todos los controles de seguridad que su investigación precisaba (King, *et al*, 2000).

Conjugando todos estos elementos, quizá sea posible encontrar una medida que convenga a un mayor número de científicos basándose en la idoneidad de la medida y así se acepte de mejor grado ser evaluados, un elemento directamente relacionado con el progreso científico. El sometimiento a un proceso de evaluación claro y objetivo es una reclamación legítima que nos ayudará a descubrir los aspectos más débiles del sistema y a considerar distintas formas de superar los problemas que interfieren en el desarrollo de una producción científica de calidad. Además, un buen procedimiento de evaluación puede proporcionar la ocasión de premiar el buen trabajo realizado por los actores del sistema del conocimiento.

6. BIBLIOGRAFÍA

- AGENCIA NACIONAL DE EVALUACIÓN DE LA CALIDAD Y ACREDITACIÓN (2004): *Los españoles y la universidad. Primera encuesta nacional sobre la imagen pública del sistema universitario español*. Madrid, ANECA.
- BECHER, T. (1989): *Tribus y territorios académicos. La indagación intelectual y las culturas de las disciplinas*. Barcelona, Gedisa.
- BELLAVISTA, J., GUARDIOLA, E., MÉNDEZ, A. y BORDONS, M. (1997). *Evaluación de la investigación*. Madrid, Centro de Investigaciones Sociológicas.
- BELLAVISTA, J., TURPIN, T., HILL, S. y DE MIGUEL, J. M. (1998): «Cultura organizativa de investigadores y entorno político y social» *Papers*, 54: 79-109.
- BUELA-CASAL, G. (2002): «Evaluación de la investigación científica: el criterio de la mayoría, el factor de impacto, el factor prestigio y los diez mandamientos para incrementar las citas» *Análisis y Modificación de la Conducta*, 28 (119): 455-475.
- BRICALL, J. M. (2000): *Universidad 2mil*. Madrid, Consejo de Rectores de las Universidades Españolas, CRUE.

- CAMPBELL, R. A. (2003): «Preparing the new generation of scientist: Social process of managing student» *Social Studies of Science*, vol. 33, n. 6, 897-927.
- DE MIGUEL, J. *et al* (1999): «Modelo de indicadores de las universidades» *Gestión y Análisis de Políticas Públicas*, 16: 29-62.
- DÍAZ DE RADA, V. (2001): *Organización y gestión de los trabajos de campo con encuestas personales y telefónicas*. Barcelona, Ariel.
- EHEVARRÍA, J. (2004). «El *ethos* de la ciencia, a partir de Merton» *Sociología de la Ciencia*, VALERO, J. A. (Coord.). Madrid, EDAF.
- FERNÁNDEZ ESQUINAS, M. (2002): «Mercados de trabajo en la ciencia. Balance de la investigación y propuesta de marco analítico» *Revista Internacional de Sociología*, nº 32: 35-75.
- GARCÍA, C. E. y SANZ, L. (2005): «Competition for funding as an indicator of research competitiveness» *Scientometrics*, 64, 3: 271-300.
- GODIN, B. (2003): «The emergente of S&T indicators: Why did governments supplement statistics with indicators?» *Research Policy*, nº 32: 679-691.
- GONZÁLEZ, A. M. (2004): *Modos de producción científica: Cultura y metodologías de investigación en la Universidad de Cádiz*. Universidad de Cádiz, Tesis doctoral.
- HACKING, I. (1998): *¿La Construcción social de qué?* Madrid, Paidós.
- INSTITUTO NACIONAL DE ESTADÍSTICA (2000): *Estadística sobre las actividades en Investigación Científica y Desarrollo Tecnológico (I+D). Indicadores básicos 2000*. Madrid, INE.
- KING, D. A. (2004): «The scientific impact of nations» *Nature*, 430:311-316.
- KING, G. *et al* (2000): *El diseño de la investigación social. La inferencia científica en los estudios cualitativos*. Madrid, Alianza.
- JIMÉNEZ-CONTRERAS, E., MOYA, F., DELGADO, E. (2003): «The evolution research activity in Spain. The impact of the National Comission for the Evaluation of Research Activity (CNEAI)» *Research Policy*, 32: 123-142.
- LLERA, F. J. y RETORTILLO, A. (2004): *Los españoles y la universidad. Primera encuesta nacional sobre la imagen pública del sistema universitario español*. Agencia Nacional de Evaluación de la Calidad y la Acreditación.
- MALTRÁS, B. (1996): *Los indicadores bibliométricos en el estudio de la ciencia. Fundamentos conceptuales y aplicación en política científica*. Universidad de Salamanca, Tesis doctoral.
- MANUAL DE FRASCATI, DE OSLO, CANBERRA Y PATENTES (2002): <http://www.connicyt.cl/indicadores/ponencias/QUINTERO.HTM> (8.1.2002).
- MERTON, R. K. (1973): *The Sociology of Science. Theoretical and empirical investigations*. Chicago, University of Chicago Press.
- MERTON, R. K. (1984): *Ciencia y tecnología en la Inglaterra del siglo XVII*. Madrid, Alianza.
- MOYA-ANEGÓN, F. y SOLÍS CABRERA, F. M., coord. (2004): *Indicadores Científicos de Andalucía (ISI, Web of Science. 1998-2001)*. Granada, Programa de Divulgación Científica de Andalucía, Parque de las Ciencias, Secretaría General de Universidades, Consejería de Educación y Ciencia, Junta de Andalucía.
- ORTEGA Y GASSET, J. (1930): *Misión de la Universidad*. Madrid, Alianza.
- ROUSSEAU, R. (2002): «Journal evaluation: Technical and practical issues» en *Library Trends*, vol. 50, n. 3, pp. 418-439.
- RUBIO LINIERS, M. C. (2001): «Bibliometría y Ciencias Sociales. Proyecto Clío» en <http://clio.rediris.es/clionet/articulos/bibliometria.htm> (03.10.01).
- STORER, N. W. (1976): «The hard sciences and the soft: Some sociological observations» en *Bulletin of the Medical Library Association*, 55 (1): 75-84.
- VALERO, J. A. (2004): «Sociología de la ciencia: un análisis posmertoniano» *Sociología de la Ciencia*, VALERO, J. A. (Coord.). Madrid, EDAF.
- WHITLEY, R. (2003): «Competition and pluralism in the public science: the impact of institutional frameworks on the organisation of academic science» *Research Policy*, nº 32: 1015-1029.