

UNIVERSIDAD DE SALAMANCA

FACULTAD DE FILOSOFÍA

DEPARTAMENTO DE FILOSOFÍA, LÓGICA Y FILOSOFÍA DE LA CIENCIA



Los indicadores bibliométricos en el estudio de la ciencia  
Fundamentos conceptuales y aplicación en política científica.

BRUNO MALTRÁS BARBA

*Tesis Doctoral*

*Salamanca, 1996*



UNIVERSIDAD DE SALAMANCA

FACULTAD DE FILOSOFÍA

DEPARTAMENTO DE FILOSOFÍA, LÓGICA Y FILOSOFÍA DE LA CIENCIA



Los indicadores bibliométricos en el estudio de la ciencia  
Fundamentos conceptuales y aplicación en política científica.

BRUNO MALTRÁS BARBA

*Tesis Doctoral*  
*Director: Miguel Ángel Quintanilla Fisac*  
*Salamanca, 1996*

*Los indicadores bibliométricos en el estudio de la ciencia.  
Fundamentos conceptuales y aplicación en política científica*

Tesis presentada D. Bruno Maltrás Barba  
y dirigida por Dr. D. Miguel A. Quintanilla Fisac  
para optar al grado de Doctor en Filosofía

Vº Bº

Salamanca, 3 diciembre 1996

A la memoria de mi hermano, Guille.

# Contenidos

<b>Prefacio</b> .....	XI
-----------------------	----

## **PARTE I. EL SISTEMA DE PUBLICACIÓN DE LA CIENCIA**

<b>Capítulo 1. Bases del funcionamiento del sistema de publicación</b> .	7
1.1. El surgimiento del sistema de publicación y su papel en la constitución de la ciencia moderna .....	7
1.2. Prioridad y reconocimiento .....	10
1.3. Los sentidos del reconocimiento .....	22
1.4. El sistema de publicación como archivo de la ciencia .....	26
1.5. El sistema de revisión por pares .....	30
1.5.1. La fiabilidad del sistema de revisión por pares .....	34
1.5.2. Algunas consecuencias del sistema de revisión por pares ....	39
<b>Capítulo 2. La producción de resultados científicos</b> .....	45
2.1. El ciclo de producción de conocimientos .....	45
2.2. Las actividades científicas .....	53
2.3. Los resultados científicos .....	56
2.4. Los agentes científicos .....	62
2.5. Los criterios de agregación .....	64
2.5.1. Temático .....	64
2.5.2. Institucional .....	69
2.5.3. Geográfico .....	74
<b>Capítulo 3. Elementos del sistema de publicación</b> .....	77
3.1. Las revistas científicas .....	77
3.2. La estructura del documento científico .....	86
3.3. Las referencias bibliográficas en el análisis bibliométrico .....	99

## PARTE II. LOS INDICADORES BIBLIOMÉTRICOS

Capítulo 4. Indicadores de producción .....	109
4.1. Los recuentos de publicaciones .....	112
4.1.1. El significado de los recuentos .....	113
4.1.2. La atribución de los resultados .....	120
4.1.3. Métodos de fraccionamiento .....	123
4.1.4. La métrica de los recuentos .....	129
4.1.4.1. La estructura de la productividad científica .....	131
4.1.4.2. Las variaciones entre disciplinas .....	144
4.1.5. La cobertura de las bases de datos .....	149
4.2. El concepto de producción científica .....	151
4.3. Indicadores bibliométricos de la producción científica y sus aplicaciones. ....	152
 Capítulo 5. El problema acerca del concepto de calidad científica .	161
5.1. La calidad científica y los indicadores bibliométricos .....	161
5.2. La noción de "calidad" en los estudios cuantitativos de la ciencia	164
5.2.1. Los pioneros: calidad como prestigio .....	166
5.2.2. Esfuerzos de precisión conceptual .....	168
5.2.3. La "salud" de la ciencia .....	177
5.2.4. Dos enfoques: indefinición versus concreción. ....	180
5.2.5. Problemas de evaluación de la ciencia .....	183
5.2.6. ¿Y si preguntamos a los científicos? .....	192
5.2.7. Conclusión .....	202
 Capítulo 6. Los indicadores bibliométricos y la calidad .....	203
6.1. Síntesis de los polos básicos en el debate sobre la calidad .....	203
6.1.1. El objeto de la calidad. ....	206
6.1.2. El nivel de agregación .....	207
6.1.3. El carácter objetivo de la calidad científica .....	208

6.1.4. La composición de la calidad .....	209
6.1.5. La relación entre calidad y éxito científicos .....	211
6.1.6. Grado de adecuación al concepto interno de calidad científica .....	214
6.2. Una calidad medible por los indicadores .....	219
6.3. El factor de impacto (FI) y los indicadores derivados .....	222
6.4. Los indicadores P-10 y %SUP como estimadores de la calidad .	234
Capítulo 7. Los indicadores bibliométricos de colaboración .....	239
7.1. La colaboración en la ciencia .....	240
7.2. La medida de la colaboración .....	244
7.2.1. Bases conceptuales y alternativas metodológicas en la cuantificación de la colaboración .....	248
7.2.2. Comentarios sobre los métodos de recuento de la colaboración .....	259
<b>Conclusiones</b> .....	265
<b>Bibliografía</b> .....	269

Los indicadores bibliométricos en el estudio de la ciencia  
Fundamentos conceptuales y aplicación en política científica.

## *Prefacio*

---

La ciencia es el fenómeno cultural y social más importante de los últimos cuatro siglos. No puede, por ejemplo, entenderse nada de la Filosofía Moderna sin atender al impacto que causaron en los filósofos las nuevas formas de conocer desarrolladas por las ciencias clásicas. El descubrimiento de la asimetría entre el conocimiento intuitivo y ordinario, de una parte, y el método (o métodos) científicos, de otra, constituyen el hilo conductor de la Filosofía Moderna desde finales del Renacimiento hasta las últimas crisis postkuhniana.

En el siglo XX la ciencia es, seguramente, la institución social más dinamizadora: sus efectos transforman desde la raíz las intervenciones tecnológicas y la propia estructura de las sociedades contemporáneas. A través de los mecanismos de división social del trabajo cognitivo se han generado instituciones científicas que han modificado completamente el panorama de la educación, la sanidad, la economía o el aparato del Estado. Esta importancia social y la propia riqueza de su estructura interna explican que, a partir de los años cincuenta y sesenta, fueran innumerables los autores que comenzaron a estudiar seriamente la dinámica de la ciencia desde el punto de vista de una institución social productora de conocimiento. En estas tres últimas décadas se ha observado una explosión de estudios sociales sobre el funcionamiento e impacto de las actividades científicas. Dentro de éstos, los estudios cuantitativos han surgido como un nuevo complejo disciplinario teórico y práctico, en cuyo seno se han desarrollado nuevos métodos de aproximación empírica a la ciencia basados en baterías de indicadores cuantitativos acerca de su institucionalización, producción y calidad.

Ya no es necesario subrayar y justificar la importancia de estos indicadores, puesto que han entrado a formar parte de las propias instituciones científicas, de sus sistemas de recompensa y de los filtros de control de calidad. Por otra parte, se han convertido en un instrumento básico mediante el que la sociedad, a través de distintos mecanismos, puede acceder a un conocimiento sobre la estructura real de la ciencia y, en su caso, orientar en lo posible y en lo deseable la dirección del desarrollo del conocimiento. Para diseñar las políticas científicas se emplean ya cotidianamente los diversos métodos de medición cuantitativa con vistas a la evaluación de los sistemas de investigación. Más que justificar su necesidad, lo que queda pendiente es una legitimación teórica y filosófica.

Esta tesis partió de una preocupación por no separar dos tradiciones en los estudios sobre la ciencia: la primera, filosófica, se había preocupado por la construcción y desarrollo de modelos "micro" para la dinámica del conocimiento; la segunda, sociológica, había concentrado sus esfuerzos en el estudio de grandes agregados, en forma de instituciones y de sistemas observables. Esta preocupación tenía, igualmente, una doble dimensión: por un lado, la de establecer un puente entre los resultados observables y su significado epistémico; por otro lado, tras esta preocupación estrictamente filosófica, subyacía un interés de fundamentación teórica de los métodos cuantitativos centrado en cuestiones como las siguientes: la taxonomía básica de indicadores, la delimitación de las magnitudes que miden estos indicadores, la relación entre las cantidades medibles y valores epistémicos inobservables, el estudio de las prácticas, hábitos, patrones de conducta y microinstituciones que sustentan la interpretabilidad de los indicadores cuantitativos sobre la ciencia, etc.

Las cuestiones que han guiado este trabajo fueron emergiendo a partir de la práctica y el uso de estos indicadores en el estudio del sistema científico español y del europeo. A pesar de su extensivo uso y de la aparente simplicidad de muchas de estas medidas, fue haciéndose para mí cada vez más ineludible la preocupación por el sustrato teórico en el que se basaban. El sentido común y una vaga referencia a los logros de algunas escuelas sociológicas parecían ser las únicas respuestas cuando se asumía una postura crítica respecto a la validez de estas medidas. Sin este sustrato teórico, parecía extraerse la impresión de que no era posible tender un puente entre la dimensión puramente descriptiva y la

necesidad normativa que exigía su uso interno dentro de las comunidades científicas como control de calidad y su uso social en las políticas científicas como instrumento de evaluación.

El centro de atención en este trabajo se reduce a los indicadores bibliométricos, entendidos como uno de los subconjuntos de las medidas cuantitativas sobre la ciencia. La bibliometría es, así, una parte de la cienciometría. *Los indicadores bibliométricos son medidas obtenidas a partir del análisis estadístico de los rasgos cuantificables de la literatura científica.* El problema que se aborda en el presente trabajo es el siguiente: ¿qué significan los indicadores bibliométricos? Conocer el fundamento y las posibilidades de la interpretación de esos indicadores es algo necesario si se aspira a hacer un buen uso de ellos.

La primera parte está dedicada al análisis del sistema de publicación de la ciencia, como lugar en el que se establece de manera prioritaria el sistema de reconocimientos internos a partir de la valoración de los resultados científicos presentados materialmente en forma de documentos. Las pautas y principios sobre las que se sostiene este sistema y las consecuencias de su funcionamiento ofrecen una base sólida para asentar la comprensión general de los indicadores bibliométricos. La reflexión sobre ese funcionamiento permitirá una caracterización más precisa del papel del documento científico, materia prima del análisis bibliométrico. La noción de *resultado científico* es la clave que permite integrar los aspectos cognitivos y sociales de la ciencia; establece el puente entre los objetivos y funcionamiento propios de la ciencia con las consecuencias observables por el análisis bibliométrico.

En la segunda parte se trata directamente de buscar un significado interpretable y operativo para tres tipos de indicadores bibliométricos: los indicadores de producción, calidad y colaboración. Puede tomarse como un muestrario de las vías posibles al lidiar con el problema del significado de los indicadores. El capítulo de los indicadores de producción se ocupa de encontrar una explicación coherente para el comportamiento de las medidas basadas en recuentos de publicaciones y de justificar las normalizaciones de esas medidas. Para los indicadores de calidad, se ha seguido una estrategia contraria: depurar un concepto hasta encontrar una versión acorde con las posibilidades del análisis bibliométrico. El dedicado a la colaboración se ocupa de explorar las alternativas técnicas y relacionarlas con diferentes conceptos de la colaboración.

De la misma forma en que el filósofo del derecho se preocupa por cuestiones sustanciales sobre la naturaleza de la justicia que trascienden necesariamente la expresión de las normas en el derecho positivo, hemos intentado no desprender el estudio cuantitativo de la ciencia de la relación con los valores epistémicos que nos permiten hablar de ésta como de una auténtica empresa social de producción de conocimientos, y no simplemente de meros productos culturales más o menos útiles. Si nos hemos acercado al objetivo que se ha perseguido, se habrá contribuido a establecer una fundamentación teórica de la práctica evaluativa de las instituciones científicas.

A lo largo de estos años he disfrutado del apoyo de muchas personas que me han hecho sentir muy afortunado. De distintos modos, han hecho posible que este trabajo llegara a su fin.

Miguel A. Quintanilla ha sido maestro y amigo. Es difícil encontrar a alguien de quien se puedan aprender tantas cosas y con tanta naturalidad. Conmigo ha sobrepasado los límites de la generosidad y de la paciencia. A Fernando Broncano le debo más de lo que soy capaz de expresar aquí. Siempre ha estado dispuesto a elevarme el ánimo y a valorar mis pequeños logros. Su amplitud de miras y su cultura extensiva han sido un foco de constante enriquecimiento. Jesús Vega me ha ayudado de muchas maneras. La admiración que siento por él sólo es superada por la estima que le tengo como amigo y compañero. Las visitas de Margarita Vázquez y Manolo Liz, además de un estímulo intelectual, han sido sanos recreos en esta larga tarea.

Trabajar con todos ellos y con los demás miembros del grupo EPOC me ha ofrecido un entorno en el que he desarrollado mucho más de lo que habría podido imaginar. Los problemas prácticos de evaluación de ciencia y de las políticas científicas, con los que diversos proyectos de investigación me han puesto en contacto, han proporcionado una constante referencia en los problemas teóricos que han guiado esta tesis.

El apoyo institucional ha sido determinante en muchos sentidos y quiero, además de hacer la justa mención, expresar mi gratitud. En mi formación universitaria recibí las ayudas públicas del Ministerio de Educación en forma de becas de estudios. Este trabajo es una consecuencia directa de la beca de investigación que he disfrutado durante los años

1991-94, concedida por la Secretaría General del Plan Nacional de I+D. A la Universidad de Salamanca debo el haberme proporcionado los medios y la estabilidad necesaria para concluir la investigación. Espero haber compensado en alguna medida todas estas deudas.

Raquel ha sido mi compañera de gozos y fatigas. Ha mantenido mi vuelo con sus pies en el suelo. A mi familia, además de todo le debo la confianza y el cariño incondicionales, que son un soporte muy de agradecer. Si me olvidara de mencionar a Antonio y Mercedes, ellos se enfadarían con razón. Todos mis amigos tendrán que preguntarme ahora otra cosa.

Los indicadores bibliométricos en el estudio de la ciencia.  
Fundamentos conceptuales y aplicación en política científica

## Parte I

---

# EL SISTEMA DE PUBLICACIÓN EN LA CIENCIA

De modo general, diremos que el *sistema de publicación de la ciencia* es el conjunto de elementos y pautas que sostienen, regulan y perpetúan el proceso por el que los investigadores hacen accesibles de modo "oficial" al resto de la comunidad científica sus pretensiones de contribuir al acervo científico. El fruto que resulta de la operación de este sistema es la *literatura científica*, ese agregado de documentos pertenecientes a una clase muy particular que delimita el escenario de los debates científicos y que hace posible que la ciencia sea una empresa colectiva.

Simplificando, el funcionamiento del sistema de publicación científica es el siguiente: los investigadores envían (sin por ello recibir una compensación económica) a determinadas revistas especializadas unos escritos en los que describen detalladamente sus descubrimientos, con la esperanza de que esas revistas los publiquen. A petición de los responsables de las revistas, ciertos expertos juzgan (también sin cobrar) si cada uno de esos escritos tiene o no valor científico; cuando encuentran que sí, recomendarán su publicación, lo que generalmente implica que en el plazo de unos meses ese escrito aparecerá en las páginas de un número de la revista correspondiente (a veces se solicita del autor una cuota para

sufragar gastos de la revista). Esos expertos también pueden a veces dar indicaciones a los autores para que mejoren sus trabajos antes de que éstos sean publicados. Así parece asegurarse que los trabajos finalmente publicados hayan superado un listón de calidad. Posteriormente, los resultados contenidos en el conjunto de trabajos publicados son puestos a prueba y utilizados por el resto de especialistas. Así emerge un conocimiento sólido y sistemático, que tiene además una enorme potencia generatriz de nuevos conocimientos de la misma clase.

Pero este dispositivo puede dejar perplejos a los observadores ajenos a la ciencia. No sólo por la posibilidad de que algo tan complejo haya llegado a constituirse, sino sobre todo por el hecho de que pueda funcionar correctamente, tal como se espera de él. ¿Por qué los investigadores no se guardan para sí sus descubrimientos, a los que evidentemente consideran valiosos, y en cambio los ponen desinteresadamente a disposición de la comunidad científica? ¿Puede confiarse en que el sistema de control no producirá abusos, que será neutral respecto a los investigadores y a los trabajos, esto es, podrá discernir adecuadamente si un nuevo resultado es o no valioso? Por otra parte, si se concede el beneficio de la duda y se admite que el sistema realmente funciona, parece inevitable la conclusión de que la ciencia es una organización social ciertamente peculiar, que sostiene unos valores muy distintos de los comunes en otras esferas de nuestra sociedad: abruma tanta generosidad y cooperación en un terreno que produce conocimientos comprobadamente eficaces, útiles, capaces de dotar de ventajas comparativas a unos grupos humanos respecto a otros.

Por todo lo anterior resulta interesante indagar cuáles son las fuerzas que mantienen el sistema de publicación; estudiar con detalle cuáles son los papeles que este sistema cumple en la ciencia y en qué sentido resulta central en el desarrollo de ésta; e intentar comprender hasta qué punto la posibilidad de que este sistema funcione hace que la ciencia sea diferente de otras partes de la sociedad.

En el momento en que un investigador presenta para su publicación un trabajo escrito acerca de los resultados que ha alcanzado en su tarea científica, entra en funcionamiento un preciso mecanismo en el que intervienen elementos diversos, organizados bajo un conjunto de reglas bien establecidas. Después de que se complete una serie de fases, ese investigador sabrá si su trabajo aparecerá o no finalmente publicado en los canales oficiales de la ciencia, es decir, en las revistas, colecciones, o

ediciones especializadas y con marchamo científico, que constituyen el núcleo en torno al que se articula el sistema de publicación. En esas fases intervendrán otros elementos de ese sistema, a los que antes se aludía, que son, además de los investigadores en su doble papel de productores y consumidores de publicaciones, y de las revistas o los periódicos científicos, soportes materiales de la publicación, otras entidades como las asociaciones, de las que a veces éstos últimos son órganos dependientes; sus editores y consejos editoriales, que intervienen en la definición y la aplicación de los criterios de publicación; los jueces anónimos que evalúan los manuscritos presentados para ser publicados; los artículos y otros tipos de documentos científicos, que adoptan un formato muy determinado por los requisitos que el sistema impone; los suscriptores y lectores de las revistas, cuyos intereses y expectativas pueden influir en las prácticas editoriales; e incluso las instituciones que mantienen las bibliotecas o las que emplean a los científicos, las cuales pueden influir sobre el comportamiento productivo de éstos. En suma, el sistema de publicación involucra en una acción coordinada a muchos de los elementos esenciales de la organización científica.

Entre las pautas que determinan el modo en que funciona el sistema de publicación, están las normas que definen cuáles son las formas de publicación aceptables, con sus convenciones acerca del modo de exposición o de cortesía académica; también otras que se derivan de los valores aceptados y que fundan los criterios que aplican los evaluadores, los editores y los mismos científicos al escribir sus trabajos, referidas al contenido científico que exhibe el escrito; además, existen protocolos comúnmente aceptados que regulan las funciones de todos los elementos anteriores y sus relaciones mutuas durante el proceso previo a la publicación formal de un documento científico.

Puede encontrarse muchas veces la expresión "comunicación en la ciencia" para designar este proceso; sin embargo, entender literalmente esa denominación implica una notoria parcialidad. Por una parte está el hecho obvio, frecuentemente señalado, de que existen otros canales y medios para la comunicación entre los científicos, como las comunicaciones personales (cartas, contactos directos, envío de materiales previos a la publicación), la colaboración eventual, los congresos y las reuniones científicas, entre otros, que suelen llamarse *informales* por contraposición al hecho de la *publicación oficial*. La comunicación en la ciencia tiene lugar tanto en los canales formales como en los informales. Pero

por otra parte, y especialmente, esa expresión es parcial porque además del intercambio de datos o la transmisión de información, en este proceso suceden otras cosas que son fundamentales en la determinación del funcionamiento de la ciencia como sistema organizado de producción de conocimientos.

Desde luego que en él la información fluye, y además de un modo especialmente rico: no sólo del investigador que presenta su trabajo hacia el resto de la comunidad, sino también en los mecanismos de control intermedios (interacción que puede modificar el manuscrito original), desde los receptores del trabajo en la evaluación pública hacia el autor (esta retroalimentación le ofrece nuevas bases para enjuiciar su propio trabajo) y a través de la vasta red que surge de los nexos generados por las referencias entre los trabajos, que por un lado localiza trabajos anteriores que pueden ser de utilidad y por otro articula esas aportaciones individuales, enriqueciendo su significación. Tampoco puede dejar de recordarse que, desde su surgimiento, lo que ahora llamamos "sistema de publicación de la ciencia" sustituyó en buena medida al primitivo sistema de comunicación por cartas, folletos y libros que se intercambiaban los colegas interesados en temas de investigación relacionados, asumiendo la mayor parte de las funciones de esta costumbre anterior.

No obstante, *existen elementos en el proceso de publicación científica cuya presencia sólo se justifica apelando a otras funciones distintas de la mera comunicación y que incluso serían un obstáculo si ésta fuera lo único o lo más importante*. Algunas de estas funciones, que se tratarán más extensamente junto a otras en las páginas que siguen, son el establecimiento de un *canal reivindicativo* de la prioridad en los descubrimientos que resulta esencial para que los científicos individuales obtengan el reconocimiento que necesitan; la operación de un *filtro protector* que salvaguarda la autoridad y prestigio de la ciencia al tiempo que "purifica" el flujo de información que los científicos han de tomar en consideración; o la edificación controlada de un *corpus de conocimiento público y consensuado*, que es una de las peculiaridades que se han defendido como propias del conocimiento científico (Ziman, 1968).

Esta multiplicidad de funciones ha de ser tomada en cuenta a la hora de comprender los determinantes que actúan tanto a lo largo del proceso de publicación como sobre la forma definitiva que adquiere el documento científico. Por una parte, justificará la *presión* que existe por

publicar los resultados de las investigaciones, así como el sometimiento de los autores a ciertas reglas y controles que aparentemente sólo añaden trabas a su propósito de comunicar esos resultados. En estos dos factores se apoyará la convicción de que el sistema de publicación formal cubre satisfactoriamente el dominio de los resultados científicos: es poco probable que un resultado valioso se quede fuera porque su autor no sienta la necesidad de publicarlo o porque sea injustamente rechazado; y es más improbable aún que algo se constituya como *resultado científico* completamente al margen de ese sistema oficial. Por otro lado, el tener presente esa multiplicidad de funciones del sistema de publicación dispondrá una base firme para analizar la estructura del documento científico en términos de *reflejo* de esas funciones.

Es fácil, por tanto, comprender la oportunidad de tratar aquí este tema. Tanto los fundamentos de la validez de los indicadores bibliométricos como la definición del dominio de sus posibilidades (y por ende de sus limitaciones) dependen de manera decisiva del funcionamiento del sistema de publicación. En primer lugar y como consideración radical, la *representatividad* de los indicadores descansa en la cobertura que el sistema de publicación formal tiene respecto a los resultados científicos. En segundo lugar, los indicadores bibliométricos requieren que las regularidades que aspiran a capturar estén firmemente garantizadas de algún modo. Si el sistema de publicación estuviera profundamente manipulado o sesgado por factores externos incontrolables y caprichosos, los indicadores bibliométricos carecerían de validez. La interpretación de estos indicadores requiere la existencia de ciertas pautas permanentes (y en algún sentido necesarias) en el sistema de publicación: dicho en otros términos, deben haberse identificado ciertas *causas motrices* que mantienen la forma en que éste se desarrolla, incluso en sus pequeños detalles, si se quieren tomar sus consecuencias observables como base para indicadores interpretables.

En los capítulos que siguen se mostrará cómo el sistema de publicación de la ciencia fue configurándose como respuesta a una diversidad de funciones, al tiempo que servía a otras cuyo alcance toca los mismos fundamentos de la ciencia que llamamos *moderna*. Al hilo del comentario sobre la constitución histórica del sistema de publicación, se irán presentando las funciones que éste iba satisfaciendo, se analizará el papel de los elementos clave en ese sistema y se resaltarán algunas de las consecuencias más relevantes en cuanto a la forma que tal mecanismo

imprimía al conocimiento científico. Se prestará una atención especial a la noción de *reconocimiento* en la ciencia, como clave en el impulso hacia la publicación de los nuevos descubrimientos.

El segundo capítulo se dedica al estudio de la conceptualización que se deriva del ciclo de producción de conocimientos científicos. La noción de *resultado científico* se presenta como necesaria para comprender ese ciclo y para satisfacer los sentidos del reconocimiento discutidos en el primer capítulo. También se tratan los aspectos formales de los agentes científicos más cercanos a la construcción de indicadores bibliométricos..

La dinámica interna generada por la existencia de las revistas científicas abrirá el tercer capítulo. El análisis de la estructura del *documento científico oficial* en términos de las funciones y motivaciones a las que cada una de sus partes remite será el corolario de la anterior exposición, y sobre él se asentarán las bases de la interpretación de los indicadores bibliométricos. Para concluir, se mostrará la conveniencia de utilizar, en lugar del documento completo, una representación de éste (su referencia bibliográfica) como la materia inmediata de los análisis bibliométricos; además, se enumerarán y discutirán las posibilidades de análisis que esa materia prima permite, es decir, se definirá el campo de construcción de indicadores bibliométricos.

## Capítulo 1

### *Bases del funcionamiento del sistema de publicación*

---

#### *1.1. El surgimiento del sistema de publicación y su papel en la constitución de la ciencia moderna*

El sistema de publicación de la ciencia es un sistema complejo, formado por una serie de elementos heterogéneos cuyo comportamiento está regido por un conjunto de precisas normas de funcionamiento. Naturalmente, este sistema no surgió de pronto. Aunque pueden buscarse sus raíces en las recopilaciones de escritos acerca de experimentos de la *Accademia del Cimento* en la Italia del siglo XVI, se considera generalmente que su antecedente directo se constituyó a lo largo de la segunda mitad del siglo XVII en varios países europeos, especialmente en Francia e Inglaterra. Puede decirse que requirió unas pocas décadas para llegar, en lo esencial, a su forma prácticamente definitiva<sup>1</sup>. La estabilidad que ha mostrado a lo largo de varios siglos este sistema es uno de los rasgos que llaman primero la atención al ocuparse de su estudio, sobre todo si se contrasta con las profundas transformaciones que, en otros aspectos, la ciencia ha experimentado durante ese mismo periodo.

---

<sup>1</sup> Ziman (1968: 137). En relación con la historia de las revistas científicas, ver KRONICK (1976), PORTER (1964).

Seguramente se deba buscar la razón de tal estabilidad en la importancia del papel que juega el sistema de publicación en relación con la ciencia que llamamos moderna, en su honda imbricación en ella. Para toda una tradición de estudios sobre la ciencia, este sistema es a la vez expresión y condición de algunos de los rasgos más característicos y propios de esa ciencia. A este respecto, Ziman llega a decir que

[...] la invención de un mecanismo por el cual pueden publicarse por partes los resultados de ciertas investigaciones detalladas ha sido un paso decisivo en el desarrollo del Método Científico. (Ziman, 1968: 135)

A lo largo del presente capítulo se explorará el alcance de esta aserción. El papel del sistema de publicación en la ciencia ha sido valorado desde varias perspectivas teóricas, y puede decirse que ha ocupado un lugar de privilegio en los análisis de la ciencia.

Durante la segunda mitad del siglo XVII, la costumbre de intercambiar cartas, libros o folletos entre personas que se dedicaban a la investigación comenzó a ser sustituida por la publicación de los resultados en revistas científicas que aseguraban un cierto control de lo que contenían sus páginas. En este proceso intervinieron decisivamente las nacientes asociaciones de científicos que, por aquel tiempo, estaban precisando sus objetivos e intereses y con ellos los medios para alcanzarlos. En un trabajo ya clásico, Harriet Zuckerman y Robert K. Merton resaltan este punto:

Las nuevas sociedades y academias científicas del siglo XVII fueron decisivas para la invención social del periódico científico, que comenzó a ocupar un lugar cada vez más importante en el sistema de intercambio científico escrito [...] Estas organizaciones proporcionaron la estructura de autoridad que transformó la mera *impresión* de trabajos científicos en su *publicación*. (Zuckerman y Merton, 1971: 582)

En su estudio muestran cómo ya en la fundación de las *Philosophical Transactions* (1665) -la segunda revista científica por un mes de diferencia, tras el *Journal des Sçavans*- al lado del objetivo principal de promover la difusión del conocimiento científico, existía en la Royal Society (la asociación promotora) la conciencia tanto de las dificultades que debían superarse como de algunas de las ventajas colaterales que se obtendrían. De ese modo, las medidas adoptadas para conseguir ese objetivo primario comenzaron induciendo o reforzando las motivaciones que

podían contribuir a alcanzarlo. Así fueron también definiéndose nuevos papeles, como el del director, responsable de la publicación, o más tarde el de los revisores o evaluadores.

Antes de que transcurriera mucho tiempo, al tratar de resolver los problemas de mantenimiento del periódico, Oldenburg, junto con colegas de la Sociedad involucrados en la cuestión, introdujo varias medidas adaptativas que terminaron por definir el rol del director. El Consejo también reconoció el problema inmediato de tener "suficiente material" para este periódico que se acababa de concebir, y se elaboraron gradualmente los recursos institucionales para inducir a los científicos a publicar en el periódico. Quizá lo más significativo aquí es que el consejo, como patrocinador de las *Transactions*, estaba implicado en su destino y quiso establecer un cierto control sobre su contenido. (Zuckerman y Merton, 1971: 583)

Uno de los primeros obstáculos que debieron superarse fue la actitud de ocultamiento que mantenían los investigadores sobre sus descubrimientos, ante la desagradable posibilidad no sólo de plagio sino de lo que por entonces se llamaba *robo filosófico* (atribuirse como propio un descubrimiento que había realizado otro; llegó a denunciarse el robo físico de originales no publicados). Principalmente este hecho, junto con el del escaso número de personas que se dedicaban seriamente a cuestiones científicas, hacía dudar de que se fuera a disponer de material suficiente para publicar. De modo que, en primer lugar, para animar a los científicos a enviar los relatos de sus investigaciones y logros se esgrimió explícitamente la garantía que supondría la pronta publicación para *proteger la prioridad de los autores*, mediante la certificación de la recepción del manuscrito proporcionada por la Royal Society y por la misma revista.

La preocupación de los científicos por asegurarse la prioridad de sus descubrimientos no responde, desde luego, a ningún comportamiento caprichoso o enfermizo. No se trata de una mera cuestión de vanidad o egotismo, sino que parecen actuar impulsados por las reglas de la comunidad científica. En pocas palabras: preservar la prioridad significa probar la originalidad, y ésta es el camino hacia el reconocimiento.

## 1.2. Prioridad y reconocimiento

El papel del reconocimiento en la carrera de los científicos individuales y su influencia en la organización de la ciencia han recibido varias interpretaciones. Hasta los primeros años setenta, la sociología de la ciencia basada en el planteamiento teórico de Merton representa la corriente dominante; en esta corriente, la noción de *reconocimiento* ocupa una posición central.

En efecto, Merton se ha ocupado de estudiar profundamente (1957; 1960; 1968) el significado del ansia por sentar la prioridad y el de las disputas, nada raras, que surgen entre los científicos por estos asuntos. Según su concepción, el valor asignado a la prioridad en el descubrimiento es una resonancia del funcionamiento del sistema de recompensas de la ciencia, que premia especialmente la *originalidad*<sup>2</sup>. La originalidad contribuye, evidentemente, al avance del conocimiento, y esto parece ser una aspiración primaria entre los científicos. Al igual que los otros sistemas sociales, la ciencia ha establecido sus modos de recompensar el cumplimiento de las normas y el ajuste a sus ideales. Para Merton la comunidad científica se organiza en torno a varios *imperativos institucionales* que actúan normativamente y definen los valores internos de la comunidad. Se suelen denominar mediante el acrónimo mnemotécnico de "los CUDEOS", y son los siguientes: comunismo, universalismo, desinterés y escepticismo organizado (Merton, 1942)<sup>3</sup>.

El *reconocimiento* se entiende entonces como el testimonio social, público, de adecuación a las normas, el símbolo manifiesto de haber hecho bien la tarea. Según Merton, el interés por el reconocimiento es la contrapartida motivacional en el individuo del valor asignado institucionalmente a la originalidad. El sistema de recompensas consiste en la pluralidad de formas honoríficas (que Merton expone con bastante detalle) mediante las que la comunidad científica otorga reconocimiento

<sup>2</sup> Originalidad y prioridad no son lo mismo, aunque están íntimamente relacionadas: dos descubrimientos independientes pueden ser ambos originales pero realizarse uno antes que otro. La prioridad es una fuerte prueba a favor de la originalidad; una vez establecida, es muy difícil para otro demostrar que se había alcanzado el mismo resultado de modo independiente. Merton (1957: 400) señala esta falta de identidad y cita además a Boring (1927) como precedente.

<sup>3</sup> Mulkay (1976) señala que posteriormente se han propuesto algunas adiciones: originalidad, humildad, independencia, neutralidad emocional, imparcialidad (Barber, 1952; Storer, 1966; Merton, 1973; Mitroff, 1974).

público a los individuos que han contribuido de modo significativo al avance de la ciencia.

Para comprender cuál es el sentido del reconocimiento es preciso tener también presente otra de las nociones básicas en la corriente mertoniana acerca del sistema de recompensas, la del *intercambio*<sup>4</sup>, según la cual el reconocimiento sería una especie de pago que la comunidad hace a los individuos por poner en común sus descubrimientos. El conocimiento se paga con reconocimiento, de modo que tanto el individuo como la comunidad salen beneficiados. La recompensa que supone el reconocimiento actúa funcionalmente al servicio de la comunidad como incentivo para que los individuos continúen poniendo a disposición de los demás sus descubrimientos.

Cuando el reconocimiento de la prioridad no se concede o se esfuma de la vista, el científico pierde su propiedad científica. Aunque este tipo de propiedad comparte con otros tipos el reconocimiento de los derechos del "propietario", se halla en agudo contraste en todos los otros aspectos. Una vez hecha su contribución, el científico ya no tiene derechos exclusivos de acceso a ella. Pasa a formar parte del dominio público de la ciencia. Tampoco tiene derecho a regular su uso por otros, retirándola si no se la reconoce como suya. En resumen, los derechos de propiedad en la ciencia se reducen exactamente a esto: al reconocimiento por otros de la parte que le corresponde al científico en el resultado alcanzado. (Merton, 1957: 387)

Este reconocimiento, la valoración pública por las contribuciones a la ciencia, se puede acumular, por así decirlo, y se traduce en el logro de un nivel de *prestigio*, o en otras palabras, en la construcción de una *reputación*. Evidentemente, el prestigio tiene un valor comparativo respecto al resto de los colegas; lo que importa es tener más prestigio que otros. Por ello, las consecuencias del funcionamiento del sistema de recompensas, especialmente en cuanto a su distribución entre los individuos, también han atraído la atención de Merton y de su escuela. El establecimiento de una estratificación social<sup>5</sup> derivada de la cantidad de prestigio que cada investigador consigue, representa un foco de atención principal a ese respecto. La pertenencia a un cierto rango define el

---

<sup>4</sup> Aparece especialmente claro en Merton (1957). Hagstrom (1965) desarrolla este concepto.

<sup>5</sup> Zuckermann (1968); Cole, Cole (1973).

*status*, que se corresponde con un determinado nivel de "honor y de estima". Así, dependiendo del prestigio alcanzado, un científico cualquiera será considerado eminente, excelente, bueno, normal o mediocre. Las consecuencias de pertenecer a una u otra de estas categorías trascienden la psicología de cada individuo, e inciden directamente en el desarrollo de la carrera científica.

Uno de los fenómenos estudiados es el del reforzamiento que favorece a los que ya han conseguido reconocimiento respecto de quienes todavía no lo han conseguido, lo se conoce como el *efecto Mateo* (Merton, 1968). Esto significa que resulta más fácil ser escuchados y valorados para quienes ya tienen un cierto nivel de prestigio que para esos otros que carecen de él. Quienes ya tienen prestigio reciben más medios y más atención, por lo que ya disponen de una ventaja comparativa; pero además, sus logros se tienden a valorar por encima de logros de mérito similar obtenidos por investigadores de un nivel inferior. Por ello, les resulta más fácil aumentar su prestigio. Se trata, en realidad, de un proceso de realimentación del éxito social. Este efecto puede parecer una desviación indeseable del sistema de recompensas, desde la perspectiva de que el reconocimiento debería ser estrictamente proporcional a los logros de cada individuo; Merton, sin embargo, incide también en algunas consecuencias de este efecto que resultan funcionales para la ciencia en su conjunto<sup>6</sup>.

Hemos visto, pues, el modo en que la preocupación que muestran los científicos por la prioridad queda justificada en el planteamiento de Merton, aunque la sinopsis aquí presentada no haga justicia a su riqueza expositiva y conceptual. La imagen ofrecida por Merton está constituida por una elegante urdimbre de relaciones complejas entre las normas y valores de la comunidad, los intereses y motivos individuales en un todo cuyas partes resultan funcionales, es decir, contribuyen al mantenimiento de la institución de la ciencia como empresa común.

Podemos resumir el sentido básico del reconocimiento en el esquema mertoniano así: *conseguir reconocimiento equivale a acumular prestigio*, a hacerse una reputación, es decir, a ser bien considerado por los colegas. Al valorar los méritos investigadores del individuo, lo que los

<sup>6</sup> También han sido estudiadas, dentro de esta aproximación, las desviaciones inducidas por el excesivo énfasis institucional en el valor de la originalidad, como por ejemplo el fraude, el plagio o la calumnia de plagio (Merton, 1957: 404 ss.) y las disfunciones del sistema de recompensas (Merton, 1960: 550; 1968: 559 ss.).

colegas reconocen es la capacidad científica de aquél; demostrar originalidad es demostrar autonomía científica. Los científicos perseguirían el prestigio guiados por una especie de aspiración moral, inscrita en las normas sociales de la comunidad científica, la cual dispone de los mecanismos (el sistema de recompensas) para premiar el triunfo dentro del seguimiento de esas normas.

Lo realmente destacable en el presente contexto es que, al ser el sistema de publicación el sostén principal de las reclamaciones de prioridad de los investigadores, este sistema es al mismo tiempo el núcleo fundamental en el establecimiento de los méritos que conducen a la adquisición de un nivel de prestigio. El sistema de publicación es el *terreno de juego* del sistema de recompensas de la ciencia, y como tal resulta capital en la construcción de la imagen mertoniana de la ciencia. Quizá con cierta exageración, puede decirse que esta imagen se basa primordialmente en las normas y valores (no distinguidos con claridad) que parecen expresarse de un modo privilegiado precisamente en el funcionamiento del sistema de publicación.

Pero esta imagen de la ciencia, que la muestra como un sistema social organizado normativa y funcionalmente en torno al objetivo de desarrollar un contenido cognoscitivo independiente de esa organización social, fue cuestionada seriamente desde principios de los años setenta (Mulkay, 1969, 1976; Barnes, Dolby, 1970; Sklair, 1972). Uno de los centros de la crítica fue la falta de adecuación del comportamiento real de los científicos al conjunto de normas que Merton había propuesto como constitutivas del *ethos* de la ciencia, es decir, como guía de la conducta de los científicos. Se consideró, por ejemplo, que la descripción de Merton era más adecuada para la ciencia del siglo XVII que para la de nuestra época, o que era inaplicable fuera del ámbito académico. Uno de los puntales que señalaron la incompletud de ese *ethos* fue Mitroff (1974), quien propuso que para dar cuenta de los comportamientos de los científicos es necesario postular la existencia de un sistema de contranormas, cada una complementaria y opuesta a las establecidas por Merton: frente al comunismo, puede operar el secretismo; el universalismo puede ser reemplazado por el particularismo; el desinterés, por el interés individual o de grupo; hay ocasiones en las que el escepticismo organizado es sustituido por el dogmatismo organizado. También los seguidores más directos de Merton se esforzaron por clarificar o desarrollar los puntos más débiles del planteamiento. Zuckerman (1977),

por ejemplo, insistió en la separación de dos dominios en el quehacer científico, el público y el privado: sólo en el primero de ellos tendría aplicación el *ethos* descrito por Merton.

Pero además de matizaciones, ampliaciones y correcciones, existía una sólida crítica dirigida a los fundamentos de un tipo de sociología de la ciencia, aquella que excluía de sus análisis los contenidos cognoscitivos de la ciencia por considerarlos neutrales e independientes de la comunidad científica, mientras se centraba en los procesos sociales que explicaban el progreso en esos contenidos. De este modo surgió una serie de enfoques, que suelen englobarse bajo el rótulo alternativo de "sociología del conocimiento científico", y que asumieron en mayor o menor grado la propuesta metodológica del *Programa Fuerte*<sup>7</sup>: estudios de laboratorio (Latour y Woolgar, 1979; Knorr-Cetina, 1981), etnometodología (Garfinkel, 1982), programa relativista (Collins, 1981; 1985), análisis del discurso (Gilbert, Mulkey, 1984), sociología de la traducción (Callon, et al. 1983), etc. Frente a la descripción mertoniana de la ciencia como una institución social funcionalmente sostenida por una estructura estable de normas y valores, estas escuelas enfatizan la contingencia de las normas que operan en la comunidad científica y la relevancia de otros aspectos menos ideales, como por ejemplo el papel de las luchas de intereses de índole diversa en la configuración de los debates científicos y en su resolución. Su centro de atención se sitúa en la descripción de los procesos de construcción social del conocimiento científico.

La constitución de estos enfoques, contrapuestos a la tradición sociológica mertoniana pero también a ciertas tradiciones filosóficas que se han ocupado de la ciencia, ha llegado a ser considerada como una revolución en los estudios sobre la ciencia. Quizá pueda considerarse que lo común a todas ellas es el objetivo de "abrir la caja negra" de la ciencia y aplicar en ese ámbito hasta entonces prohibido<sup>8</sup> los métodos sociológicos para explicar los procesos de generación y validación del conocimiento científico. Dado que representan una alternativa en muchos sentidos radical, conviene, por lo tanto, preguntarse: ¿cuál es la vigencia

<sup>7</sup> El núcleo de la propuesta se configuró en Bloor (1976) y Barnes (1974; 1977).

<sup>8</sup> La delimitación de un ámbito inabordable por el análisis sociológico está expresada en la clásica distinción de Reichenbach (1934) entre *contexto de descubrimiento* y *contexto de justificación*. La sociología mertoniana había aceptado esta distinción y asumido que su análisis sólo podía referirse al contexto de descubrimiento.

de la noción de *reconocimiento* en estas nuevas perspectivas acerca del quehacer científico? ¿Ha sido reformulada o, simplemente, ha perdido su interés? Si la noción de reconocimiento dependiera exclusivamente de un tipo de aproximación a la ciencia, seriamente cuestionado en lo que respecta a ciertos puntos, perdería su valor como garante del interés de los científicos por investigar y publicar, e indirectamente podría quedar maltrecha la confianza en que el sistema de publicación ofrezca un fenómeno representativo para el análisis cuantitativo de la ciencia.

Pues bien, resulta muy llamativo que en las nuevas teorizaciones sobre la ciencia el significado del reconocimiento haya cambiado muy poco respecto al propuesto por Merton. Sigue siendo el principal de los motores que explican el afán de los científicos individuales por realizar descubrimientos originales y por comunicarlos; se considera también que se consigue principalmente mediante la publicación certificada de resultados propios, conformes a ciertas reglas sostenidas por especialistas; y se entiende que el reconocimiento genera el prestigio, la reputación, alcanzados por la valoración de las propias contribuciones por parte del resto de colegas. No se cuestiona en absoluto la existencia de un gran interés de los científicos por conseguir reconocimiento; al contrario, más bien se refuerza la importancia que se asigna al papel de ese interés en la interpretación de las relaciones internas de la comunidad científica, aunque no puede decirse que haya sido objeto de un tratamiento extensivo dentro de esta tradición alternativa.

Pero al lado de estas coincidencias básicas, hay que colocar las divergencias que en este punto separan de la sociología de la ciencia mertoniana a estos enfoques. Aunque puede decirse que una noción de reconocimiento muy parecida a la mertoniana (es decir, como base de la reputación) todavía ocupa una posición privilegiada en los nuevos esquemas conceptuales acerca de la ciencia, el papel que juega en estos esquemas es bien distinto del que tenía en la corriente mertoniana. Las diferencias más importantes se refieren, por un lado, a los *motivos* que impulsan a los científicos a buscar reconocimiento, y por otro, a los *mecanismos* mediante los que finalmente el reconocimiento es otorgado. Esto es fácilmente comprensible, ya que la existencia de un determinado sistema de normas y valores es uno de los blancos atacados, y en la descripción mertoniana este sistema es el fundamento, al mismo tiempo, de la aspiración "moral" de los investigadores y de la correcta distribución de reconocimiento por parte de la comunidad entre los individuos.

Como consecuencia de las diferencias señaladas, se ha producido una cierta extensión del sentido básico de reconocimiento que lo entendía como simple reputación o prestigio basados en los méritos de las contribuciones.

Un claro resumen de la alternativa que sustituye los motivos que inducen a los científicos a buscar reconocimiento nos lo ofrece Barry Barnes (1985: 41ss), al insistir en el interés práctico que guía a los científicos: obtener reconocimiento les es directamente útil, y en cierto modo imprescindible, para el desarrollo de las carreras individuales. En una situación como la de la ciencia contemporánea, en la que los recursos e infraestructuras dedicadas a la investigación son escasos en relación al número de investigadores y a sus necesidades para mantenerse en primera línea, la competencia por conseguir mayor prestigio se comprende porque es precisamente éste el que puede marcar las diferencias en cuanto al acceso a una mayor cantidad de recursos o de instalaciones, a la posibilidad de obtener financiación para proyectos de investigación, etc., todo lo cual permite la exploración sistemática de una línea de trabajo sin demasiadas fatigas o distracciones externas. Como el prestigio científico se consigue mediante la demostración de competencia y originalidad, la prioridad proporcionada por la publicación certificada resulta directamente útil a los investigadores. De hecho, esta vía es prácticamente imprescindible. De la misma manera pueden tomarse en consideración los aspectos más materiales de la carrera científica como motivos relevantes para la publicación: el reconocimiento de méritos para optar a plazas académicas pasa, literalmente, en muchas ocasiones por la revisión y evaluación de la lista de publicaciones de los concursantes. Esto, desde luego, hace descender a un nivel bastante concreto lo que en la escuela mertoniana se concibió fundamentalmente como una mera aspiración moral al reconocimiento de la capacidad científica y del valor de las contribuciones propias, es decir, de ajuste a los ideales de la comunidad científica.

Según la posición descrita por Barnes, aunque cabe admitir que en ocasiones la necesidad de reconocimiento sea psicológica o moral, el reconocimiento funciona como *moneda* en el sistema de recompensas de la ciencia. Esta moneda, exclusivamente emitida por la comunidad científica, representa para los individuos el acceso tanto a las recompensas científicas (hombres, medios para la investigación...) como a las recompensas de la sociedad en general, porque es convertible, por ejemplo, en

mejoras salariales. De este modo, es la base del poder que la comunidad científica tiene sobre los individuos que la conforman y al mismo tiempo garantiza su autonomía respecto al resto de la sociedad. Puesto que nadie más que esa comunidad puede proporcionar reconocimiento científico y éste es una necesidad básica para los investigadores, para obtener reconocimiento hay que satisfacer ciertos requisitos impuestos por la comunidad. En relación a los individuos, no puede entenderse la persecución del reconocimiento sin hacer referencia a su valor práctico, es decir, a su carácter de condición imprescindible en el desarrollo de la carrera científica de las personas que se dedican a la ciencia: para los científicos, obtener reconocimiento es una necesidad concreta.

Valga lo anterior como ejemplo de una concepción alternativa respecto a la motivación que sustenta la búsqueda de reconocimiento por parte de los científicos. Lo principal es notar que, aunque las explicaciones de esta búsqueda puedan ser varias (y no necesariamente incompatibles), lo que no se cuestiona es el impulso que ejerce hacia el descubrimiento y la publicación.

En cuanto al mecanismo de distribución de reconocimiento, hay que destacar la severa crítica que Latour y Woolgar (1979: cap. 5) dirigen contra el sistema de intercambio, eje de la explicación mertoniana acerca la distribución de recompensas. Esta crítica se culmina con la propuesta de la noción de *crédito*, que consideran más general y potente que la de reconocimiento como recompensa. El crédito juega un papel fundamental en el sistema de inversiones de los científicos que gira en torno al valor de la credibilidad de los investigadores, todo lo cual descansa en último término, según su esquema, en la solidez de los contenidos científicos que cada uno aporta.

El sistema de intercambio esbozado por Merton fue desarrollado con más detalle y ortodoxamente por Hagstrom (1965), y es ésta la exposición explícitamente atacada por Latour y Woolgar. El sistema descrito por Hagstrom está regido por una especie de cortesía entre los individuos, derivada del respeto al conjunto de normas característico de la comunidad científica. Puesto que, aparentemente, no se da transacción monetaria alguna, lo concibe como un "intercambio de regalos" entre los investigadores y la comunidad: cuando un investigador pone a disposición de los demás su descubrimiento, éstos le dan educadamente las gracias, y con la valoración de sus méritos el investigador se siente suficientemente recompensado.

Este trueque, según Latour y Woolgar, refleja la asunción de un esquema económico precapitalista, arcaísmo que ha quedado obsoleto en las otras esferas de la sociedad moderna y cuya existencia es, por tanto, difícil de justificar en las relaciones entre los científicos. Además, se trata de una explicación innecesariamente complicada, con elementos inobservables y ocultos por definición: lo esperable, si esas normas operan realmente, es que los científicos nieguen que busquen esa recompensa. Estos autores también hacen referencia a otro modelo, el de Bourdieu (1975), que prescinde de las normas y aplica un modelo económico más complejo, de tipo capitalista. Pero ni éste ni el anterior consiguen explicar un punto fundamental, la *demanda de información fiable*, porque ninguno de los dos se ocupa de los contenidos de la ciencia. Ninguno de los dos modelos, dicen, ayuda a entender el interés que tienen los científicos por leerse los unos a los otros.

Esta insuficiencia proviene de la parcialidad del concepto de reconocimiento (crédito) como recompensa, que atribuyen a la tradición mertoniana; para Latour y Woolgar, existe otra dimensión que es necesario distinguir: el *crédito como credibilidad*, que tiene que ver con la capacidad de los científicos para hacer ciencia realmente. Lo que buscan los científicos no es una mera recompensa sino, ante todo, la credibilidad, porque ésta es la que les abre la entrada al ciclo de conversión de la credibilidad:

Para resumir, sería erróneo tomar la recepción de recompensa como el objetivo último de la actividad científica. De hecho, la recepción de recompensa es sólo una pequeña porción de un gran ciclo de inversión en credibilidad. El rasgo esencial de ese ciclo es la obtención de credibilidad, la cual capacita para la reinversión y la subsiguiente obtención de credibilidad. En consecuencia, para la inversión científica no hay otro objetivo último más que la continua reorganización de los recursos acumulados. Es en este sentido en el que comparamos la credibilidad de los científicos con un ciclo de inversión de capital. (Latour, Woolgar, 1979: 197-8)

Participar en este ciclo es necesario para desarrollar una carrera científica. En esta propuesta, los factores económicos, sociales y epistemológicos actúan juntos, en el mismo plano explicativo de las elecciones profesionales y las trayectorias de los investigadores.

La noción de credibilidad posibilita la conversión de dinero, datos, prestigio, credenciales, áreas de problemas, argumentos, artículos, etc., entre sí. Mientras muchos estudios de la ciencia se centran en alguna pequeña sección de este círculo, nuestro argumento es que cada una de estas facetas no es sino una parte de un ciclo sin fin de inversión y conversión. Si, por ejemplo, retratamos a los científicos como motivados por la búsqueda de recompensa, sólo se puede explicar una pequeña parte de la actividad observada. En cambio, si suponemos que los científicos están comprometidos con la búsqueda de credibilidad, somos más capaces de dotar de sentido tanto a sus diferentes intereses como a los procesos por los que un tipo de crédito se transforma en otro. (Latour, Woolgar, 1979: 200)

El mecanismo por el cual este ciclo se mantiene en marcha descansa en la interdependencia que existe entre los investigadores en la producción de los descubrimientos. Esta dependencia mutua produce una demanda constante de información creíble, lo cual genera una especie de mercado en el que los conocimientos sólidos adquieren un valor determinado por las fuerzas de la oferta y la demanda. Muchos de los movimientos de los investigadores, como cambios de área, de problema, de métodos, de lugar de trabajo, etc., están guiados por la intención de invertir (tiempo, medios, credibilidad) del modo más rentable a su alcance, en términos de obtención de mayor credibilidad, hecho que puede ser percibido por el científico a través de signos de carácter muy diverso. Este esquema explicativo es independiente de las motivaciones concretas de los individuos, puesto que no excluye ni requiere ningún estímulo particular de los que suelen aducirse para justificar las conductas de los científicos. Pero para progresar en la rueda de la ciencia, es inevitable plegarse a estas reglas.

Para un científico en activo, la pregunta más vital no es: "¿saldé mi deuda con él en la forma de reconocimiento por el buen artículo que escribió?", sino: "¿es suficientemente fiable como para ser creído? ¿Puedo confiar en él / en su afirmación? ¿Va a proporcionarme hechos sólidos?" Los científicos se interesan así los unos en los otros no porque vean obligados por un sistema especial de normas a reconocer los logros de los otros, sino porque cada uno necesita al otro para aumentar la propia producción de información creíble. (Latour, Woolgar, 1979: 202)

Para comprender de un modo más ajustado la propuesta de Latour y Woolgar, hay que señalar también que la mercancía básica de este ciclo no es exactamente la información fiable. Una de las características de la credibilidad en este proceso es que se aplica sin hacer una distinción entre la afirmación y quien la sostiene: cuando una afirmación se pone en duda, se está cuestionando la capacidad de quien la produjo y la de sus defensores; cuando una afirmación se impone, la credibilidad de su autor crece. Esto significa que lo que los investigadores ponen, por así decirlo, en el mercado, es su *capacidad para producir información fiable*, y sus logros anteriores son solamente su aval, no su valor. Por eso, cuando convierten esa credibilidad en medios para una investigación, deben producir los resultados esperados si no quieren perder parte de su credibilidad.

De este esquema explicativo se desprende que el reconocimiento a un cierto científico, en el sentido de crédito, es otorgado explícitamente por el resto de los investigadores cuando hacen uso de las afirmaciones de éste, admitiéndolas como válidas y mostrando que son productivas, que son útiles para generar nuevos conocimientos; así, la distribución del reconocimiento no depende de la cortesía de los otros, sino que, al actuar en interés propio en la producción de información fiable, se selecciona y emplea la que se considera mejor, con lo que se deposita la confianza en ciertas afirmaciones y aumenta con ello la credibilidad de los autores de éstas. La base del crédito de un científico, por lo tanto, es la estabilidad y utilidad de sus afirmaciones anteriores. Aunque las menciones formales, como las citas, invitaciones, premios, etc. pueden ser reflejo del crédito de un investigador, la cantidad acumulada se muestra más bien en el grado de confianza que se deposita en ese investigador. En realidad se trata de una posición virtual, estimada, en ese sistema de confianzas, la cual puede tener efectos concretos en la definición de las posibilidades de movimiento y de actuación del investigador.

La explicación de Latour y Woolgar sobre el modo en que se obtiene el reconocimiento de los colegas científicos es muy plausible. Nos sitúa ante un mundo real, en el que no es necesario suponer ideales subyacentes para dar cuenta de comportamientos cuyas causas no son extrañas al sistema general en el que tienen lugar. Proporciona además una imagen en la que se integran en una suave continuidad los motivos y los mecanismos de distribución del reconocimiento; la comparación con el esquema mertoniano resalta aún más este aspecto.

Sin embargo, el retrato que hacen de la posición mertoniana no es completamente justo. El principal punto en el que esto es patente es en el de reducir la noción mertoniana de reconocimiento a la de recompensa. Para Merton y para su escuela, el reconocimiento forma parte de un sistema de intercambio en el cual funciona como recompensa; pero el reconocimiento no consiste en eso. No se reduce a las formas honoríficas por las que se hace manifiesto. El reconocimiento está para ellos relacionado con la valoración del mérito y de la capacidad científicos, basada en la apreciación positiva de las contribuciones realizadas a la ciencia. Es realmente difícil encontrar diferencias significativas con la noción de reconocimiento que proponen Latour y Woolgar como alternativa. La diferencia es radical en cuanto a los mecanismos de obtención del reconocimiento, pero no en cuanto al significado básico de la noción de reconocimiento.

Esto significa que la noción de reconocimiento ha sido central en dos corrientes sociológicas de la ciencia que, aunque difieren en aspectos fundamentales, comparten los rasgos esenciales en lo que respecta a esa noción. No cabe excluir que los mecanismos que cada una ha propuesto para explicar los procesos de distribución del reconocimiento sean en buena medida complementarios: la que insiste en la credibilidad como mercancía parece ajustarse a la mayor parte de la práctica cotidiana de la ciencia; la que sitúa al reconocimiento como centro del sistema de honores funcional para la ciencia en su conjunto, parece reflejar más bien un subconjunto particular, quizá un tanto elitista, el de los científicos de cierta eminencia, que alcanzan una consideración generalizada en el seno de la comunidad científica.

Sin embargo, es justo reconocer que la imagen propuesta por Latour y Woolgar encaja mejor con algunos aspectos importantes. En este esquema se comprende más fácilmente la urgencia de los investigadores particulares por obtener reconocimiento, ya que enlaza esa búsqueda con la particular precariedad de la condición del científico: su continuidad siempre está en peligro, por lo que hay que afirmarla constantemente. En el esquema mertoniano, esa urgencia se hace depender del éxito en la socialización, en la transmisión e incorporación por parte de los individuos de las normas de esa comunidad. En segundo lugar, el mercado de la credibilidad es más creíble (más probable que funcione) que el paraninfo de los honores. Éste peligra con las conductas egoístas, mientras que aquél abarca sin menoscabo los comportamientos

interesados y los generosos. El reconocimiento de los otros es una condición ineludible para hacer valer los intereses propios, es decir, para obtener el propio reconocimiento. La credibilidad propia se basa en utilizar los resultados que dan credibilidad a otros, en un ciclo de dependencia mutua.

En resumen, puede sostenerse que los científicos necesitan reconocimiento y que sólo pueden conseguirlo de los otros colegas, mediante la publicación de resultados originales que puedan ser considerados significativos. Los científicos rara vez escuchan lo que no haya pasado por los cauces formales. Luego existe una alta probabilidad de que cualquier cosa que un científico considere un descubrimiento interesante sea enviada a estos canales: dejar algo interesante en los canales informales supone un elevado riesgo de perder una buena parte del reconocimiento que se podría obtener por ello. En la mayor parte de los casos, tiene el valor de una credencial, puesto que en los *curricula* no se analiza el trabajo en sí, sino sus títulos y lugares de publicación. Si la contribución sobresale, puede convertirse en el camino para la obtención de honores más elevados, con las ventajas adicionales que esto conlleva.

### 1.3. *Los sentidos del reconocimiento*

De la discusión anterior emerge una apreciación más clara de los elementos que intervienen en la comprensión de la noción de *reconocimiento*. En primer lugar, está el contenido mismo del concepto, su *significado*. Luego, los *motivos* que pueden impulsar a los científicos a buscarlo. En tercer lugar, el modo en que se obtiene, que incluye un *acto concreto* y lo que puede llamarse un *mecanismo de distribución*.

Las posiciones que se han comentado comparten el mismo *significado*: el reconocimiento es la *valoración por los otros de la competencia científica de un investigador*, de su capacidad de hacer ciencia en relación con sus colegas. En cuanto a las otras dos dimensiones, ya se han visto las divergencias: para la escuela mertoniana, los individuos buscan el reconocimiento guiados por la internalización de los valores de la comunidad científica y como sanción pública de su conformidad a ellos; lo obtienen al poner sus descubrimientos a disposición de la comunidad y *demostrar originalidad*, gracias al funcionamiento de un sistema de recompensas basado en el cumplimiento de las normas por parte de los

otros individuos. En el esquema propuesto por Latour y Woolgar o en el descrito por Barnes, la necesidad de reconocimiento está vinculada a la continuidad del investigador como tal. Los primeros, además, consideran que la *producción de información fiable* es al mismo tiempo la condición para obtener reconocimiento y lo que configura el mecanismo por el que, sobre un comportamiento egoísta, es necesario dar valor a las aportaciones de los otros científicos para convencer del valor de las propias, ya que la fiabilidad propia se construye sobre la fiabilidad de otros.

Pero aún es posible comprender de otro modo la noción de reconocimiento. Por un lado, es pertinente señalar otros dos significados de éste, además del de *valoración de la competencia científica*: el reconocimiento puede entenderse también como *admisión* en la comunidad científica y como *convalidación* de los resultados obtenidos. Además, como se verá más adelante, puede precisarse de otra manera *qué* es aquello que tiene que hacer un investigador para ser reconocido y *cómo* se le otorga, cuál es el mecanismo por el que alcanza estos tipos de reconocimiento.

Para los científicos, la permanente necesidad de reconocimiento no tiene que ver solamente con asegurar su continuidad, sino también con el refrendo de su inclusión en la comunidad. Esto es así porque, para *ser un científico*, no basta con haber conseguido un título académico o un puesto como investigador; alguien es un científico cuando *es considerado* científico por el resto de los científicos. En eso consiste, primariamente, el reconocimiento: en ser admitido como colega, es decir, como alguien capaz de comprender, aplicar y desarrollar el saber propio de la ciencia o de cierta especialidad de ella. El reconocimiento sirve así para obtener la *admisión* en el grupo<sup>9</sup>. Nadie puede hacer valer su condición de científico sin este tipo de reconocimiento. Quien no es reconocido en este sentido no es un científico más allá del ámbito de los colaboradores inmediatos, es decir, de quienes son testigos directos de su competencia en las actividades científicas; pero en una dimensión más general resulta difícil obtener este tipo de reconocimiento si no es mediante la demostración de una aportación propia.

Este sentido del reconocimiento, como admisión, está implícito lógicamente en el reconocimiento como prestigio o como credibilidad,

---

<sup>9</sup> Pueden encontrarse embrionariamente este sentido en Ziman (1968: 89).

porque es una condición necesaria de estos. Sin embargo, no ha recibido una atención especial por parte de ninguna de esas tradiciones sociológicas. Ser admitido como científico es independiente de la valoración que luego se consiga; no marca las diferencias entre los científicos, sino entre el *dentro* y el *fuera* de la ciencia. Si la capacidad de los científicos se distribuye de una manera tan asimétrica como la sugerida por la ley de Lotka (que no sólo se da en cuanto al número de publicaciones, sino también en el número de citas recibidas), entonces *para la mayor parte de los investigadores el reconocimiento como admisión es prioritario, y pocos pueden aspirar a obtener una valoración particular.*

Estas dimensiones no son excluyentes entre sí (su influencia puede actuar al mismo tiempo), y pueden fundar por separado tanto la funcionalidad de proteger la prioridad como los motivos individuales para reclamarla. No es fácil determinar el peso que cada una de ellas pueda tener en cada caso particular como motor de la publicación, pero es difícil negar que alguna de ellas actúe necesariamente. La continuidad de los científicos como tales está en juego en ambas; si la continuidad de los individuos peligra, también lo hace la comunidad entera.

Pero aún debemos ocuparnos de otra dimensión más del reconocimiento en la ciencia, no derivada de la búsqueda de la prioridad aunque sí vinculada al impulso hacia la publicación de los nuevos descubrimientos. Como las dos descritas, guarda también una intensa relación con la originalidad, pero aquí es de signo opuesto: mientras que en las anteriores demostrar la originalidad era la finalidad que se perseguía, en ésta, al contrario, la necesidad de reconocimiento procede precisamente del carácter problemático de la originalidad. Esta tercera dimensión del reconocimiento es la de la *convalidación por otros* del descubrimiento propio<sup>10</sup>. Procede del hecho de que los conocimientos nuevos llevan asociada una cierta sospecha y son tratados con precaución por los científicos. El principio metodológico del escepticismo (que Merton había elevado a la categoría de imperativo institucional al hablar del *escepticismo organizado*, Merton, 1942) es aplicado por los mismos descubridores en forma de duda ante lo nuevo: la aceptación exige la superación de una crítica severa y lo más radical posible, en otras palabras, los nuevos conocimientos tienen que ser *puestos a prueba*<sup>11</sup>.

<sup>10</sup> Este sentido del reconocimiento por parte de los otros como una salida al solipsismo tiene una cierta tradición filosófica; la conveniencia de su inclusión en este apartado fue sugerida en conversaciones con Fernando Broncano y Jesús Vega.

Tras haber repasado cuidadosamente los procedimientos y razonamientos empleados, el investigador individual tiene que enfrentarse a la posibilidad de un sesgo, error o incompetencia propios. La aceptación de las propias limitaciones<sup>12</sup> por parte de cada individuo forma parte del proceso por el cual los conocimientos son puestos a prueba en la ciencia. Abrir la puerta a esta eventualidad, la de que uno pueda estar equivocado aun cuando cree estar seguro de algo, da entrada al mismo tiempo a la necesidad que tiene cada individuo de asegurarse de que los descubrimientos propios son conocimiento objetivo fiable, independiente de su subjetividad.

La validación por parte de los otros científicos introduce los nuevos conocimientos en el ámbito de lo intersubjetivo y por tanto pueden aspirar a ser objetivos: al lado de las dificultades y alternativas filosóficas para definir la objetividad, parece claro que la intersubjetividad es una condición necesaria para considerar que algo sea objetivo. No es muy arriesgado decir por tanto que el criterio de objetividad pasa por la intersubjetividad aunque esto no agotara la definición de "objetivo". Si creemos que la ciencia esta constituida por conocimiento científico, entonces para que un conocimiento alcance la categoría de *científico*, entre otras cosas habrá de ser compartido por varios investigadores.

La publicación es el principal modo de solicitar esa confirmación a la comunidad entera, es decir, de perseguir el grado más general del reconocimiento como convalidación. Existen, por supuesto, grados inferiores de reconocimiento en este sentido, como por ejemplo el que se obtiene al mostrar o repetir un experimento ante un colega próximo o ante los compañeros habituales de trabajo. Sin embargo, resulta complicado cuantificar la medida en que esta convalidación restringida resulta suficiente para los investigadores y a cuántos de ellos inhibe de buscar una convalidación más general. A pesar de esto, como en las otras

---

<sup>11</sup> Una de las consecuencias paradójicas de la superación de una crítica rigurosa es el fortalecimiento de los conocimientos una vez aceptados, que puede derivar en dogmatismo. La resistencia de los científicos ante lo nuevo (Barber, 1961) ilustra este hecho. Kuhn (1962) señala que los científicos se adhieren dogmáticamente al paradigma ante los casos que lo ponen en cuestión. Mitroff (1974b) en su crítica a la concepción mertoniana de los imperativos institucionales, lo considera una contranorma del escepticismo, también funcional para la ciencia en determinadas circunstancias.

<sup>12</sup> Nos encontramos aquí con una de las dimensiones de lo que Merton (1957: 397) interpreta como la norma institucional de la *humildad*.

dimensiones del reconocimiento que se han tratado con anterioridad, es más que plausible tomarlo como una de las fuerzas que sustentan la costumbre de publicar los nuevos descubrimientos. En este caso, además, desempeña un papel más cercanamente constitutivo de los contenidos de la ciencia, puesto que no se trata de un reconocimiento de los individuos, de sus capacidades, o de la conveniente realización de unas ciertas actividades, sino de la sanción pública de los resultados en sí mismos.

En resumen, es necesario atender a tres sentidos o dimensiones del reconocimiento en la ciencia: como *admisión* en la comunidad científica, como *valoración de la capacidad* en relación a otros colegas, y como *convalidación* de las contribuciones. En los siguientes capítulos se mostrará la conexión de estos tres sentidos con la noción de *resultado científico*, y cómo la publicación formal es prácticamente el único modo de aspirar al reconocimiento en cualquiera de esos sentidos. El mecanismo de distribución de reconocimiento en la ciencia se presentará como una consecuencia necesaria de la constitución de los *resultados científicos*.

#### 1.4. *El sistema de publicación como archivo de la ciencia*

Regresemos ahora al comentario sobre la constitución del sistema de publicación. Además de apreciarse el valor de asegurar la prioridad, tempranamente se advirtió también que la publicación periódica de los nuevos resultados iba a ofrecer el valioso fruto que representaría un *archivo de la ciencia* constituido por la colección de aportaciones que irían sumándose: el cuidado de ese archivo aseguraría la *conservación de las contribuciones* recogidas en él. El valor otorgado a esta permanencia refleja la confluencia de intereses particulares y comunitarios, aunque éstos últimos quizá no siempre fueran plenamente conscientes. Resultaba inmediatamente atractivo para los individuos, pues abría una sólida vía hacia la posteridad<sup>13</sup>. Pero al mismo tiempo cimentaba el carácter acumulativo y social del conocimiento, propiedades esenciales en la ciencia moderna<sup>14</sup>. Gracias a la accesibilidad y al orden proporcionado por ese

---

<sup>13</sup> Zuckerman y Merton (1971: 586) citan una explícita carta de Boyle a Oldenburg, primer director de las *Philosophical Transactions*.

<sup>14</sup> Ziman (1968) ha caracterizado al conocimiento científico como *cosensible y consensuado* y ha vinculado estas características al proceso social que rodea a la publicación formal de los trabajos científicos. Merton (1968: 567) también ha insistido en términos

archivo, las nuevas investigaciones pueden, de modo bastante fidedigno, tomar a las precedentes como punto de partida y aprovechar sus logros. La consolidación del uso de la cita de referencia, especialmente durante el siglo XIX, parece un signo evidente de ese carácter acumulativo, que refleja una actitud distinta a la de las épocas precedentes (Solla Price, 1963: 112). Las conexiones explícitas con la literatura anterior crean un conocimiento vivo, orgánico, resultado de integrar las aportaciones individuales en un proceso continuo de aprovechamiento y de redefinición (Ziman, 1968). Surge así un corpus común en el que el consenso se alcanza mediante la exposición a la crítica de todos.

Esto no significa que cada aportación suponga una mera suma o una extensión del conocimiento que ya existía. En realidad, ese archivo guarda mucho más que simples informes supuestamente concluyentes. Lo que atesora es una continua fuente de inspiración para los científicos en activo que, además de heredar los logros pasados, reciben problemas pendientes, métodos y técnicas, y también indicaciones para no desperdiciar esfuerzos en caminos estériles. Por eso incluso los resultados más revolucionarios, aquellos que conducen a una revisión profunda de los supuestos más básicos, mantienen una impagable deuda con la tradición científica forjada por sus antecesores<sup>15</sup>. Las colecciones de revistas científicas constituyen un depósito común de experiencias cuya utilidad va más allá de la que pueden proporcionar los logros consolidados. El sistema formal de publicación es la mejor garantía de que una contribución, sea cual sea su valor, pueda formar parte del legado que recibe cada nueva generación de científicos. Los logros no publicados corren el riesgo de desaparecer, que es casi lo mismo que no haber existido para la ciencia.

Por eso es inexcusable el señalar que este acervo histórico proporciona la *base para la educación científica* de los nuevos investigadores. Eso le hace desempeñar, como se advierte enseguida, un papel crucial en la perpetuación de la misma comunidad científica. En primer lugar, resulta esencial en la formación y transmisión de una cultura propia de la

---

parecidos sobre el mismo punto.

<sup>15</sup> Estas palabras despiertan enseguida, claro está, el recuerdo de las posiciones kuhnianas que minan la ingenua concepción del progreso científico como una suma de conocimientos. Pero las cuestiones acerca de la *incommensurabilidad* entre paradigmas surgidas del planteamiento de Kuhn en su celeberrima obra (1962) no permiten dudar del papel constitutivo de la tradición en la ciencia (Kuhn, 1977).

ciencia, con sus pautas y técnicas socialmente construidas y asumidas. Aunque los procesos de socialización a través de las prácticas y del contacto directo de los discípulos con los investigadores son en la ciencia de la mayor importancia<sup>16</sup>, la existencia de un registro escrito ofrece por su parte un marco fijo, común y universal en cierto sentido, del que sustancialmente se toma el contenido aprendido por los científicos en su formación. Más allá de esta consideración, este registro puede ser en muchos casos la experiencia directa de los científicos:

Mi argumento aquí es que obtenemos nuestro conocimiento de la ciencia leyendo y pensando los experimentos de otros, que raras veces repetimos nosotros mismos. Esto sería casi imposible en la práctica, y, de hecho, nadie lo favorece, salvo en casos excepcionales; [...] Por lo general, dependemos de la honradez y competencia de los observadores originales. (Ziman, 1968: 53)

Paralelamente, desempeña un importante papel en la diseminación de un lenguaje y un conjunto de conocimientos comunes; esta base compartida les permite a los investigadores mantener los vínculos comunicativos de un modo muy eficiente, no sólo por el uso de lenguajes especiales y aceptados, también porque posibilita la exposición abreviada, dando por supuesto -es decir, omitiendo- una gran parte de lo necesario para comprender el mensaje completo, de modo que sólo hay que transmitir sus aspectos más particulares y diferenciadores. El grado de codificación de las disciplinas ha sido tomado a veces como rasgo relevante en la caracterización de las diferencias entre éstas (Zuckerman y Merton, 1972: 633). Estas consideraciones inciden en el mismo punto: es difícil que algo llegue a ser una auténtica contribución a la ciencia fuera del sistema formal de publicación.

Pero existen otras consecuencias relevantes para el funcionamiento de la ciencia. Si la publicación *certificada* se percibió como un modo de asegurarse la prioridad, la conservación de un archivo de la ciencia sirve también, complementariamente, como sostén de la originalidad cuyo valor está implícito en la persecución de esa prioridad. Algo es nuevo cuando no tiene precedente exacto; la existencia de un registro de las contribuciones pasadas hace legítimo considerar que algo es novedoso cuando no se encuentra en ese catálogo histórico ningún precedente

---

<sup>16</sup> Mulkay (1976). Kuhn resalta la influencia que tienen los procesos de formación en la asunción de un paradigma.

exacto. De ese modo, el archivo sirve también a los intereses de los recién llegados, y no sólo a los de los que ya están en él: puede decirse entonces que sirve para garantizar tanto la prioridad pasada como la futura, porque establece un inventario oficial sobre el que contrastar la prioridad.

Por otra parte, la entrada en ese archivo, con la permanencia que ello supone, *oficializa* la contribución concreta. Esto vale tanto para lo bueno como para lo malo, ya que los errores que puedan cometerse en los documentos dados a conocer a través de este cauce formal también se hacen *oficialmente públicos*. Así se explica en buena parte la cautela que muchas veces muestran los investigadores, especialmente los más experimentados, ante la premura por publicar sus nuevos resultados. El carácter oficial que toman los trabajos publicados equilibra la urgencia por la obtención de reconocimiento con la precaución por la posible pérdida de credibilidad. De este modo se refuerza el autocontrol que lleva a los autores a expurgar de su trabajo cualquier aspecto que consideren dudoso o problemático para evitar una posible mácula en su trayectoria, que sería perenne aunque pudiera llegar a ser compensada por otros aciertos.

Puede verse, por lo tanto, al sistema de publicación de la ciencia como el origen de un archivo bien ordenado de las aportaciones pasadas. Con ello no solamente asegura la permanencia de esas contribuciones, sino que indirectamente incide de varios modos en el desarrollo la ciencia. Puesto que la mayor parte del patrimonio científico que reciben los investigadores (o al menos casi toda la parte "tangible") pueden encontrarlo en ese archivo, ha de resultar muy difícil para ellos imaginar que sus propias contribuciones a la ciencia pueden perdurar al margen de ese sistema de publicación. Al adoptar el punto de vista de la permanencia, se tiende a tomar esas contribuciones como ya constituidas, como algo terminado. Más atrás se señaló que todas las aportaciones conservadas podían ser útiles en algún modo, por ejemplo ahorrando tentativas estériles. Pero es interesante insistir en el proceso de construcción de las contribuciones científicas. Una idea o un escrito que acaba de desprenderse de su autor no es todavía una contribución científica. Al margen de la constante reinterpretación que pueden sufrir los textos a medida que surgen nuevas ideas, es importante apercebirse de cómo es en el sistema de publicación donde las aportaciones son puestas a prueba, evaluadas, absorbidas, en fin, adquieren su sentido y son colocadas en una red

conceptual más general. Los siguientes apartados se ocuparán de los principales aspectos relacionados con este punto.

### 1.5. *El sistema de revisión por pares*

La entrada en ese archivo no carece de restricciones. Desde el principio, se sintió la necesidad de llevar a cabo un control previo a la publicación, un paso que sirviera como filtro protector contra los posibles errores. La publicación en las *Transactions* suponía en la práctica recibir el aval de la Royal Society, por lo que la revisión de los manuscritos era un requisito más que aconsejable si se quería mantener y aumentar el buen nombre de ésta. La revista se convertía de hecho en el boletín oficial de la Royal Society y era el prestigio y la credibilidad de ambas lo que se ponía en juego. Este hecho ha sido interpretado como un eco de los importantes cambios de rol social que se estaban produciendo con la constitución de una comunidad científica organizada (Zuckerman y Merton, 1971: 587-588). El establecimiento de un mecanismo de control de la calidad y adecuación a los cánones científicos estaba orientado a mantener el crédito y reputación de los miembros de la asociación científica emergente y de la asociación misma, que comenzaba a rivalizar con otras instituciones sociales. Pero tuvo además otros efectos:

Esta autoridad basada en una probada competencia tuvo consecuencias mutuamente reforzadas para los científicos, en su triple rol de miembros de la Royal Society, colaboradores de las *Transactions* y lectores de ellas. Estas consecuencias moldearon la temprana evolución del periódico científico y el sistema de árbitros de varias maneras. (Zuckerman y Merton, 1971: 587)

Estos autores señalan algunas de esas consecuencias, para las que también aducen sostén documental. En primer lugar, los científicos supieron apreciar enseguida la ventaja de ser evaluados por jueces competentes en su propia materia: una revisión cuidadosa podía someter sus hallazgos a una exigente prueba y enriquecerlos. Además apareció la costumbre de escribir directamente para publicar en un periódico científico, lo que indujo una presentación y cuidado especiales en los trabajos, por la seguridad de que serían revisados por personas competentes. Esto supuso otra potente base para la autoevaluación que ya se

ha mencionado e influyó, sin duda, en el formato de los escritos. La consecuencia directa más importante fue el desarrollo progresivo de la función del árbitro, que conlleva la responsabilidad de evaluar rigurosamente un trabajo y la aplicación de ciertos procedimientos para conseguirlo. Por último, el interés de los científicos en cuanto consumidores de información pudo influir en el reforzamiento de un control de calidad mediante las protestas que hacían llegar cuando se publicaban materiales dudosos. Como lectores, consumidores de información que podían emplear sus propias investigaciones, a los científicos les interesaba poder confiar en la autenticidad de la obra de los otros, lo que aconsejaba una revisión previa a la publicación.

Todos estos factores fueron influyentes en el establecimiento del sistema de revisión previa a la publicación de los trabajos. El carácter científico de las revistas se aseguraba mediante la revisión de los manuscritos por jueces competentes que rechazaban aquellos cuyos contenidos no se ajustaban a la exigencia de rigor de la ciencia. En poco tiempo, el procedimiento se perfeccionó para alcanzar un elevado grado de eficacia (cumplimiento de los objetivos) y de eficiencia (coste de ese cumplimiento).

Diversos autores han descrito y estudiado aspectos de este sistema tan familiar para los científicos que publican. Pero, más que los detalles concretos, interesa en este lugar un comentario más bien formal de este procedimiento. Podemos resumir en tres palabras el carácter de esta revisión: *paridad, pluralidad, anonimato*. Con la primera nos referiremos al hecho de que la revisión debe ser realizada por colegas del autor, es decir, por personas de la misma condición o *pares*. La necesidad de que sea así proviene del obvio requerimiento de que los jueces o evaluadores posean una capacidad contrastada: no pueden, entonces, ser éstos personas que carezcan de formación científica, tales individuos serían *inferiores* en cuanto al nivel desde el que emiten su juicio; tampoco pueden, por definición, ser personas con unos conocimientos *superiores* a los científicos en ese campo, pues en tal caso, esos otros conocimientos serían considerados los científicos, y el autor de lo que se está evaluando no sería un científico. Por lo tanto, y concediendo en principio al autor que somete su obra al editor la consideración de científico, su trabajo sólo puede ser evaluado competentemente por otro científico experto en el mismo campo, porque lo que se espera de un trabajo científico, lo que lo hace aceptable, es que se adecue a las normas y criterios

establecidos internamente en la ciencia; nadie mejor para verificar ese punto que alguien que está dentro de ella, porque si estuviera fuera sería necesariamente un inferior no competente en la cuestión.

Por supuesto, unos científicos pueden ser mejores que otros, pero al hablar de *iguales* o *pares* nos referimos a que todos ellos son científicos y a que no existe una instancia superior. De modo que estos jueces o árbitros eran científicos cuya capacidad en la especialidad habían demostrado con anterioridad a juicio del editor de la revista, que era el responsable de su selección.

Este primer rasgo de la evaluación, que hemos llamado *paridad*, funda un *control interno* en la ciencia, lo que significa que ésta, en cuanto sistema, se autocontrola. Lo que más interesa señalar aquí es el hecho de que las *acciones y los efectos derivados de este autocontrol deberán ser interpretados en términos de intereses, normas, valores u objetivos propios de la ciencia*, incluyendo naturalmente en ella a sus agentes. Esta exigencia debe ser tenida en cuenta al tratar con indicadores bibliométricos, cuya dependencia de este proceso es patente, y servir, por ejemplo, para eludir las interpretaciones que puedan estar sesgadas por la aplicación o el uso al que se pretende destinar un indicador determinado.

La *pluralidad* es el segundo término definitorio del proceso de evaluación. Se refiere a que no basta aquí con un único juez o árbitro para alcanzar un veredicto fiable, sino que se demanda el dictamen de varios árbitros (dos o más) sobre el mismo trabajo. Lo que se busca es el acuerdo entre varios juicios independientes, ya que la coincidencia de pareceres disminuye la posibilidad de error. Se persigue, por lo tanto, la fiabilidad de la evaluación. Así, cuando dos árbitros discrepan en su opinión<sup>17</sup> sobre el valor de un trabajo, se suele repetir el trámite y someter ese trabajo a la consideración de otros nuevos árbitros, hasta que se obtenga un acuerdo suficiente como para que el editor tome una

---

<sup>17</sup> Bakanic, McPhail y Simon (1987) analizan con detalle el proceso de revisión. Distinguen, como habituales, las siguientes categorías para clasificar el veredicto de los jueces: a) rechazado o enviar a otra revista; b) revisar y someter de nuevo; c) aceptado condicionalmente (con cambios sugeridos); d) aceptado sin condiciones (tal cual). Encontraron que las decisiones finales de los editores se fundan en las recomendaciones de los jueces. Sostienen, junto con otros autores, que el desacuerdo entre los jueces es la norma, al contrario que Zuckerman y Merton (1971:580) que observan y citan un acuerdo elevado y claramente superior a lo esperable por azar. Hay que notar que estos últimos autores emplean sólo dos categorías (aceptado - rechazado) y las posibilidades de acuerdo descienden al aumentar el número de categorías.

decisión acerca de la publicación. La pluralidad es, por una parte, el correlato del reconocimiento en la ciencia de la *falibilidad* incluso respecto a los individuos más preparados, y por otra, la adopción de la *confirmación intersubjetiva*<sup>18</sup> como medio de minimizar los posibles errores.

En tercer lugar, se señalaba el *anonimato* como rasgo característico del proceso de evaluación de los trabajos previo a su publicación en revistas científicas. Consiste en mantener oculta la identidad tanto del autor del trabajo como de los árbitros elegidos para evaluarlo. De este modo, el autor no sabe quiénes le están juzgando; y los jueces no saben quién es el autor del trabajo que están revisando ni tienen noticia sobre el resto de los jueces o sobre sus veredictos. Los árbitros reciben el trabajo y no saben quién lo ha escrito ni quién más lo va a revisar. El anonimato es una condición para que la pluralidad antes descrita tenga los efectos perseguidos, porque protege la independencia de los juicios. Además, focaliza el interés sobre el objetivo principal, que es evaluar si un trabajo determinado es o no aceptable científicamente e interesante, y no, pongamos por caso, juzgar acerca de la capacidad o trayectoria de su autor, de los medios con los que ha contado o sobre cualquier otra característica extrínseca al producto en sí. Con el anonimato se pretende, en especial, asegurar la limpieza del proceso, es decir, que todos disfruten de las mismas condiciones de partida en cuanto a la evaluación: se intenta desactivar la posible ventaja que algunos podrían obtener por medio de influencias o relaciones personales, así como evitar los posibles sesgos psicológicos inducidos en los árbitros bien por disfrutar el autor de un prestigio elevado o, por el contrario, por ser un autor desconocido. Todo ello supone además un estímulo para presentar un buen documento, puesto que se espera que será juzgado solamente por sus méritos intrínsecos.

Si todo marcha correctamente, mediante el funcionamiento del *sistema de árbitros* que se ha caracterizado se obtendrá un veredicto fiable y honesto acerca de si un trabajo determinado es o no *científicamente aceptable*. La capacidad de discriminación de este veredicto es muy reducida, ya que coloca finalmente a todos los trabajos en una de las dos siguientes categorías: aceptados o rechazados. En cuál de ellas acaben los considerados *dudosos* dependerá de la política editorial de la revista.

<sup>18</sup> Recuérdese el tercer sentido del *reconocimiento* que se distinguió en las páginas precedentes.

Seguramente los árbitros podrían decir, en el caso de que tuvieran que comparar directamente varios trabajos<sup>19</sup>, que éste es mejor que aquél otro, e incluso quizá ordenarlos todos por su valor o importancia. Sin embargo, si es que los hubiera, no trascienden los juicios que permitirían establecer diferencias de calidad, interés o relevancia entre las versiones finales de los trabajos aceptados; es más que probable, por supuesto, que los rechazados sean inferiores a los aceptados, pero no se ofrece una valoración escrupulosa que ordene los trabajos evaluados, ni se dispone de información pública sobre los rechazados. Por lo tanto, del hecho de que un trabajo sea publicado tras una revisión como la descrita podemos deducir que ese trabajo contiene *algún valor científico* según los parámetros al uso en la disciplina correspondiente. Supone, en fin, que el trabajo satisface unas exigencias mínimas y que aporta alguna novedad interesante en cierto grado. Además hay que tener siempre presente el hecho obvio de que lo que se ha juzgado es un documento escrito del que se confía sea una especie de informe honesto sobre una investigación, y naturalmente no se ha tenido contacto directo con el trabajo científico que pretende reflejar. Por ello, algunos tipos de posibles errores o engaños no podrían ser detectados con este filtro de la revisión por pares.

### 1.5.1. *La fiabilidad del sistema de revisión por pares*

El modelo que aquí se ha expuesto representa más bien un ideal que una concreción. El sistema de revisión por árbitros puede, en la práctica, no ajustarse plenamente a los requerimientos que se han expuesto, de lo que cabría temer, en principio, que la eficacia y la equidad de la evaluación se vieran afectadas negativamente. Esto nos conduce a la reflexión más general sobre la medida en que este sistema cumple la función que de él se espera, y más concretamente sobre los factores que pueden alejarlo de ese desempeño. Puesto que uno de los objetivos

---

<sup>19</sup> Los jueces evalúan cada trabajo individualmente, y cada trabajo tiene jueces diferentes; por eso, en realidad no se dispone de juicios estrictamente comparables. Pero existen dos comparaciones indirectas: por un lado, los trabajos publicados previamente configuran el marco de lo aceptable, los criterios de los jueces; por el otro, el editor establecerá también un orden de preferencia entre los trabajos recomendados por los jueces si no puede publicar todos. Chase (1970) se ocupa de los criterios empleados para la publicación.

prioritarios parece ser la aplicación de un control de calidad que filtre los trabajos científicamente aceptables, en la operación de este procedimiento pueden cometerse dos tipos de errores: o permitir la publicación de trabajos de una calidad insuficiente o rechazar completamente trabajos relevantes, condenándolos al olvido. Además de ser ambos perniciosos en cuanto a la concesión de reconocimiento por la injusticia que implican, provocan también una pérdida de eficacia en la empresa común: el primero hace perder el tiempo a los lectores, el segundo supone desperdiciar logros útiles. La cuestión es, entonces, si el funcionamiento real de este sistema de control presenta una tendencia significativa hacia uno u otro tipo de error o los errores son aleatorios; en caso de una alta tasa de errores cabría preguntarse si los errores cometidos tienen consecuencias apreciables en el desarrollo de la ciencia.

Distinguiremos tres principales fuentes de error en la operación de este sistema de control por árbitros. En primer lugar, estarían los atribuibles a las equivocaciones en el juicio de los árbitros, ya sean debidos a algún tipo de sesgo o prejuicio psicológico, ya sean producto de la falta de diligencia, de pericia o incluso de honradez. En éstos, lo que se pone en duda es la capacidad del árbitro para emitir un dictamen justo. En segundo lugar, están los posibles incumplimientos formales del procedimiento, es decir, que los jueces designados no sean *pares* estrictamente, o que no se respete la pluralidad o el anonimato; la causa es aquí una mala aplicación que desactiva los mecanismos de protección de la fiabilidad. En tercer lugar, y frente a los de los dos tipos anteriores que podemos considerar *internos* al proceso, hay que tener presente la posible acción de factores *externos* a ese proceso: se introducen en él a través de las preferencias o de las decisiones finales de los editores<sup>20</sup>, que éstos pueden tomar ignorando las recomendaciones de los jueces, guiándose por motivos ajenos a la calidad del trabajo o influidos por la presión de grupos de interés tales como las instituciones o los llamados *colegios invisibles*. Merece la pena que nos ocupemos brevemente de estas posibles causas con el fin de estimar la magnitud del peligro que acecha el funcionamiento justo del sistema de revisión por pares.

Como ya se ha visto, las formas adoptadas para el procedimiento responden a la necesidad por superar las limitaciones como individuos

---

<sup>20</sup> El papel activo de los editores en el proceso de revisión fue estudiado por Crane (1967). Más recientemente ha sido tomado en consideración por Bakanic, McPhail y Simon (1987).

de quienes actúan como jueces y lograr una evaluación satisfactoria. Desde un principio, por lo tanto, está presente la posibilidad de error o de manipulación humana y esas normas se han establecido precisamente para evitar tanto uno como otro. Por lo tanto, puede parecer que los fallos que hemos incluido en la primera clase fueran una mera consecuencia de los fallos de la segunda clase, es decir, del funcionamiento inadecuado (mala aplicación) del procedimiento, y que entonces sólo merece la pena ocuparse de las causas de la segunda clase de fallos. Sin embargo, existe la posibilidad de que, aún aplicando correctamente el procedimiento, el juez no actuara correctamente, por lo que ambas clases de fallos tienen algún grado de independencia.

La sospecha de que los fallos de la primera clase tengan efectos significativos tiene seguramente su origen en la aparente falta de responsabilidad personal de los jueces en relación al veredicto. La cuestión es que los jueces conservan el anonimato incluso después de que hayan tomado su decisión, lo que hace pensar a algunos que aquellos no tienen motivo para comprometerse en realizar bien su tarea. En otras palabras, puesto que no se difunde el nombre del evaluador, éste está a salvo de las consecuencias de sus errores y puede actuar irresponsablemente en el sentido más completo. Se ha llegado a proponer la eliminación del anonimato posterior a la evaluación (Zuckerman, Merton, 1971: 616) con la pretensión de que el aumento de responsabilidad conllevaría una mejora en el desempeño de la función. Pero esta propuesta presenta contrapartidas negativas. La independencia de los jueces debe ser salvaguardada, y ésta también depende de la imposibilidad de recibir represalias originadas en el sentido de sus veredictos. Autores y jueces pueden intercambiar sus papeles en diferentes ocasiones, y es necesario evitar tanto las venganzas por juicios negativos anteriores como la complacencia para inducir juicios benévolos sobre las propias obras en el futuro. Esta connivencia originada en el proceso de publicación también podría incluir otros ámbitos científicos y afectar a la limpieza en el acceso a plazas académicas, a infraestructura o a honores. De modo que no está asegurada la inocuidad de un cambio del procedimiento en este aspecto, sino más bien todo lo contrario.

Es importante señalar que la aparente ausencia de responsabilidad personal por parte de los evaluadores no implica que esa responsabilidad no sea asumida por alguien. De hecho, son los editores de las revistas científicas quienes cargan con ella. La responsabilidad última de lo

que sea publicado en una revista corresponde al editor, así como la del posible rechazo de obras relevantes. El editor tendrá buen cuidado de seleccionar jueces apropiados para cada evaluación. La confianza que deposita en ellos es el resultado de las evaluaciones correctas que haya realizado antes; si un científico se equivoca demasiado en su función de árbitro, pronto se dejará de solicitar su concurso (Merton, Zuckerman, 1971; Ziman, 1968). Existe, pues, un filtro previo sobre la competencia de los jueces. Pero también existe un control de sus evaluaciones, ya que deben ofrecer las razones en que han fundado su decisión, y éstas pueden no resultar convincentes para el editor o editores. En caso de discrepancia entre los dictaminadores, suele repetirse el proceso. Además, se da una fuerte renovación o rotación entre los científicos que desempeñan el papel de árbitros, por lo que la posibilidad de errores sistemáticos (que alguien sea siempre evaluado por el mismo juez sesgado) es pequeña; es necesario reconocer, sin embargo, que existe un cierto margen para que se produzcan injusticias puntuales. En todo caso, cualquier autor que crea que una revista le ha rechazado injustamente su trabajo puede intentar someterlo a la consideración de otra revista (Ziman, 1968: 146): los veredictos no son inapelables. El tipo opuesto de fallo, la publicación de algo irrelevante o erróneo, producirá la indiferencia o la refutación contundente por parte de la comunidad científica, con el precio que ello supone para la revista que lo ha publicado<sup>21</sup>.

Por todas estas razones, parece existir un amplio acuerdo acerca de que es pequeña la posibilidad de que el mal juicio de un árbitro tenga consecuencias graves. Además, la costumbre se inclina muchas veces a favorecer la publicación de los casos dudosos, cuando éstos satisfacen ciertos mínimos, con lo que resulta más probable cometer el error de publicar trabajos insuficientes que el de relegar trabajos buenos. De todos modos, resulta como es obvio muy difícil estimar la posible pérdida de contribuciones relevantes. La existencia de algunos casos notorios ha dado lugar al llamado "síndrome Mendel"<sup>22</sup>, que tiende a elevar la proporción de genialidades desperdiciadas.

---

<sup>21</sup> El comportamiento efectivo de los árbitros ha sido estudiado en muchos trabajos. Para una revisión reciente véase Cicchetti (1991): Algunas aportaciones interesantes son Cole, Cole (1973). Crane (1967), Cronin (1981), Garfield (1986).

<sup>22</sup> Garfield (1979). Ziman (1968) cita el redescubrimiento de un manuscrito rechazado por las Transactions en las que un tal Waterson anticipaba las teorías de Joule. Respecto a este último caso, Ziman señala la difícil comprensión de la exposición. Pero (-grullo),

Más importante en cuanto a la eficacia del sistema de revisión, parece la segunda clase de fallos posibles, la de aquellos debidos a la mala aplicación de las normas que rigen el procedimiento.

Para empezar, puede suceder que el anonimato no siempre sea completamente respetado o, sencillamente, que no haya modo de garantizarlo. Aunque se elimine el nombre del autor, el manuscrito suele contener tantas pistas para un evaluador competente que éste podrá deducirlo con escasa dificultad. Si se eliminaran todas esas pistas (referencia a obras anteriores, tema, métodos e infraestructura, etc.) "no quedaría nada digno de ser publicado" y, por tanto, evaluado. En algunas ocasiones es difícil evitar que el árbitro sepa quién es el autor o autores de lo que está juzgando, lo que, en principio, podría sesgar su valoración.

Hay que recordar que esto no constituiría una violación completa del anonimato tal y como se ha descrito con anterioridad; éste tiene varias dimensiones independientes, cada una de las cuales protege de un peligro distinto. Además del anonimato del autor para los árbitros, que evitaría sesgos debidos principalmente al nivel de prestigio o *status* del autor, existe el anonimato de los jueces respecto al autor, que aísla a aquéllos de presiones de cualquier tipo por parte de éste, y el de los jueces entre sí, que asegura la independencia de los juicios. Parece que estas dos últimas dimensiones del anonimato en la evaluación se cumplen en general satisfactoriamente, por lo que no cabe la duda radical acerca de la eficacia del anonimato. Sin embargo, no se puede obviar sin más la influencia que puede ejercer la consideración social del autor en las decisiones de los jueces. Para Zuckerman y Merton (1971: 605 ss.), la diferencia más significativa se encuentra en el tipo de jueces que son asignados según la categoría del autor: los indicios apuntan, por ejemplo, que es más probable para un científico de primera categoría que sus jueces sean también de primera categoría, en comparación con los autores de categoría inferior. La diferencia relativa de *status* entre autor y evaluador, fuera positiva o negativa, no parecía producir sesgos significativos. Los índices de aceptación crecen ligeramente con la categoría, pero esto es un resultado que entra dentro de lo esperado, puesto que es de suponer que los científicos de categoría superior son capaces de presentar mejores trabajos. Aunque no encontraron confirmación de un

---

claro, los casos conocidos sólo pueden ser casos de pérdidas temporales.

sesgo que favoreciera a los científicos de rango elevado, tampoco puede descartarse que, especialmente en los casos dudosos, disfruten de una pequeña ventaja. Pero, por el contrario, sí hicieron notar que los jóvenes parecían recibir un trato más suave en disciplinas como la física. A la misma conclusión llegaron Bakanic et al. (1987), quienes tampoco encontraron pruebas de un trato preferencial debido al *status*.

Otra posible "negligencia" en la aplicación estricta del procedimiento de evaluación previa es la carencia de una pluralidad de juicios sobre el mismo trabajo. Puede ocurrir que el editor considere suficiente la opinión de un árbitro para tomar su decisión de publicar o rechazar un trabajo, o que incluso no consulte a nadie en absoluto. Dos breves comentarios sobre esta posibilidad: en primer lugar, los editores suelen ser ellos mismos especialistas en alguno de los ámbitos cubiertos por la revista y, por lo tanto, ser considerados árbitros válidos para algunos de los trabajos que reciben; en segundo lugar, no son raros los casos que dejan poco sitio para la duda, ya sea por su calidad notoria, ya sea por sus bochornosas carencias o por sus pretensiones infundadas de originalidad, o más sencillamente, porque sus resultados sean poco más que rutinarios y el riesgo de error muy pequeño. En conclusión, la pluralidad estricta en la evaluación es muy aconsejable, pero no siempre necesaria; es especialmente útil ante los casos dudosos.

### 1.5.2. *Algunas consecuencias del sistema de revisión por pares*

El modo en que se realiza la revisión por árbitros en la ciencia se ha mantenido prácticamente inalterado desde su constitución, sobreviviendo a revoluciones políticas, sociales, económicas e ideológicas. Tal fortaleza inclina a pensar que cumple alguna función relevante para la ciencia con un grado de eficacia más que aceptable. El examen del sistema de evaluación ha mostrado su solidez al tiempo que pone de manifiesto el origen de los posibles errores. Pero una valoración cabal de este sistema de control debe tener en cuenta la naturaleza de su finalidad: no trata de filtrar solamente los trabajos excelentes, sino de *exigir la superación de un nivel mínimo según los criterios al uso en la disciplina correspondiente*. Este control puede permitir quizá la publicación de obras mediocres y vetar ocasionalmente obras trascendentes; es evidente también que existe una gran diferencia de importancia entre las obras

publicadas. Pero la valoración última no corresponde a este sistema de control; éste sólo funda el hecho de que la publicación científica no es completamente libre, sino avalada por ese control. En realidad, el hecho de la publicación es el comienzo de una valoración más precisa y general de los méritos e importancia de cada obra, valoración que realiza virtualmente la comunidad entera de cada disciplina y que deja trazas objetivas de la recepción que ha tenido tal o cual trabajo científico, es decir, de la medida en que ha sido fructífero para la actividad del resto de los colegas.

A lo largo de este capítulo se han tratado las principales ventajas que la revisión por pares reporta, como el papel purificador del flujo de información o el refuerzo para presentar mejores trabajos, dado que la posibilidad de rechazo supone que un relajamiento en la tarea puede convertirse en que todo el esfuerzo se eche a perder. Pero cabe atribuir otras consecuencias al establecimiento del sistema de árbitros, además de las ya descritas. Por ejemplo, en el formato de los escritos científicos, que además de haber alcanzado una gran uniformidad, se ha mantenido estable a lo largo de mucho tiempo. Es bastante sensato suponer que la existencia de una revisión previa como la que se ha descrito jugó un papel importante en la normalización de la presentación, es decir, en la convergencia hacia una estructura común de los escritos científicos realizados para la publicación. La forma de los documentos científicos y las funciones de cada parte serán analizados con cierto detalle en el capítulo dedicado a ello, pero podemos ahora señalar las principales características de esos documentos: planteamiento claro de una hipótesis o problema; separación de datos e interpretaciones; exposición pormenorizada del método empleado; mención explícita de los precedentes científicos y de la pretensión de novedad. Como se observa enseguida, esta forma difiere abruptamente de lo que sería el relato cronológico y preciso de una experiencia científica<sup>23</sup> hecho a partir de las notas de investigación, y es también distinto de la mera comunicación de resultados

---

<sup>23</sup> La forma literaria que adoptan los documentos científicos es el resultado de un proceso que puede rastrearse históricamente. Ziman (1968:53) señala el "curioso" carácter impersonal y expurgado del escrito científico. La transformación literaria que da origen a los documentos científicos y la forma que adquieren ha atraído la atención como un proceso interesante que oculta tras una forma manifiesta intereses y actitudes de los científicos. Varios estudios han enfocado así su análisis de este punto, y entre ellos hay que destacar el trabajo de Knorr-Cetina (1981).

relevantes, o de la mera reclamación de una prioridad. ¿Qué influencia pudo tener el control previo a la publicación sobre la adopción de esta forma expositiva?

Es evidente el interés de los autores al presentar sus trabajos por obtener el *sí de los jueces*. Facilitar la revisión y eliminar las sombras de duda en el ánimo de los jueces aumenta las posibilidades de obtener una evaluación favorable. Esto condiciona la forma y organización del texto, por ejemplo, reforzando la exposición detallada de los procedimientos y la especificación clara de la trascendencia de la contribución. Los recursos para convencer al evaluador de la capacidad del autor -como mostrar conocimiento de los precedentes y de las alternativas, etc.- también debían estar presentes en el texto pues, en principio, el autor permanece anónimo. A medida que se fue imponiendo esta estructura, conformarse a ella se convirtió en símbolo de *estar dentro*, de conocer las pautas y costumbres propias de los científicos, por lo que una variación creativa y peculiar sería recibida con recelo por los evaluadores. Al mismo tiempo, éstos pueden actuar también como garantes de las normas de estilo, redacción y presentación dictadas explícitamente por los editores de la revista que les solicita su concurso.

El control previo fue al mismo tiempo un incentivo para los autores, puesto que la publicación regulada de ese modo revestía a los trabajos de una pátina de rigor y seriedad que beneficiaba inmediatamente el prestigio del autor y aumentaba las posibilidades de ser escuchado por una audiencia selecta, precisamente por aquella por la que le interesaba ser reconocido. Esa misma audiencia se vio favorecida, como ya se ha comentado, con la garantía de haber superado un cierto control de calidad, hecho que aminoraba sustancialmente las posibilidades de perder el tiempo leyendo basura al consultar revistas de carácter científico con evaluación previa. Este hecho debió de suponer, naturalmente, un incremento de la eficacia en la comunicación, al disminuir la cantidad de ruido difundido en forma de trabajos dudosos, erróneos o infecundos.

Desde el punto de vista de los receptores debe señalarse aún otro efecto beneficioso de la implantación del control de las publicaciones, y no es otro que el servicio que prestó para proteger y estimular la competencia leal entre los científicos. Los autores que publican son colegas de los lectores, lo que implica que son también competidores por el reconocimiento. Si lo publicado ha superado un determinado control, las pretensiones de originalidad y capacidad que contiene no serán

completamente vacuas o espúreas. Esto otorga algún valor *a priori* a las publicaciones y las aparta de la sospecha constante, con lo que la mera publicación confiere ya algún grado de reconocimiento (Hagstrom, 1974) y justifica por sí el esfuerzo.

Pero también se han señalado algunas consecuencias potencialmente perniciosas del sistema de revisión previa por pares, derivadas de la posición de ventaja y poder que se adquiere en la función de dictaminador. Una de ellas es la adquisición de información privilegiada, ya que los jueces tienen acceso a los trabajos antes de que éstos sean publicados. Este hecho puede ser interpretado desde varios puntos de vista. Por un lado, el disponer de una información de vanguardia puede ser considerado como una de las compensaciones por realizar un trabajo anónimo y gratuito. Quizá los autores pueden sentir el temor, o incluso la confirmación, de que algunos jueces han proseguido luego, como investigadores, las líneas sugeridas en trabajos evaluados por ellos con anterioridad. Pero, aunque no es despreciable la ventaja de varios meses que puede disfrutar un evaluador respecto al resto de la audiencia, quizá la fuerza de esta posibilidad haya que buscarla en la minuciosa atención con que revisan los trabajos que evalúan, superior a una lectura superficial, lo cual les permite vislumbrar posibilidades a veces no sospechadas por los propios autores (Zuckerman, Merton, 1971: 617). Por otra parte, los evaluadores suelen ser científicos competentes y activos en la misma especialidad que juzgan, por lo que no es sorprendente que trabajen en temas relacionados. En todo caso, la pluralidad en la revisión suele ser un freno suficiente para este tipo de prácticas deshonestas.

Otra de las consecuencias negativas que suelen temerse es la posible formación de círculos de poder que controlan la publicación. De vez en cuando trascienden quejas de algunos científicos acerca de manipulaciones ocultas. No es fácil discernir cuándo esas acusaciones reflejan injusticias reales y cuándo son meros intentos de disfrazar la propia incapacidad científica. Pero tampoco se puede menospreciar la posibilidad de que los poderes establecidos ejerzan su influencia favoreciendo o silenciando la publicación de trabajos inscritos en determinadas líneas<sup>24</sup>; ni se puede descartar que el ansia de este tipo de poder esté en la base de los esfuerzos por crear y mantener revistas científicas. En todo caso,

---

<sup>24</sup> Hull (1988) estudió detalladamente la influencia de las posiciones teóricas de los editores en la aceptación de escritos, sin encontrar sesgos que perjudicaran a los escritos con posiciones contrarias, aunque sí un ligero favor hacia las propias.

hay que señalar que los monopolios estrictos en la ciencia son prácticamente impensables.

En resumen, puede decirse que el sistema de revisión por pares es una de las piezas clave en la formación y sostenimiento de la comunidad científica: difunde reglas, normas y pautas, al tiempo que establece un espacio público para la discusión y para la prueba. Todo apunta a que cumple sus funciones con un elevado grado de fiabilidad, aunque no excluye absolutamente la posibilidad de errores puntuales. Se trata de un procedimiento sostenido socialmente, cuya eficacia, a pesar de los mecanismos dispuestos para maximizarla, no puede sobrepasar en ciertos casos a la del conjunto de individuos que en él participan. El resultado visible de la operación de este procedimiento es la publicación oficial de documentos científicos, de los que se garantiza que poseen un cierto valor científico. De todo lo dicho, se deduce fácilmente que es uno de los puntales que sostienen la validez de los indicadores bibliométricos, que se construyen a partir de conjuntos estadísticamente significativos de documentos publicados en canales oficiales, lo que hace despreciable la influencia de los escasos errores que puedan darse en el funcionamiento del sistema de revisión. Pero como la publicación oficial sólo garantiza la superación de un mínimo, y dada la gran variabilidad en el valor de las contribuciones publicadas, hay que decir que la anterior confianza incluye la recomendación de extremar las precauciones cuando se trabaja con conjuntos pequeños de documentos científicos. Volveremos sobre esto más adelante.



## Capítulo 2

### *La producción de resultados científicos*

---

Una vez tratadas las funciones y los efectos en que se apoya la profunda imbricación constitutiva del sistema de publicación en la ciencia, e identificadas las claves que aseguran la estabilidad de sus pautas, nos ocuparemos ahora de los fundamentos generales de los indicadores bibliométricos que pueden derivarse del funcionamiento de ese sistema. La noción de *resultado científico*, que emerge de la inspección del proceso de generación de los conocimientos científicos, es la respuesta que se propone como base de la construcción de esos indicadores. Esta noción satisface los sentidos del reconocimiento distinguidos en el capítulo anterior y da un sentido particular a los documentos científicos, como se verá en este capítulo y en el siguiente. La definición de criterios de distribución de los resultados afecta de modo general a la aplicación y uso de los indicadores bibliométricos; aquí se comentarán algunos de los aspectos de los principales criterios de agregación.

#### *2.1. El ciclo de producción de conocimientos*

Una de las cuestiones fundamentales en relación con la validez de los indicadores bibliométricos para el análisis de la ciencia es la de su alcance: ¿qué parte de la actividad científica o de sus productos puede ser reflejada por estos indicadores? ¿qué proporción de los frutos valiosos para la ciencia puede escapar o permanecer indetectable mediante el uso de este tipo instrumentos?

Puesto que la base de los indicadores bibliométricos es la literatura oficial de la ciencia, lo que suelen llamarse *canales formales*, es muy conveniente fijar un esquema del papel que juega ese tipo de literatura en el proceso global de la ciencia, tanto para valorar su importancia en él como para tener una idea estimativa de lo que puede quedar fuera de esa literatura oficial.

Esta literatura está constituida por las publicaciones aparecidas en las revistas científicas establecidas, en monografías o incluso en actas de congresos sometidas a un control institucionalizado. Los escritos que pertenecen a esta clase han sido producidos por los autores con el cuidado propio de un informe acabado, sabedores de que carecerá de restricciones en su difusión; además, han superado algún tipo de control por los pares para asegurar que merecen atención, es decir, portan un sello que avala su seriedad.

La importancia de esta literatura, no sólo en la diseminación, sino sobre todo en la constitución del conocimiento científico, ha sido percibida desde hace mucho tiempo:

La producción de nuevo conocimiento [...] no está realmente completa hasta que ha sido transmitida a los otros y deja de ser solamente el conocimiento de una persona. (Machlup, 1962: 14)

Desde la perspectiva de la sociología de la ciencia, la publicación forma parte esencial del proceso por el que la ciencia es una empresa fructíferamente colectiva:

Pero, para que la ciencia avance, no basta concebir ideas fructíferas, elaborar nuevos experimentos, formular nuevos problemas o establecer nuevos métodos. Las innovaciones deben ser efectivamente comunicadas a otros. A fin de cuentas, esto es lo que entendemos por *contribución* a la ciencia: es algo que se da al fondo común del conocimiento. En última instancia, la ciencia es un cuerpo de conocimiento socialmente compartido y convalidado. Para el desarrollo de la ciencia, sólo importa la obra efectivamente conocida y utilizada por otros científicos inmediatamente. (Merton, 1968: 567)

La publicación permite el control institucionalizado que hace que el conocimiento científico sea compartido y asumido por la comunidad, sea el más fiable disponible por haber superado la crítica más exigente, y no pueda ser tomado como conocimiento de iluminados.

Al lado de este tipo de literatura, sin embargo, existe otra nutrida red de intercambio de información entre los investigadores a través de la cual circula otro tipo de escritos que no tienen que someterse a ningún control independiente ni han de exhibir un cuidado o terminación especial en su presentación. Responden a una comunicación ágil entre colegas, de primera mano, habitualmente acerca de trabajos en curso para informar o pedir opinión sobre puntos específicos, y se dan, por ejemplo, mediante comunicaciones personales a colegas, en reuniones científicas o en la coordinación de equipos geográficamente distantes. Suele llamarse *canales informales* a las vías por las que circula esta clase de documentos, algunos de los cuales constituyen la *literatura gris*.

Una limitación atribuida a los indicadores bibliométricos se basa precisamente en la existencia de estos canales informales: se arguye que éstos contienen una parte muy importante del proceso científico que quedaría al margen de la literatura oficial. La impresión que transmite esta objeción es la de concebir que los canales informales son una alternativa que puede ser muchas veces equiparable a los canales formales, como si el que algo aparezca en uno o en otro pudiera depender del azar o de las circunstancias y como si los efectos en la ciencia pudieran ser los mismos en cualquiera de los dos casos.

Una tibia respuesta a esta dificultad ha consistido en señalar que la mayor parte de lo que circula en los canales informales acaba publicado en los formales, por lo que sólo es necesario tener en cuenta estos últimos (Martin, Irvine, 1980; Moravcsik, 1977). La literatura oficial sería, así, una muestra suficientemente representativa. Al mismo tiempo, se ha hecho notar que atender a los dos tipos de canales ocasionaría redundancias incontrolables, debidas a la multiplicidad de versiones que puede tener un mismo trabajo.

No se puede negar que los canales informales cumplen funciones importantes en el proceso científico. Por ejemplo, sostienen los vínculos personales que están en la base de los *colegios invisibles* (Price, 1963), principal modo de mantenerse en la vanguardia de los problemas abiertos. Incluso puede admitirse que estos canales pueden llegar a proporcionar una comunicación más eficiente en algunos casos, transmitiendo información valiosa y precisa que no aparece en los documentos publicados oficialmente o de manera más rápida que en éstos, y también que pueden fecundar investigaciones en curso. A pesar de todo ello, la anterior crítica debe ser rechazada con total rotundidad por varias razones.

Los canales informales y formales no son paralelos entre sí. Algunos documentos publicados proceden de versiones previas que han circulado por canales informales; pero sería completamente absurdo hacer una versión informal de un documento ya publicado oficialmente para que circule por estos canales incontrolados. En lo que ambos circuitos tienen en común, el oficial funciona como filtro depurador de lo estimable que ha aparecido en los canales informales. En este sentido manifiesta Ziman su rechazo a los intentos por mejorar la circulación de los borradores o manuscritos:

El argumento principal es que no deben borrarse las distinciones entre las comunicaciones científicas formales y las informales. El documento científico oficial en una prestigiosa revista no es un anuncio ni una simple noticia; es una contribución al consenso del conocimiento público" (Ziman, 1968: 141)

De modo más radical se expresa Price:

La ciencia [...] debe consistir en artículos científicos que son citados, contados. Por ello, propongo como definición formal considerar a la ciencia como lo que es publicado en los artículos científicos. (Price, 1982: 167)

Estas advertencias no deben tomarse como simples declaraciones de principios. Existe una diferencia esencial y necesaria entre los canales informales y la publicación oficial: los autores sólo pueden aspirar a ser reconocidos mediante la publicación oficial. Es en los canales formales, y sólo en ellos, donde se constituyen los *resultados científicos*, que son el modo de probar la condición y la competencia como científico. Esto hace que exista una fuerza poderosa (aunque no infalible) que empuja a que todo lo digno de consideración termine, bajo una forma acabada, publicado oficialmente.

La noción de *resultado científico*, que se definirá en el siguiente apartado, es clave tanto en el mecanismo de distribución de reconocimiento como en la constitución del conocimiento científico. Para comprender plenamente su significado es preciso ocuparse con cierto detenimiento del proceso concreto del que surge ese conocimiento científico. Existe un esquema generalmente aceptado, que ha sido llamado "ciclo de comunicación y evaluación de la investigación" (Barnes, 1985) y también

"ciclo de producción de conocimientos" (Callon, Courtial, Penan, 1991) que servirá aquí, en lo básico, como guía.

Con cierta arbitrariedad, porque se trata de un ciclo, puede comenzarse con un científico individual o un equipo que está realizando una investigación. Una investigación científica es una *actividad* (es decir, una secuencia de acciones concretas llevadas a cabo por sujetos particulares situados en un contexto material) de la que tradicionalmente se ha considerado esencial su *contenido cognoscitivo*. Para empezar, se trata de una tarea que persigue sistemáticamente la ampliación de los conocimientos disponibles en cierto ámbito. El éxito de esa actividad es, por lo tanto, la obtención de un tipo particular de conocimientos.

Además, la secuencia de acciones concretas que constituye la actividad científica carecería por completo de sentido sin la referencia a un conjunto de saberes que rigen la aplicación de determinadas técnicas: la misma realización de la investigación depende así de una cierta competencia basada en conocimientos. Y para terminar, hay que notar que en el origen de esa tarea está también (en varios modos) un cuerpo heredado de conocimientos. En este último aspecto, lo más evidente es que la superación de las limitaciones o imperfecciones detectadas en ese *corpus* es el principal impulso de la investigación científica (Bunge, 1969: 19).

Así, su punto de partida se encontrará seguramente en la literatura científica precedente, que habrá delimitado o sugerido algún problema pendiente o alguna línea de trabajo prometedora. El enorme peso que el contenido cognoscitivo tiene en la comprensión de la actividad investigadora ha llevado a muchos estudiosos de la ciencia hasta el punto de subordinar por completo las acciones concretas de ésta, como mero instrumento al servicio de los objetivos de conocimiento, y tomar como simples contingencias la situación y el modo en que se desarrollan.

Pero por lo que toca a ese contenido cognoscitivo, enseguida se observa la dualidad inherente a una búsqueda organizada de algo desconocido. A pesar de la puesta en práctica consciente de una serie de operaciones metódicamente dirigidas a aclarar un aspecto habitualmente bien definido, esta labor no está en absoluto libre de sorpresa. En la práctica, esta tarea no suele ser tan lineal como pretenden sus narraciones o sus *reconstrucciones racionales* (Latour, Woolgar, 1979; Knorr-Cetina, 1981). Aunque esté guiada por la comprobación de alguna hipótesis de trabajo, los resultados intermedios son muchas veces inesperados y exigen ciertas decisiones coyunturales que pueden variar el

camino previsto, produciendo un desarrollo zigzagueante y tentativo. Los recursos que los investigadores tienen a su disposición, que incluyen cosas tan heterogéneas como instrumental, ayudantes y técnicos o consumibles, marcan a menudo el campo de actuación, ya sea restringiendo las posibilidades y conduciendo la investigación por caminos alternativos, ya sea ofreciendo medios potentes que animan a afrontar retos complicados. La actividad concreta y las condiciones materiales en las que ésta se desarrolla han de ser tenidos en cuenta como factores centrales en el proceso de generación del conocimiento científico porque determinan tanto los objetivos planteados como los resultados efectivamente alcanzados. Incluso el tipo de institución al que pertenecen los investigadores imprimirá un carácter a su investigación y determinará cuáles son las prioridades.

En algún punto de esa investigación, en la que se toman las precauciones que se consideran necesarias para aislar los posibles factores explicativos, el investigador puede obtener algo que atraiga su atención. Probablemente sea imposible dar una lista completa de las cosas que pueden despertar esa clase de interés; de la manera más general, puede decirse que son aquellas en las que se percibe que pueden suponer una mejora en el estado del conocimiento científico. Representan alguna clase de información, pero ésta puede ser de tipos muy diversos: nuevos datos que confirmen, refuercen o cuestionen una teoría o una hipótesis; alguna técnica útil para otras investigaciones; planteamiento de hipótesis nuevas, síntesis de explicaciones o de principios explicativos; establecimiento de conexiones entre problemas diversos... En fin, una promesa de transformación positiva en el conocimiento disponible.

Lo común es que, tras una reflexión propia, comente ese interés con algún colega muy próximo, quien le dará otra opinión sobre la importancia y fortaleza de lo hallado, o lo someterá a una primera valoración en el seno de su equipo de investigación. Si sus compañeros más inmediatos encuentran que ese interés está justificado, seguramente el investigador preparará uno o varios textos en los que describe los que considera puntos principales, con la finalidad que algunos colegas con los que tiene una relación más estable o a quienes considera expertos en el tema puedan escrutar la viabilidad o potencialidades acerca de esos puntos concretos y le devuelvan sus comentarios. El tipo de documento y su presentación puede variar de un caso a otro; pero no suele ser muy

elaborado, sino solamente una comunicación rápida de lo esencial, quizá con una contextualización genérica de la cuestión.

Estos colegas le dan su opinión, que sirve para que el investigador contraste su propia valoración del trabajo. También pueden indicar algunas sugerencias para mejorar o aclarar algunos puntos del texto o para continuar la investigación y profundizar o explorar en ciertos aspectos.

A veces se trata de datos útiles, pero que son producto de un trabajo rutinario que carece de una valoración especial. En estos casos, ésa u otra versión mejor presentada del documento queda en los canales informales y termina el proceso. Pero si el contenido es más trascendente, es probable que el investigador prepare un informe más trabajado, en el que incorpora las sugerencias recibidas, y lo exponga en alguna reunión científica, como un seminario restringido o un congreso. En esta fase puede recibir apreciaciones por parte de una audiencia especializada.

Si la aceptación es satisfactoria, redacta un trabajo ateniéndose a unas pautas formales muy determinadas, en el que presenta de una forma acabada su contribución, especificando claramente su conexión con algún problema o cuestión, así como las consecuencias positivas de su aportación. Habitualmente, al redactar el artículo el autor ya tendrá una idea aproximada de a qué revista enviará su escrito, por su concordancia con el tema de éste; sin embargo, el enfoque de la revista que elija puede influir en el énfasis en unos o en otros aspectos del trabajo. Puede también que se plantee escribir varios artículos, para varias revistas, profundizando en cada uno de ellos en aspectos diferentes de los resultados alcanzados en su investigación.

Pero además de la adecuación de los contenidos, existen otros criterios que intervendrán en la selección de la revista por parte del autor. La difusión que puede proporcionar la revista o el prestigio de ésta son los principales. Es más fácil obtener reconocimiento si se publica en las revistas de mayor difusión y prestigio. Como es de esperar que todos los autores busquen el máximo reconocimiento por sus trabajos, la competencia por publicar será más aguda en las mejores revistas, con lo que el listón que habrá de superar será más exigente. Sin embargo, los autores tienen una estimación del nivel de su propio trabajo, por lo que, en general, no lo enviarán si no tienen alguna posibilidad de ser aceptados. La estrategia más adecuada es enviarlo primero a la revista

con mayor prestigio de entre las que supone que podrían aceptar su escrito.

Enviar un manuscrito supone ciertos compromisos por parte del autor y de la revista. El autor, además de aceptar las normas formales de la revista, asegura la exclusividad a esa revista en caso de que sea publicado. La revista debe responder lo antes posible acerca de su decisión acerca el trabajo, tras juzgarlo de manera imparcial en relación a sus criterios habituales.

El editor de la revista, habitualmente alguien experimentado en el campo, hace una primera evaluación del interés y calidad del artículo y de su adecuación a la línea editorial del periódico. Si encuentra que su publicación es viable, escoge habitualmente dos especialistas reconocidos que, anónima e independientemente, darán una opinión cualificada sobre el trabajo; éstos evaluadores (*referees*) propondrán cambios en caso necesario o lo rechazarán si descubren algún tipo de insuficiencia seria. Si no hay acuerdo entre los dictaminadores en cuanto a la aceptación o el rechazo, el editor puede solicitar a otros una nueva evaluación y repetir esta fase. El editor deposita su confianza en los evaluadores, quienes se convierten en garantes del nivel científico de la revista; por eso son cuidadosamente escogidos. Al mismo tiempo, ser elegido como evaluador es una forma de reconocimiento de su competencia, con lo que éstos ponen en juego su consideración como científicos.

El juicio de los evaluadores es la base principal de la decisión que toma el editor en cuanto a la publicación del escrito. Envía al autor el veredicto y los comentarios de los evaluadores. Si el artículo es aceptado, el autor incorpora las sugerencias necesarias (que se referirán a cuestiones de detalle) y el artículo aparece, en un plazo de unos meses, en un número de la revista bajo el nombre de su autor. También puede ser aceptado condicionalmente, lo que significa que se requiere una recomposición más profunda del escrito. Si el artículo es rechazado, pero aún considera que tiene algún valor, el autor puede realizar algunas modificaciones e intentará publicarlo en otra revista, y se repetirá el proceso de evaluación.

Cuando un artículo aparece publicado en una revista científica o en otro canal oficial, es virtualmente accesible a toda la comunidad científica de esa especialidad. Gracias a que se suponen todos estos controles, cuando los investigadores interesados en esa materia necesiten obtener información científica de una cuestión particular, solamente

consultarán estas revistas. Para los autores, haber superado este filtro es entrar de modo legítimo en la corriente de aquello que merece la atención de los científicos. Por eso la mera publicación supone ya un reconocimiento, en el sentido de *admisión* que se trató en el anterior capítulo. Pero lo que es más importante es que, en ese momento, el trabajo se ha constituido como un *resultado científico*, es decir, como aquello que puede ser tomado en consideración en el debate del que se genera el conocimiento científico. Es muy improbable que algo llegue a ser un resultado científico al margen de este tipo de mecanismos.

Los resultados científicos sirven a los investigadores en la producción de nuevos resultados en varios sentidos. Pueden fundamentar, marcar una línea de trabajo, incitar alguna comprobación o complemento, etc. Algunos resultados acapararán los puntos centrales del debate y otros tendrán una presencia marginal o nula. Los autores que han producido los resultados más fructíferos recibirán explícita e implícitamente el reconocimiento de su capacidad para realizar buenos trabajos científicos. Los resultados menos significativos no harán que su autor deje de ser considerado científico, pero no le proporcionarán una valoración especial que le distinga. Por otra parte, la existencia de errores en los resultados no es algo que convulsione las relaciones entre los científicos. En caso de ser graves o reiterados pueden cuestionar la confianza en un científico, pero todo este sistema de control es necesario precisamente para depurar los posibles errores.

La materia inmediata para el proceso intelectual del que emerge el nuevo conocimiento científico son los resultados científicos. Los mecanismos concretos que pueden intervenir en este proceso son variados y no vienen al caso. De cualquier manera, del debate en el que los resultados científicos juegan un papel central, surge un conocimiento sólido y sistematizado, al que cada resultado ha contribuido en grados y modos distintos. La detección de ciertas limitaciones en este conocimiento impulsará la búsqueda de otros resultados que puedan contribuir a superarlas, con lo que se cierra el ciclo.

## 2.2. *Las actividades científicas*

Puede decirse que las *actividades científicas* son aquellas necesarias para la consecución de resultados científicos. Tomadas en un sentido

amplio, éstas cubren un variado espectro, pueden estar guiadas por diferentes objetivos inmediatos y producen asimismo una serie de efectos diversos. Un agente científico puede preparar un experimento, estudiar las conclusiones de algún trabajo previo, formar discípulos, comentar un trabajo a petición de un colega, mejorar su capacidad técnica en algún aspecto determinado, contactar con colegas y convencerles para cooperar, reflexionar sobre el ajuste de ciertos datos a las previsiones teóricas, organizar un equipo de investigación, derivar consecuencias de los principios admitidos, comprobar las predicciones de la teoría, argumentar para justificar la necesidad de invertir tiempo y dinero en su línea de trabajo, y otras cosas. El fin último de estas actividades es, desde luego, la consecución de resultados científicos que lleven a la generación de conocimiento, pero la contribución de cada una a este fin puede ser más o menos directa. Además, es posible que no todas las actividades obtengan el efecto deseado. Diferentes grados de eficacia o de fortuna pueden determinar la cantidad y el valor de los resultados alcanzados.

Es importante por lo tanto insistir en la distinción entre la *actividad* científica y los *resultados* de esa actividad. Dos son las principales dificultades: por un lado, de la ausencia de resultados científicos no se puede inferir la falta de actividad; por otro, tampoco es fácil fundar la suposición de una relación constante y universal entre la actividad y los resultados que nos sirva para determinar cuánta actividad se ha desarrollado sobre la base de la cantidad de los resultados.

Respecto a la primera dificultad, hay que decir que para atribuir alguna actividad científica puede ser necesario ofrecer algún resultado científico: la consideración como agente científico puede depender de haber obtenido resultados. En este punto nos encontramos con que los indicadores bibliométricos pueden servir como criterio para determinar quiénes forman parte de la colección de *agentes científicos efectivos*, es decir, de aquellos que han producido resultados científicos, y como ayuda en la caracterización de sus patrones productivos. Sin embargo, en un periodo cualquiera no todos los agentes científicos son efectivos y esa falta de resultados no implica una pasividad científica. Los resultados científicos llegan a menudo tras un largo y costoso proceso de preparación, de comprobaciones minuciosas y de algunos fracasos. En otras palabras, los agentes científicos *efectivos* pueden ser sólo un subconjunto del total de agentes científicos que han trabajado durante el periodo estudiado, es decir, de los *activos*, que a su vez pueden ser una

parte de los agentes científicos de los que dispone un sistema científico determinado, pues puede darse por ejemplo el caso de que no todos los agentes cuenten en un cierto periodo con los medios para investigar o el de que se hallen temporalmente alejados de la investigación; en suma, que estén dedicados a otro tipo de actividades. Podemos llamar por lo tanto a éstos *agentes potenciales*. El conjunto máximo de agentes científicos que podría movilizar en un momento determinado un sistema si sus recursos no fueran limitados es el de los *agentes disponibles* por ese sistema. La medida de la productividad (producción media por agente) ha de tener en cuenta la anterior distinción, ya que define distintos modos de cálculo. Las ratios entre agentes *efectivos* y los otros tipos de agentes también pueden servir para caracterizar la situación productiva de un sistema científico.

En cuanto a la segunda dificultad (suponer que se dé una relación constante entre actividad y resultados), la existencia de tal relación constante implicaría aparentemente asumir que la *eficacia* de la actividad científica es siempre la misma. Sin embargo, saber cuál es la proporción entre la cantidad de actividad desarrollada y el resultado que de ella se ha obtenido en cada caso particular es precisamente uno de los objetivos más apetecidos por quienes se dedican a evaluar o a gestionar la ciencia. Esto es así porque piensan que existen diferencias interesantes y significativas en el valor de esta proporción. La detección de estas diferencias puede servir, por ejemplo, para favorecer preferentemente a los más eficaces o para identificar los factores contextuales cuya manipulación mejore la eficacia del resto. Saber cuánta actividad científica ha desarrollado un determinado agente parece entonces ciertamente algo interesante, pero por desgracia conseguir esta medida no es un empeño que resulte trivial.

Es difícil concebir un modo aceptable de sumar directamente las heterogéneas actividades que antes se han citado, o de obtener un valor que represente este agregado. Este valor se ha llegado a identificar con la medida de los recursos consumidos (Moravcsik, 1977; Martin e Irvine, 1980), incluyendo en éstos tanto los recursos materiales (infraestructura y dinero) como el tiempo empleado (jornadas totales). En lugar de sumar las actividades se sumaría así su precio, con lo que el indicador resultante es el coste total que esas actividades han supuesto. Pero la actividad científica no es lo mismo que gasto de recursos en ciencia, existe una clara diferencia conceptual (la misma actividad puede

desarrollarse de un modo más barato por unos agentes que por otros) cuya eliminación no aporta ninguna ventaja. Pero deriva el problema hacia el de la medida de la *eficiencia* de la actividad científica, ámbito en el que sí tiene sentido la comparación entre el coste de unas actividades y otras en relación con su producción.

El verdadero problema es que la noción de *actividad científica* es en sí muy vaga. Simplemente responde a nuestra convicción de que los resultados científicos no caen del cielo, sino que exigen el esfuerzo de los científicos, que *hacen cosas* para obtener esos resultados. Por eso no se podía, al empezar, dar la *lista completa* de las actividades científicas. Lo que sí debe quedar claro es que *los indicadores bibliométricos tienen relación directa con los resultados científicos y no con la actividad científica*. Es muy peligroso, por lo tanto, llevar demasiado lejos las inferencias acerca de las actividades partiendo de la base de esos indicadores, que se refieren primariamente a los resultados.

### 2.3. Los resultados científicos

Se ha definido la actividad científica como aquella que está orientada a la consecución de resultados científicos. Está claro que el esfuerzo y los recursos empleados en estas actividades no se justifican solamente por el mero placer de dedicarse a ellas: aunque esa dedicación pueda ser gratificante por sí misma, es necesario reconocer que la esperanza en la obtención de algún tipo de resultados guía y explica la realización de la actividad científica, con su enorme coste de trabajo y medios empleados. De hecho, ya desde la constitución de la ciencia moderna, era explícito el deseo de obtener conocimientos útiles para el hombre, prefiriéndolos a ese otro tipo de conocimientos especulativos, alejados de la realidad humana y sin ninguna aplicación práctica.

El concepto de *resultado científico* se aplica intuitivamente y no parece plantear problemas en su uso cotidiano, es decir, al referirse coloquialmente a la ciencia. Sabemos que la ciencia nos ha proporcionado a los humanos una gran cantidad de conocimientos nuevos que han aumentado nuestra comprensión de la naturaleza y nuestro aprovechamiento de ella, o que han contribuido a organizarnos de un modo más eficaz. Creemos también que esa fuente de conocimientos se basa en el esfuerzo de los científicos, que aplicando una serie de reglas y métodos

encuentran soluciones a interrogantes complicados. Sin embargo, esta pragmática clara del concepto *resultado científico* no está asociada a una semántica igualmente transparente. ¿Qué es exactamente un *resultado científico*?

Hay que enfrentarse al hecho de que existe un uso ambiguo o vago del concepto de producto o resultado científico, cuyo significado varía según el contexto y la ocasión. Para empezar, algo puede ser un resultado científico considerado desde *dentro* del sistema científico y otras cosas pueden serlo cuando se consideran desde *fuera* del sistema científico<sup>25</sup>.

Desde el punto de vista *externo* al sistema científico (desde el que miran las personas ajenas a la comunidad científica) la noción de resultado científico parece acercarse más a la de *consecuencias* o *aplicaciones* de los nuevos conocimientos, de los descubrimientos que realizan los científicos. Aquí interesa la utilidad o aplicabilidad de los conocimientos científicos, su rentabilidad, el buen o mal uso que se haga de ellos. En este sentido, los resultados científicos han sido apreciados por diversas partes de las sociedades debido a variados motivos: por proporcionar una base para una potencia bélica superior; por aumentar el prestigio de la nación; por contribuir al bienestar y la salud de los habitantes del país; por los beneficios económicos derivados de su aportación a la potencia industrial; o incluso por elevar el nivel cultural general de la sociedad. Todos estos efectos pueden ser considerados como resultados de la actividad científica (aunque no puedan atribuirse exclusivamente a ella) y han sido la justificación esgrimida para la promoción de la ciencia que han llevado a cabo sociedades o gobiernos a lo largo de los últimos siglos.

---

<sup>25</sup> Cada tradición de estudios sobre la ciencia coloca en un lugar la frontera que separa lo interno de lo externo. Para la filosofía de la ciencia, los factores internos son los relativos a los contenidos y a la lógica propia del conocimiento científico, y los externos son aquellos que proceden de la psicología de los científicos, de sus intereses y organización social o del resto de la sociedad (económicos, culturales, tecnológicos). Los factores externos se conciben a menudo como los distorsionantes que desvían al desarrollo científico de la perfecta coherencia deductiva; estos contaminantes son el precio que la ciencia tiene que pagar por ser real y además humana, son las constricciones o condiciones materiales. Aquí, sin embargo, nos acercaremos más al planteamiento de la sociología de la ciencia e incluiremos en la ciencia no sólo al conocimiento establecido y depurado sino también al que está en proceso o sometido a debate, a los propios científicos y a la especial comunidad que forman.

El modo en que los logros alcanzados mediante la investigación científica son finalmente aprovechados por la sociedad puede variar de un caso a otro; en muchas ocasiones se requiere un proceso complejo y no gobernado directamente por la comunidad científica sino más bien por factores y agentes sociales no siempre determinables con facilidad. Se conocen desde hace tiempo la influencia que pueden ejercer los intereses sociales, y las limitaciones o las posibilidades que se derivan de cada tipo de organización económica, industrial, etc., en la difusión de las aplicaciones de los nuevos conocimientos. Estos intereses no suelen ser explícitos, sino que aparecen al reconstruir explicativamente los acontecimientos pasados.

Pero además, el impacto de los descubrimientos científicos en la sociedad resulta muy difícil de medir también por la naturaleza cualitativa de los cambios que pueden inducir. Las aplicaciones de algunos descubrimientos han modificado tan profundamente la estructura de la sociedad en que se han desarrollado que se puede decir que han creado sociedades nuevas.

Estas someras consideraciones bastan para acordar que no podemos juzgar como resultados científicos sólo aquellos descubrimientos que posean un reflejo directo en la sociedad en forma de aplicaciones útiles. Por un lado, los efectos pueden manifestarse a través de intrincados procesos. Esto dificulta la identificación precisa de la "causa científica", que además puede no haber sido suficiente. Por otro, muchos de los descubrimientos científicos son capitales para que el desarrollo científico siga su curso, y sus consecuencias sobre la sociedad se proyectan a través de otros descubrimientos dependientes de los primeros.

De modo que no disponemos de una noción consistente de lo que sea un resultado científico considerado "desde el punto de vista externo", sino más bien de una conciencia general del hecho de que los descubrimientos científicos tienen un impacto en la sociedad en forma de aplicaciones o de cambios que no serían posibles sin esos descubrimientos. La carencia de esa noción precisa hace verdaderamente difícil concebir una medida cuantitativa que se refiera a ella.

Sin embargo, al reflexionar sobre ese sentido externo se descubre algo interesante: fijarse en las aplicaciones de la ciencia implica algún trato con la idea de efectividad de los logros científicos. Para explicar esa efectividad suele recurrirse a nociones como las de verdad, la de identificación de las causas, o la adecuada descripción de las regularidades de la

naturaleza. En otras palabras, se establece una conexión con lo que se llama conocimiento científico, algo sistemático, general y fiable. Esto conduce a sostener una determinada posición sobre a qué, en sentido interno, se puede llamar un resultado científico, la cual reduce lo valioso de lo obtenido por la ciencia al conocimiento científico. Esta restricción es injustificada y se basa en una visión empobrecida del proceso por el que la ciencia genera su conocimiento.

Para determinar qué debe ser considerado *resultado científico* en el sentido *interno*, es necesario volver al ciclo de producción de conocimiento que se ha descrito anteriormente. En él se presenta la necesidad de una distinción radical entre lo que sea un *resultado científico* y lo que sea el *conocimiento científico*. Tradicionalmente, la caracterización del conocimiento científico ha girado en torno a conceptos tales como *verdad, sistematización, estructura, generalización, capacidad explicativa y predictiva, fiabilidad*, etc., cada uno de los cuales ha sido objeto de tratamiento y énfasis distinto por parte de varias posiciones teóricas aún en conflicto. Por el contrario, se ha visto cómo los resultados científicos pueden incluso ser erróneos, aislados o concretos; aquí se los ha descrito como una fase previa y necesaria para el surgimiento del conocimiento científico. De un modo tentativo, puede decirse que los resultados científicos son una clase particular de información, organizada de un modo especial, vinculada a preguntas de un tipo muy restringido, y con ciertas garantías de fiabilidad, de la cual se espera que contribuirá a la formación de conocimiento científico.

Para acentuar esta contraposición entre resultado y conocimiento científicos, puede atenderse también a los propios agentes que producen ambos. Un investigador es considerado un científico por ser capaz de producir *resultados científicos*, pero no deja de serlo por no haber logrado *conocimiento científico*. Lo que se puede esperar razonablemente de un investigador es que consiga *resultados* interesantes para la ciencia, es decir, que sepa aplicar los métodos de su especialidad y comprender los principios y conocimientos asumidos en ésta, detectar sus limitaciones e intentar superarlas de manera sistemática. La producción de conocimiento científico no suele ser algo completamente individual, y rara vez puede señalarse un responsable único de cada contribución, aunque sí quienes han sido los responsables más determinantes o más directos. De todo esto se deduce que la noción de *resultado científico* es la que puede ofrecer una base sólida para fundar la de *reconocimiento* en la ciencia.

Una vez marcada la distinción entre resultados y conocimiento, y asumida la dificultad de definir este último de manera incontrovertible, ha de intentarse una caracterización algo más precisa de los *resultados científicos*. Una vez asegurado su *carácter científico*, lo que define a estos resultados es su *novedad* y su *relevancia* en relación al conocimiento científico establecido o a los resultados precedentes.

El carácter científico de los resultados procede de que se hayan tomado las *precauciones disponibles* para excluir la ambigüedad o la influencia de otros factores ajenos a los considerados. En cuanto a los otros dos rasgos, ha de señalarse que tanto la novedad como la relevancia requieren de un *marco* para su definición, que es el de lo establecido por la ciencia. La novedad sólo tiene sentido sobre un fondo de contraste que permita resaltar sus diferencias respecto a lo ya existente. La relevancia adquiere su sentido en la conexión con lo establecido: al solucionar o plantear problemas, al simplificar o al integrar teorías, al abrir vías prometedoras, etc. Novedad y relevancia expresan dos movimientos contrapuestos en relación al marco de lo establecido: la novedad separa, la relevancia une.

El conjunto del saber científico es, así, la condición para la constitución de los resultados: prescribe cuáles son las precauciones necesarias, sirve de pasivo contraste para la novedad y demanda respuestas sobre las que se construye la relevancia. Pero, además, su comprensión permite definir el espacio de posibilidades, en el que aparece lo que es nuevo y relevante.

La importancia de unos resultados para la ciencia puede juzgarse desde dos perspectivas. La primera, según el *grado* de novedad y de relevancia. La segunda, según la *transformación* que ha producido efectivamente en el conocimiento científico. Esto es otro modo de decir que los resultados pueden ser prometedores y, sin embargo, defraudar luego las expectativas que han despertado. O que resultados que fueron poco atractivos pueden ser capitales en el desarrollo del conocimiento.

La consideración tanto de la novedad como de la relevancia de un resultado es contextual y depende de la comunidad científica; no se puede establecer un algoritmo que permita asignar novedad o relevancia *a priori* y fuera de ese contexto. Algo es nuevo cuando presenta alguna diferencia respecto a lo ya establecido y, al mismo tiempo, tiene un lugar en lo establecido gracias a que forma parte del conjunto de posibilidades o de expectativas que tenía lo establecido antes de que surgiera lo nuevo

(era algo *imaginable*, aunque nadie lo imaginara antes). La novedad es reconocida por quien conoce la situación previa y está formado para comprender la nueva.

En cuanto a la relevancia, también es percibida: depende en último término del juicio o acuerdo de un subconjunto suficiente de agentes científicos, que otorgan un *papel* en el conjunto de relaciones admitidas. Si algo puede jugar un papel decisivo en el entramado de relaciones conceptuales pero nadie lo ve, no es juzgado relevante por ese grupo. La base objetiva de la relevancia es por lo tanto una condición necesaria, pero no suficiente; algunas cosas capaces de desempeñar una función pueden quedarse sin su papel. Definiremos, por lo tanto, que algo obtenido mediante la actividad científica es un *resultado científico* cuando es asumido como nuevo y relevante en una comunidad científica determinada.

Para terminar, un par de comentarios acerca de las implicaciones de esta manera de concebir los resultados científicos. Primero, no puede dejar de notarse que esta noción es independiente tanto de la *verdad* como del *consenso*, de cualquier manera que se definan ambos. Esto la aparta de una clase de polémicas cuyo fin no excluye su resurrección. También supone que los resultados científicos no contribuyen necesariamente al progreso de la ciencia (o al menos no directamente), como quiera que éste se defina. Siempre acecha la tentación de reducir la ciencia a sus grandes protagonistas. Pero no se puede prejuzgar que todo lo que no pasa a la historia de la ciencia es algo inútil, redundante, o equivocado; muchos datos, intrascendentes cada uno, cuando son reunidos pueden inspirar una nueva teoría. Incluso los errores pueden abrir vías inesperadas.

Segundo, esta noción de *resultado científico* da un sentido completo y bastante natural a la de *reconocimiento* en la ciencia. El camino para la obtención de reconocimiento es la consecución de resultados científicos. Todos los sentidos del reconocimiento están ligados en definitiva a la *capacidad de hacer ciencia*, y conseguir resultados es una prueba de esa capacidad. El mero logro de resultados otorga la inclusión en la comunidad, es decir, la consideración como científico, la *admisión*. Este es el nivel elemental de reconocimiento. Cuando los resultados juegan un papel destacado en el debate, proporcionan a su autor una valoración específica como científico, que permite situarlo respecto a otros colegas. Si sus resultados suponen una contribución más directa al

conocimiento, éstos serán *convalidados*, lo que supone un reconocimiento de un nivel superior a los anteriores.

Esto permite componer un esquema de la distribución de reconocimiento que prescinde de ciertos supuestos. Los científicos pueden pedir reconocimiento simplemente por haber obtenido resultados, no se exige contribución permanente o verdad duradera. No es necesario recurrir a la existencia de valores o compromisos mutuos cuyo cumplimiento depende del respeto a normas. El valor de la originalidad y la importancia de la prioridad pierden su significación y se convierten en algo secundario. Lo que importa es la *demonstración de capacidad científica* a través de la autonomía en la consecución de algo nuevo, relevante y con carácter científico. Esto ocurre de manera automática cuando se considera que unos resultados pueden servir para fundar otros resultados en el sentido que aquí se ha delimitado. Ser reconocido, entre otras cosas, es el único modo de que alguien pueda decir "soy un científico"; sólo puede hacerse legítimamente esto después de haber obtenido resultados científicos.

#### 2.4. Los agentes científicos

Un *agente científico* es alguien capacitado para llevar a cabo las actividades que pueden conducir a la obtención de *resultados científicos*. El modelo para la noción de *agente científico* es, claro, el investigador individual, la persona que después de una formación y aprendizaje de muchos años es capaz tanto de comprender las teorías de una disciplina como de plantear o resolver problemas pertinentes en el seno de esas teorías.

Esto no significa que cada investigador trabaje de manera autónoma y que haya sido autosuficiente en la consecución de sus resultados. La imagen romántica del investigador solitario apenas se sostiene en la ciencia contemporánea, cuyas actividades requieren muchas veces la especialización funcional y el trabajo en equipo. Es muy probable que haya necesitado una infraestructura, unos recursos materiales, unos ayudantes y un personal técnico o de apoyo; muy seguramente además, su trabajo se habrá visto favorecido por los comentarios y sugerencias de compañeros y otros colegas. Pero al menos los resultados podrán

atribuírsele en virtud de su papel esencial en la orientación y planificación de la investigación.

Cabe pensar también que los verdaderos agentes atómicos de la actividad científica sean en muchos casos grupos de investigación y no investigadores particulares. En este caso puede concebirse un agente científico como una *unidad funcional* que sigue un esquema de acción y de planificación coherente, que se organiza en cuanto al consumo de recursos y cuyas partes se supeditan a la persecución de un objetivo común. La investigación en colaboración puede llegar a producir verdaderos *agentes colectivos* en este sentido.

En el caso de los indicadores bibliométricos, es importante además tener en cuenta que los resultados de varios agentes pueden ser sumados, por lo que nos interesa también conocer los criterios que permiten definir *agregados de agentes* que no sean arbitrarios. Los agregados de agentes se constituyen sobre rasgos compartidos por sus componentes, con la intención de explorar los condicionamientos que pueden estar detrás de cada pauta.

Por último, un conjunto de agentes constituyen un *sistema científico* cuando éstos están estructurados (mantienen algún tipo de relación o de influencia entre ellos), comparten un mismo ámbito, están afectados por un conjunto común de reglas o restricciones, tienen una dependencia común en cuanto a los recursos, o están afectados por un mismo sistema externo de actuación y planificación (el mismo poder público, por ejemplo).

Hay que distinguir claramente estos tres niveles. La *colaboración* se refiere a un modo concreto de producir resultados, en el que varios componentes han actuado juntos. La *agregación* de los resultados responde a intereses concretos del análisis, el cual se puede realizar en varios niveles. Es independiente de que haya existido o no colaboración; simplemente aplica un criterio para sumar de un modo significativo un determinado conjunto de resultados científicos. La identificación de los componentes de un *sistema científico* puede realizarse de manera rutinaria, asumiendo la dependencia administrativa o la radicación de los agentes, por ejemplo, o bien como algo problemático, que busca encontrar si la estructuración real entre una serie de elementos define a éstos como un sistema científico.

Los análisis bibliométricos, que parten de la literatura científica, acceden al sistema científico a través de los *resultados*. Sólo serán

identificados mediante este método, por lo tanto, los agentes que anteriormente se han denominado *efectivos*, es decir aquellos que hayan producido resultados durante el periodo. Este hecho obvio ha de tenerse presente, por ejemplo, al calcular indicadores de productividad atendiendo sólo a los datos bibliométricos. La identificación completa de los agentes de un sistema requiere que estos datos sean complementados con otras fuentes.

### 2.5. Los criterios de agregación

Los indicadores bibliométricos se calculan habitualmente para comparar sistemas científicos entre sí o para conocer algo sobre la constitución de algún sistema científico particular. Puesto que el nivel de análisis no está predeterminado, debe escogerse, dentro de las posibilidades, aquél que más se adapte al propósito del estudio. Diferentes criterios pueden ser imaginados con la finalidad de agrupar los resultados; no obstante, los datos bibliométricos disponibles (es decir, los que se pueden obtener del análisis de la literatura científica o de las referencias bibliográficas de esa literatura) imponen un límite sobre la lista de factores que podrían suponerse influyentes en el proceso de producción. En concreto, puede conseguirse información que permita clasificar la producción científica a través de los siguientes criterios: *temático, institucional, geográfico, temporal*.

Cada uno de estos criterios puede ser aplicado en diversos niveles de agregación, solo o en concordancia con alguno de los otros, lo que da lugar a una variedad de combinaciones de entre las que elegiremos aquella que más se adapte a la finalidad del estudio y a las posibilidades metodológicas.

#### 2.5.1. Temático

La distinción de *zonas* en la actividad científica sobre la base de afinidades de contenidos es casi tan vieja como la propia ciencia; tradicionalmente se ha establecido de un modo natural y aproblemático ateniéndose a la especificidad del objeto o de los métodos. No hace falta un largo discurso para que alguien ligeramente familiarizado con la

ciencia en general comprenda las diferencias entre la biología y la física, por ejemplo, o entre la física de partículas y la física del estado sólido, y mucho menos a un investigador que se dedique a alguna de estas áreas o disciplinas. Sin embargo, nunca ha sido una tarea fácil delimitar con precisión cada campo científico, y cada vez lo es menos debido a las fuertes redes interdisciplinarias que se están creando continuamente en la ciencia actual. Al lado del aumento de especialidades se produce también una mayor transferencia de problemas, métodos, resultados y paradigmas entre unas y otras. La ciencia, paradójicamente, multiplica sus ramas, pero aumenta su unidad interna. El problema de las clasificaciones disciplinares siempre está pendiente porque la realidad de la ciencia es más rica que la estructura cerrada de una clasificación. El acuerdo acerca de una clasificación sobreviene, por tanto, cuando se admite el elemento convencional y parcialmente arbitrario que separa cada campo científico de los otros, y se aceptan las limitaciones impuestas por la finalidad de la clasificación.

En la práctica, parece que los científicos, unidos en un principio por compartir objetivos y problemas de investigación, se relacionan para intercambiar información y experiencias sobre temas determinados. De estas relaciones pueden derivar grupos estructurados, o comunidades disciplinares, cuyos miembros lo son *de hecho*, pues para pertenecer a una de estas comunidades sólo es necesario trabajar en el campo de interés de esa comunidad y respetar ciertas normas básicas de relación científica. Estas comunidades son una de las expresiones más evidentes de la especialización temática de los científicos (aunque es posible que un mismo científico pueda sentirse integrado en varias comunidades), por lo que podemos decir que existe entre sus miembros un cierto consenso acerca de las fronteras de cada disciplina y en especial de la propia. La aceptación de *voces autorizadas* acerca de determinada área o disciplina es el reverso del hecho de que se establecen dominios de competencia, territorios vigilados contra toda intromisión de los no expertos en el tema.

Varias son las causas que producen el fraccionamiento de la ciencia en disciplinas. La que parece más evidente es que la enorme profundización en los contenidos crea especialistas por las limitaciones de la capacidad humana. Sin embargo, esto es más el efecto que la causa, y para comprender mejor el motor de este proceso de diferenciación entre las disciplinas hay que recurrir al fuerte carácter competitivo que existe en

el seno de la comunidad científica. La búsqueda de prestigio y de recursos en este entorno favorece que los científicos intenten abarcar todas las posibilidades abiertas y que se muestren interesados y audaces ante las promesas de las nuevas: cuanto menos competidores se tengan, más posibilidades de conseguir medios, alcanzar el éxito y el reconocimiento. Cuando estos nichos han atraído a un número suficiente de investigadores comienzan a actuar las fuerzas estructuradoras de las pautas científicas, primero en forma de unión defensiva contra las especialidades invadidas y más tarde organizando el sistema de control y recompensas propio de las comunidades establecidas (Hagstrom, 1965). Por otra parte, existe además una especialización funcional (tanto temática como técnica) entre los miembros de los grupos de investigación que aumenta la productividad de éstos, de manera análoga a lo que sucede con la división del trabajo en la organización industrial. Esto supone una ventaja respecto a los grupos sin división funcional o los investigadores de formación generalista, lo que hace disminuir las oportunidades de éstos últimos para desarrollar su actividad. De este modo, los individuos (y a veces sub-equipos) potencian sus capacidades en una dirección muy específica, por ejemplo la realización de una determinada técnica experimental o un problema muy restringido que forma parte de una línea de investigación, en detrimento de una visión más completa (Ziman, 1968). Otros factores, como la disponibilidad de nuevos instrumentos, también pueden ser decisivos en el surgimiento de nuevas especialidades (Edge, Mulkay, 1976; Lemaine et al. (eds.), 1976).

Estas especializaciones pueden tener su reflejo en la estructura institucional (con notable claridad en la académica) cuya organización puede seguir, aunque a veces con cierto retraso, las pautas marcadas por el funcionamiento de las comunidades. Por esta razón es necesario tener en cuenta que *la organización disciplinar observable en la producción científica no coincide exactamente con la que se deriva de la organización institucional*. Las inferencias sobre el ámbito institucional basadas en consideraciones sobre las áreas deben ser por lo tanto tomadas con gran cautela, en especial cuando se refieren a instituciones multidisciplinares cuyos departamentos pueden haber solapado sus áreas de trabajo.

En resumen, las clasificaciones temáticas de la producción científica pueden asentarse en al menos tres dimensiones: *proximidad objetiva* de contenidos y métodos, *práctica científica real* e *institucionalización establecida*. Entre estas tres dimensiones podemos esperar una coincidencia

a grandes rasgos, pero lo más habitual es encontrar fuertes discrepancias al descender en el análisis. Las principales causas de esta divergencia hay que buscarlas en el distinto dinamismo de las tres dimensiones y en el diferente grado de posibilidades que cada una ofrece. La *proximidad objetiva* de los contenidos admite innumerables posibilidades en un momento determinado, y éstas cambian con cada nuevo progreso del conocimiento; la *práctica científica* se desarrolla en torno a ciertos núcleos, elegidos de entre los posibles mediante un consenso implícito alcanzado por los agentes; la *institucionalización de las especialidades* en los organismos de investigación llega más tarde (responde lentamente a los cambios sufridos por la práctica de los científicos) y favorece además un carácter más general y estable dentro de cada especialización. Puede estar fuertemente influida además por factores contingentes o históricos, ya que depende del momento en que se fijan las estructuras, lo cual puede depender de la disponibilidad o la influencia de investigadores relevantes o del grado de desarrollo de una disciplina en un momento concreto.

Al desarrollar y aplicar una clasificación temática sobre los indicadores de producción se ha de tener en cuenta la existencia de las tres dimensiones citadas. Debe especificarse explícita y claramente en cuál de ellas se basa el criterio que se ha adoptado, y mostrarse cuidadoso al extraer conclusiones que pueden estar contaminadas por las interferencias entre esas dimensiones.

Las tres dimensiones citadas son accesibles mediante estudios bibliométricos y también son aplicables a los indicadores que resultan de ellos. En la literatura científica, y en particular en las referencias bibliográficas de ella que constituyen la principal fuente para estos estudios, existen datos para definir las áreas o para identificar la pertenencia de un ítem a ellas desde cualquiera de las dimensiones anteriores. Los análisis de contenido no son sencillos de realizar, y la estructura disciplinar establecida en las instituciones presenta fuertes divergencias entre los países e incluso entre las instituciones de un mismo país y tipo, como sucede por ejemplo con la organización departamental de las universidades. Por eso, la base más utilizada para definir las áreas o disciplinas en que se ramifica una producción científica determinada es la de las prácticas que hacen explícita la pertenencia a una especialidad.

Las revistas científicas especializadas son los foros más importantes para la presentación y discusión de resultados científicos. Se consagran a

un núcleo temático determinado, cuya generalidad varía de unas a otras, y se convierten en canales de comunicación y discusión de los investigadores en ese tema. Cuando la especialidad tiene muchos practicantes, pueden coexistir varias revistas con un alto grado de solapamiento en sus contenidos, aunque, al igual que hacen los investigadores, los editores de las revistas tienden a buscar su propio nicho. Este carácter especializado sirve para definir áreas temáticas, mediante la agrupación de revistas afines y relacionadas objetivamente entre sí (mediante las redes de citas, por ejemplo). La clasificación temática que se deriva de estas agrupaciones de revistas es la base para muchos estudios bibliométricos: un documento científico se puede considerar perteneciente o relativo al área en la que está clasificada la revista en que apareció publicado. Nuestro equipo ha aplicado una metodología basada en la adscripción temática de las revistas indizadas por la base de datos multidisciplinar *Science Citation Index* (Maltrás, Quintanilla, 1992; 1995) que ha permitido realizar estudios con un grado apreciable de resolución para la producción científica española. Esta aproximación es útil cuando se están estudiando grandes cantidades de documentos cuya asignación temática no puede realizarse con un examen particularizado y cuando se persigue una caracterización de *grano grueso*.

Sin embargo, las limitaciones de este procedimiento se hacen patentes cuando se intenta acceder a niveles de resolución más finos, en los que puede cambiar además el mismo problema de delimitar las áreas. Más que buscar el modo de aplicar el consenso implícito acerca de la delimitación de las áreas o disciplinas, puede interesar la detección de frentes o polos de investigación que pasan inadvertidos incluso a los propios practicantes implicados en ellos. Con esta finalidad se han venido empleando los análisis de co-citas (o referencias compartidas por los documentos científicos), que se han mostrado útiles en la localización temprana de focos de interés de la investigación (p.e. Small, Griffith, 1974; Small, 1977; Small, Crane, 1979). Varios grupos han desarrollado métodos basados en análisis de palabras compartidas (Courtial, Law, 1989; Courtial, Callon, 1991; Braam, Moed, van Raan, 1988). En una línea similar, en el grupo EPOC de la Universidad de Salamanca se están ensayando otros métodos para la detección de áreas emergentes o de complejos disciplinares surgidos por una relación intensa y firme entre áreas bien establecidas, en particular técnicas de análisis léxico del vocabulario de los documentos científicos (Broncano, Maltrás, Vega, 1995).

### 2.5.2. Institucional

La profesionalización de los científicos es uno de los caracteres más relevantes y peculiares de la constitución de la ciencia moderna. Ya fuera por el interés práctico que comenzaron a despertar sus triunfos tempranos o por el mecenazgo cultural que le brindaron las instituciones públicas, desde el s. XVII los científicos pudieron empezar a dedicarse a la ciencia desde puestos de trabajo en instituciones académicas o de investigación. Este proceso tuvo una gran influencia en la organización de la comunidad científica, ya que los grupos de profesionales empezaron a definir las normas de ingreso y comportamiento que distinguían a la profesión de científico de otras profesiones (Beaver, Rosen, 1978a). También pudo satisfacer el creciente coste de los experimentos requeridos para el progreso de las investigaciones. El resultado, evidente en nuestros días, es que los científicos trabajan como empleados, con los compromisos, exigencias y limitaciones que esto supone. El trabajo de los científicos individuales es posible gracias a la compleja red institucional sobre la que se sostiene la ciencia actual. Los recursos necesarios para la investigación son puestos a disposición de los científicos por los centros académicos, públicos, industriales y de otros tipos en los que éstos desempeñan su tarea. Aunque los investigadores conservan cierta libertad en sus iniciativas, la investigación puede asimismo estar orientada por el programa general de cada centro o institución, y los modos de trabajo quedan fuertemente determinados por su organización interna. La importancia funcional de los centros es tal que, en cierto sentido, pueden ser considerados agentes en el proceso científico. Se convierten desde esta perspectiva en objeto de estudio y de evaluación, atrayendo sobre sí la atención de los responsables en política científica y de los gestores.

La unidad básica en la que se pueden agregar los agentes individuales siguiendo el criterio institucional es el *equipo de investigación*. Operativamente, los equipos de investigación suelen ser los elementos fundamentales de la actividad científica de nuestros días. Su composición puede variar según los casos, pero lo habitual es que en ellos trabaje un grupo de científicos y de técnicos o personal de apoyo bajo la dirección de un investigador principal. También es común la presencia de jóvenes investigadores en formación, lo que acentúa la función esencial que desempeñan los equipos de investigación en la ciencia contemporánea.

El director de un equipo de investigación desempeña un papel central en casi todos los sentidos. No basta un individuo con un talento científico superior, su liderazgo ha de sostenerse ineludiblemente también en otras cualidades. Debe poseer, desde luego, una buena visión general del problema que se intenta abordar, tener cierta capacidad para la prospectiva y unas buenas dotes organizativas; pero no es menos importante que muestre habilidad para seleccionar con tino a sus colaboradores, que sepa extraer el máximo del potencial de éstos, que les inspire la confianza necesaria para perseverar cuando apenas se alcanzan resultados... La dimensión social de los equipos puede ser clave en su funcionamiento. En muchos casos, el director debe exhibir además una vertiente práctica que guíe su consecución de medios para la prosecución de las investigaciones.

Naturalmente, es necesaria la existencia de un cierto grado de permanencia y de estabilidad para que un grupo de investigadores que trabajan juntos pueda ser considerado como un equipo. Sin embargo, los equipos de investigación presentan en general una escasa consolidación formal en la estructura de las instituciones. Este carácter volátil se debe a que estos equipos se constituyen normalmente con objetivos bastante específicos como, por ejemplo, la realización de un proyecto de investigación concreto. No obstante, tras cumplir ese compromiso los miembros pueden emprender en común otros nuevos. Una vez conseguidas unas relaciones de colaboración satisfactorias entre sus miembros, la continuidad de la mayor parte de los equipos depende de haber alcanzado resultados exitosos en los que apoyar su demanda de más recursos. En el nivel de los equipos de investigación, la lucha por captar medios es en cierto sentido una lucha por la supervivencia; es fácil comprender así por qué los equipos están entre los principales protagonistas en el día a día de la política científica, lo que llamamos la *gestión*. La eficacia de ésta puede depender fuertemente en último término de haber contado con el conocimiento adecuado para identificar a los grupos con mayores probabilidades de éxito o para cultivar las condiciones que favorezcan el sostén y la creación de equipos cualificados.

Los equipos de investigación no son necesariamente locales: los miembros de un grupo pueden trabajar en centros, lugares e incluso en instituciones diferentes. Por lo tanto es posible, y no raro, que el equipo traspase los confines de la estructura organizativa establecida. Esto es una forma de decir que a veces los equipos de investigación no

*pertenecen* a una institución determinada, lo que, unido a su falta de estabilidad institucional, explica la dificultad que existe en la práctica para determinar cuáles son los grupos o equipos de investigación activos en un momento dado. Conscientes de la importancia funcional de los equipos, algunas instituciones ya han comenzado a elaborar listas de los equipos que trabajan en ellas. Pero continúa siendo difícil identificar con precisión a los equipos de investigación cuando se llevan a cabo estudios sobre sistemas enteros de investigación en los que están integradas muchas instituciones. Los datos bibliométricos, en concreto los de los autores y los de sus lugares de trabajo, permiten abordar análisis de las redes de coautorías que pueden ofrecer pistas muy relevantes para determinar al menos cómo están constituidos los grupos de investigación *efectivos*; pero es necesario reconocer que al lado de las posibilidades sin explotar, esta línea también tiene sus limitaciones intrínsecas (no siempre es posible reconstruir la composición del grupo a partir de las publicaciones) y que debe ser complementada con otros tipos de información (Vidal, 1995; Vidal, Villarroel, 1995).

En los niveles de agregación superiores al del equipo de investigación se produce un reconocimiento formal de las divisiones o las jerarquías que conforman la estructura institucional. Tratamos ahora con entidades cuyas funciones, ámbito de competencia, derechos y deberes están generalmente regulados mediante estatutos o conjuntos de normas explícitos, es decir, disfrutan de algún grado de reconocimiento legal (en sentido amplio). Son pues entidades fijadas institucionalmente, con una estabilidad elevada y con alguna clase de autonomía para elegir tanto sus objetivos como el modo de alcanzarlos.

Por encima de este umbral, el nivel inferior está representado por los equivalentes funcionales de las Unidades y de los Departamentos. Están muy cerca del nivel operativo, aunque muchas veces se trata de divisiones administrativas. Suelen poseer mecanismos colegiados de decisión capaces de influir, aunque sea indirectamente, en la orientación general de la investigación que llevan a cabo sus miembros o en la cantidad de medios de los que disponen para realizarla. Las unidades o secciones tienen unas competencias más limitadas que los departamentos, en los que suelen estar incluidas, y una especialización más definida que éstos. En ocasiones coinciden con equipos de trabajo.

Los periodos presupuestarios constituyen la base de su programación en cuanto al gasto de recursos y también parcialmente pueden

serlo de su planificación de actividades, aunque estos presupuestos no siempre incluyen los gastos corrientes que soporta el centro o la institución matriz. Este es un detalle que hay que tener presente al evaluar la eficiencia de los departamentos y unidades. Por otra parte, salvo casos muy específicos las actividades no suelen ser exclusivamente de investigación, pueden existir otras prioridades. Particularmente relevante es el diferente peso que suponen las actividades de investigación en los departamentos de cada tipo institucional: un departamento de universidad tiene importantes obligaciones académicas, por lo que sus miembros consumen una parte importante de su tiempo en preparar y dar las clases, incluso cuando la investigación sea una actividad priorizada. En los departamentos de otras instituciones la investigación puede ser la única actividad o ser un subproducto, cuestión importante al comparar departamentos de distintas instituciones.

Gracias a que son entidades institucionalmente reconocidas, su identificación plantea menos problemas que los que se presentan con los equipos de investigación; las propias instituciones de las que dependen pueden facilitar las listas de sus departamentos. Al realizar estudios de producción, sin embargo, se hace evidente la diferencia de costumbres entre los autores en cuanto al modo de señalar en sus publicaciones su lugar de trabajo. Se va generalizando la pauta de indicar primariamente el departamento de trabajo y además la institución matriz, pero no siempre se expresa con claridad el departamento ni tampoco se siguen unas normas comunes: pueden encontrarse una variedad sorprendente de denominaciones para el mismo departamento, incluso dadas por un mismo autor, debido al empleo de distintas abreviaturas e idiomas (de Bruin, Moed, 1990). La agregación de la producción en el nivel de los departamentos y las unidades requiere por tanto un considerable trabajo de normalización de los datos bibliométricos, a veces con la necesidad de acudir a fuentes complementarias.

El siguiente nivel de agregación lo representan los *centros de investigación*. Muchas veces son complejos organizados en torno a unas instalaciones concretas e insertos en alguna especialidad. Pueden ejercer un control riguroso sobre la investigación que se lleva a cabo en su interior, ya que suelen disfrutar de una cierta autonomía en cuanto a la determinación de sus prioridades y al uso de sus recursos. Sus presupuestos son adecuados para estimar el coste real de la investigación que realizan.

Los centros de investigación son, habitualmente, dependientes de *organismos* de un nivel superior, tales como universidades o empresas, los cuales pertenecen a un nivel superior de agregación. Cada una de esas instituciones al completo puede ser un objeto de análisis interesante, debido a que la actividad desarrollada en sus centros dependientes responde en una buena parte a la planificación y la gestión realizada en este nivel. Por eso, la suma de los productos de sus componentes forma un agregado con sentido, puede iluminar aspectos que son clave para la detección de diferencias entre las instituciones.

Hasta aquí, los niveles de agregación son apropiados para análisis pormenorizados de partes concretas de un sistema. Sin embargo, cuando se aborda el estudio de sistemas completos de un cierto tamaño, como puede ser un país, o se intenta comparar varios de esos sistemas, los anteriores criterios de agregación pueden oscurecer el análisis e incluso hacerlo inviable. Para distinguir, siguiendo el criterio institucional, partes significativas de sistemas nacionales, es mucho más útil aplicar el criterio del *tipo institucional* al que pertenecen los agentes productores de los resultados. Es común distinguir varios tipos de organismos o de centros dedicados a la investigación, ya que tanto el tipo de actividad científica que se desarrolla en cada uno como su peso puede ser muy diferente. La distinción más general es la que separa los centros de investigación *académicos* de los *industriales*. Pero en la práctica se suelen emplear otras clasificaciones más específicas. Una dificultad para establecer comparaciones entre los países es que cada uno presenta ciertas peculiaridades en su organización institucional. El acuerdo a este respecto avanza poco a poco, impulsado por el esfuerzo normalizador de organismos internacionales como la OCDE (OCDE, 1980; 1993) o por el efecto mimético producido respecto a los países con estudios más avanzados sobre sus propios sistemas científicos, como puede ser el caso de la NSF (*National Science Foundation*) en los Estados Unidos.

Para el análisis del sistema español, en el grupo EPOC hemos definido una clasificación (Maltrás, Quintanilla, 1992; 1995) que puede ser fácilmente transportable al resto de los países europeos. Distingue los siguientes seis tipos institucionales: universidades (UNIV), hospitales (HOSP), centros del Consejo Superior de Investigaciones Científicas (CSIC), empresas (privadas o públicas) (EMPR), fundaciones y otros organismos de investigación no incluidos en los tipos anteriores (OTRO), y el conjunto de las instituciones extranjeras que han

colaborado con otras españolas (EXTR). El estudio de la estructura institucional del sistema español basado en esta categorización ha revelado interesantes aspectos en cuanto a su articulación y equilibrio (Quintanilla, Maltrás, 1993; Maltrás, Quintanilla, 1995).

### 2.5.3. Geográfico

El lugar donde trabajan los investigadores responsables de los resultados científicos, es decir, donde radican sus centros de investigación, ofrece un criterio básico para realizar distribuciones que ayuden a mejorar el conocimiento sobre los sistemas científicos. La localización geográfica de los centros productores de los resultados proporciona de inmediato un mapa que puede contrastarse, por ejemplo, con el mapa de los recursos científicos. Pero también permite seguir la pista de las influencias que mantienen entre sí las instituciones científicas y otras esferas de la sociedad.

Pueden definirse unidades geo-políticas en varios niveles, pero siempre han de aplicarse criterios justificados. Las principales razones que fundamentan estas distribuciones son las siguientes. En primer lugar, las instituciones científicas pueden agruparse porque disfrutan o padecen las acciones de los mismos poderes públicos. Las políticas científicas (explícitas, o planificadas expresamente, e implícitas, a través de efectos colaterales de otras políticas, como la fiscal o la industrial) generan un ámbito de condiciones que son compartidas por los centros dependientes de la misma administración. Si se trata de valorar comparativamente los efectos de esas políticas, es obvia la conveniencia de separar lo relativo a cada una de ellas. Además, el diseño de cada una de esas políticas requiere un conocimiento separado y susceptible de comparaciones sobre la demarcación correspondiente.

En segundo lugar, la pertenencia al mismo entorno social, cultural y económico define un marco común en cuanto a la cantidad y tipo de los medios disponibles, lo cual incluye las infraestructuras materiales y los recursos humanos. La formación, los modos de trabajo, o las conexiones con el tejido industrial, por poner algunos ejemplos, son factores compartidos que pueden ser tenidos en cuenta a la hora de analizar las diferencias entre este tipo de agregados.

Por último, y derivada de las anteriores, las instituciones pueden estar sometidas a un mismo conjunto de prioridades para su investigación, unas veces impuestas por las administraciones y otras por el contexto social y natural en el que se encuentran. Los problemas o las posibilidades de ese contexto pueden ser determinantes en la orientación de los temas y del tipo de investigación que se emprende, de modo que un agregado de instituciones puede ver delimitado de modo parecido su conjunto de objetivos.

En algunos niveles, la agregación geopolítica puede coincidir prácticamente con la definición de *sistemas científicos* cuando sus elementos mantienen entre sí una integración estructurada. Éste es el caso de los países. Pero en éstos también pueden tener sentido los subconjuntos definidos con el criterio geográfico, los cuales enriquecen la descripción del sistema al que pertenecen.

Constituyan o no un sistema científico, existen otros niveles además del de los países. Por debajo de éstos, el principal es de las regiones, especialmente si éstas tienen una marcada autonomía administrativa o política como es el caso de España o de otros países europeos. La trasposición al caso estadounidense se refiere a los estados federales. La distribución por provincias puede ser útil en casos específicos, cuando se requiere mucho detalle. Por encima del nivel de los países, es habitual comparar grandes zonas geopolíticas como continentes o confederaciones de países.

En resumen, el criterio geográfico produce agregados del máximo interés para la política científica, tanto en lo que respecta a la comparación con otros sistemas científicos como en lo que se refiere a un conocimiento más preciso de la constitución y condicionamientos del propio sistema.



## Capítulo 3

### *Elementos del sistema de publicación*

---

Para concluir la línea expositiva sobre el modo en que los indicadores bibliométricos pueden encontrar su fundamento en el sistema de publicación, toca ahora enfrentarse a los aspectos más concretos de este sistema. En este capítulo, el análisis descenderá a un nivel que puede llamarse, en cierto sentido, material, ya que trata las cuestiones de carácter más técnico en relación con la construcción de indicadores bibliométricos a partir de los elementos básicos del sistema de publicación: las *revistas* científicas y los *documentos* que éstas publican. De las primeras, se resaltarán el papel activo que desempeñan en el proceso de publicación. Los segundos se interpretarán como cristalización de los resultados científicos. El comentario de la estructura de la *referencia bibliográfica*, representación de los documentos que es habitualmente la materia prima de los análisis bibliométricos, cerrará el capítulo.

#### *3.1. Las revistas científicas*

La adopción de un sistema de examen riguroso de los trabajos por expertos fue la marca distintiva que señaló el nacimiento de las revistas científicas, ya que dotó de carácter científico a un cierto tipo de publicación frente a otros tipos que no ofrecían garantías. Pero, una vez establecida esta demarcación básica, la aplicación rigurosa del procedimiento de revisión por pares admite diversos grados de exigencia en la selección, independientemente de que otros contenidos,

además del nivel de exigencia, puedan formar parte de los criterios editoriales que se asuman. Por ello, este mecanismo se puede convertir en uno de los instrumentos básicos en la construcción de la idiosincrasia de las distintas revistas científicas.

El sistema de publicación de la ciencia se organiza en torno a las revistas científicas, que definen el modelo de qué sea una publicación oficial. También pueden surgir canales eventuales, como actas de congresos revisadas, monografías u otros tipos de publicación, pero su consideración dependerá de que sean capaces de asumir correctamente los controles de calidad propios de las revistas científicas. El conjunto de estas revistas y publicaciones que constituyen los canales de publicación de la ciencia en el mundo<sup>26</sup> dista de ser estático: algunas revistas decaen en su prestigio, otras ascienden; nuevas revistas surgen cada mes, mientras otras desaparecen o se transforman mediante escisiones, fusiones o cambios de área. La dinamicidad del conjunto de revistas científicas es un rasgo llamativo a la hora de preguntarse acerca de las fuerzas que gobiernan este sistema y sobre algunas de sus consecuencias (Leydesdorff, Cozzens, 1992).

La instauración del sistema de revisión previa de las publicaciones científicas introdujo un control de calidad que repercutió también sobre las revistas que lo aplicaban. Clasificar un trabajo cualquiera como *acceptable* o como *no acceptable* no depende solamente de determinados parámetros objetivos; los criterios de la aceptabilidad pueden contener también cierta arbitrariedad sobre cuál sea el nivel mínimo exigible, y de la altura en que ese nivel se sitúe resultará un tamiz más o menos fino. La calidad media de los trabajos aparecidos en una revista guardará una innegable relación con la severidad impuesta en la selección. Esto ofrece un resorte que permite intervenir en la calidad de una revista, lo que significa que ésta puede jugar un papel activo: no se limita a publicar todo lo que recibe, sino que lo selecciona según sus criterios. De este modo, la competencia entre las revistas no tardó en cobrar sentido, no sólo por la aspiración a ser la mejor revista sino también por salvar la propia continuidad. Las revistas que aceptan trabajos de escaso nivel matan el interés de los lectores por consultarlas y el de los buenos autores por enviar sus trabajos, con lo que su permanencia peligra

<sup>26</sup> La estimación de su número depende de lo estrictos que sean los criterios en cuanto a calidad, difusión, periodicidad, ámbito de influencia, estabilidad, etc.; puede ir desde unos pocos millares hasta varias decenas de millares.

seriamente. Por el contrario, las que publican buenos trabajos seguramente aumentarán su difusión y atraerán la colaboración de los mejores autores. Las revistas poco rigurosas desaparecen más fácilmente del mercado editorial científico y no sólo por razones económicas. La coexistencia de distintas revistas con ámbitos solapados puede verse entonces como el fundamento del mantenimiento de un equilibrio razonable en el nivel de exigencia.

Las revistas científicas son, así, más que un mero vehículo de información, de luchas de intereses (individuales y colectivos) o de posicionamiento en ciertas redes socio-cognitivas. Sea cual sea la interpretación que se haga de sus funciones, hay que tener en cuenta su carácter activo, su autonomía y la influencia que pueden ejercer. Las revistas pueden actuar como agentes en el proceso científico porque, además de canalizar y poner en común la información sobre la que se edifica el consenso en el conocimiento científico, pueden imponer ciertas restricciones y determinar en parte la forma y el contenido de lo que se publica. Pero también deben cuidar su propia permanencia. Incluso cuando se tengan garantizados los fondos suficientes para sostener los costes de edición, es preciso mantener un nivel de prestigio y de credibilidad. Por eso es necesario cultivar las propias ventajas competitivas. Los dos mecanismos básicos sobre los que se puede edificar la superioridad son el aumento de la calidad de los trabajos publicados y la especialización temática.

En primer lugar, dada la necesidad por definir su puesto de prestigio entre el resto de revistas y de asegurar su supervivencia, los directores de las revistas habrán de tomar las medidas que contribuyan a mantener una calidad elevada y una difusión amplia. Calidad y difusión no son ni reducibles entre sí ni enteramente independientes. Por un lado, una mayor calidad de los trabajos publicados no asegura por sí sola una vasta difusión, pero sí la favorece; por otro, los autores buscan la mayor diseminación de sus trabajos, por las razones ya mencionadas, con lo que preferirán las revistas más difundidas y éstas podrán elegir entre más trabajos. La calidad y la difusión refuerzan el prestigio de la revista; así aumenta la demanda por parte de los lectores y la motivación de los mejores autores para enviar sus trabajos, ya que publicar en mejores revistas dispensa un reconocimiento superior. La rivalidad entre las revistas tiene lugar entonces en los dos extremos que conecta, es decir, en la fuente y en el destino, por el emisor y por el receptor: se traduce en competir por los mejores autores y por el mayor número de lectores;

cuanto más cualificados estén estos últimos, tanto mejor, porque el prestigio de la revista se construirá en gran parte sobre el reconocimiento concedido por los autores que han sido previamente sus usuarios, sus lectores.

Pero (en segundo lugar) la búsqueda de la continuidad no implica necesariamente la concentración de todos los esfuerzos en mejorar la calidad y la difusión. Podría aplicarse aquí una analogía biológica: para sobrevivir, una especie no necesita ser la mejor adaptada, puede bastar con que encuentre su nicho ecológico, es decir, un ámbito donde sus posibles depredadores o competidores no estén presentes. De igual modo, las revistas pueden buscar su *nicho intelectual* mediante la especialización de sus contenidos, es decir, concentrando su atención en un tema determinado. La definición de un ámbito propio hace que la revista se dirija a una audiencia restringida, interesada particularmente en los temas tratados en ella y favorecida tanto por el ahorro de tiempo respecto a buscar artículos de su tema en revistas más generales como por las mayores garantías que ofrece. La especialización de la revista también aumenta, naturalmente, la capacidad de atraer a los autores de esa área porque éstos buscan el reconocimiento preferentemente entre los científicos de su especialidad, que son precisamente los que consultarán las revistas del área. La especialización genera así una ventaja competitiva para la revista, pero además se ve reforzada por los intereses de los dos polos que une. Es por eso por lo que la fragmentación de la ciencia en disciplinas está íntimamente relacionada con la disponibilidad de canales específicos: la consolidación de una especialidad emergente pasa muchas veces por el surgimiento de una revista representativa (Kara-Murza, 1986; Leydesdorff, Cozzens, van den Besselaar, 1993), el cual suele ser calurosamente recibido y apoyado por el conjunto de investigadores que empiezan a sentirse inscritos en la nueva especialidad.

De hecho, la inmensa mayoría de las revistas científicas muestran un fuerte grado de especialización, sobre todo disciplinar (Narin, Carpenter, Berlt, 1972), pero a veces también incluso en cuanto a métodos o técnicas. La organización disciplinar de la ciencia puede rastrearse en un cierto nivel por los dominios definidos por las revistas si éstas se interpretan como reflejo de la institucionalización de las áreas de conocimiento (Leydesdorff, Cozzens, 1992). La creación de nuevas revistas y los cambios de área o las escisiones de las ya existentes expresan la dinámica disciplinar de la ciencia no sólo como consecuencia observable de

esta dinámica, sino también como uno de los factores que la determinan. Las revistas científicas ofrecen por ello un medio privilegiado para intentar delimitar los campos temáticos en que se articula la ciencia (Leydesdorff, Cozzens, van den Besselaar, 1993). Las relaciones entre los distintos núcleos temáticos que se generan son mucho más complejas y dinámicas que lo que una clasificación al uso puede captar (Small, 1976; Leydesdorff, Cozzens, 1992; Leydesdorff, 1993); admitiendo esta limitación, sin embargo, es posible encontrar relaciones objetivas entre las revistas que ayuden a construir clasificaciones con utilidad práctica de las que no puede decirse que sean completamente arbitrarias. El análisis estadístico de los patrones de citación entre los artículos de distintas revistas ha sido empleado para establecer clasificaciones con una apreciable base objetiva, como las desarrolladas por el *Institute for Scientific Information*, que mantiene los conocidos *Citation Index* (Carpenter, Narin, 1973; 1980a).

El sistema de publicación de la ciencia se organiza de este modo en torno a los canales establecidos para la comunicación de los resultados científicos, las publicaciones periódicas institucionalizadas, las revistas científicas. Estos canales establecidos juegan un papel activo en este proceso, relacionado básicamente con la competencia que surge entre ellos y con las decisiones editoriales de sus responsables. Las condiciones para la existencia de las revistas y de los otros canales formales influyen en el proceso de selección y difusión de manuscritos de un modo capital. La tendencia hacia la calidad máxima de los manuscritos publicados se ve limitada por la disponibilidad de un número suficiente de ellos y, simétricamente, la de publicar todo lo que tenga algún interés encuentra un freno en la cantidad máxima digerible por la comunidad receptora y en el descenso de calidad que ello supone. La fragmentación en áreas progresivamente más especializados demarca "campos de batalla" bastante independientes entre sí, cada uno de los cuales está sometido a unas condiciones particulares; la tendencia hacia la maximización de las posibilidades de obtener reconocimiento hace que los investigadores se distribuyan entre todas las áreas disponibles (Hagstrom, 1965). Se alcanza así en cada área un equilibrio en cuanto al nivel exigido a los manuscritos, de lo cual cabe esperar que el esfuerzo requerido para producir un trabajo pueda variar de unas áreas a otras; esto puede influir, por tanto, en la productividad media (publicaciones por autor) de cada área. Además, pueden desarrollarse pautas formales características de esa

especialidad, que afectan, por ejemplo, al modo de exposición o a los patrones de citación.

Todo esto proporciona una explicación más que plausible del hecho cierto de la existencia de determinadas diferencias significativas en cuanto al comportamiento bibliográfico de las distintas áreas<sup>27</sup>. La formación y consolidación de criterios de publicación dispares puede provocar desemejanzas en la calidad o el contenido mínimos exigidos en cada área, así como en el uso de otros elementos sobre los que se asienta la construcción de indicadores bibliométricos (citas y vocabulario principalmente; por ejemplo, Garfield (1976) muestra que el número medio de citas que contienen los artículos presenta grandes diferencias entre áreas). Las comparaciones entre las áreas han de ser, entonces, necesariamente normalizadas. En lugar de comparar los valores brutos de cada área, deben concebirse indicadores normalizados que eliminen o minimicen los efectos de las diferencias debidas a las peculiaridades que puede producir el desarrollo independiente del sistema de publicación en cada disciplina o especialidad. Las especialidades se erigen así en los marcos de referencia adecuados para valorar los datos sobre productividad o calidad que se obtengan del análisis estadístico de la bibliografía científica (Leydesdorff, 1985).

Dentro de cada especialidad, las revistas están sometidas a unas similares condiciones básicas en cuanto a las fuerzas de oferta y de demanda de la comunidad de investigadores en esos temas. Los documentos pertenecientes a ese campo tampoco habrán de tener grandes diferencias formales entre sí. Es una idea generalizada que, en cada circunscripción temática, por así decirlo, existe un acuerdo básico acerca del nivel de prestigio relativo del que cada revista goza, derivado evidentemente de la calidad o interés que se esperan de los artículos que cada una publica. La expectación que cada revista genera se debe en gran medida a la trayectoria pasada de la revista. Esto, naturalmente, no excluye el que puedan aparecer artículos mediocres o poco relevantes en revistas prestigiosas, o que todo lo que aparezca publicado en una revista oscura sea poco apreciable; lo que significa es que es considerablemente mayor la probabilidad de encontrar buenos trabajos y menor la de hallar trabajos intrascendentes en revistas prestigiosas, mientras que sucede lo contrario en revistas poco valoradas, lo cual resulta casi obvio, porque el

<sup>27</sup> Las principales se refieren a la productividad y a los patrones de citación. En los capítulos dedicados a los indicadores de producción y de calidad se volverá sobre el tema.

prestigio de la revista se asienta sobre la relevancia de los trabajos publicados con anterioridad. Guiados por esta convicción, han sido numerosas las propuestas que, a partir de medidas aproximativas del nivel de prestigio de las revistas, han pretendido atribuir un cierto nivel de calidad a agregados de documentos científicos<sup>28</sup>. Estas medidas aproximativas se han estimado principalmente mediante la consulta a expertos y a través de diversos indicadores basados en las citas recibidas por las revistas.

Como se ha visto, la especialización temática de las revistas está reforzada desde varios frentes. Las principales consecuencias de este proceso son la de favorecer una profundización en los contenidos y el consiguiente aislamiento entre las disciplinas, producido por la dedicación más exclusiva a que conduce la mayor especialización. Así se afianza un modo de distribución de las contribuciones científicas muy selectivo en cuanto a la difusión, mediante el que sólo los investigadores especializados en un campo tienen noticias de los avances que se producen en éste. El problema que surge por la especialización excesiva puede encontrar respuesta en iniciativas que buscan la sinergia entre áreas. Las revistas, además de ocupar todas las áreas que cuenten con un *mercado* suficiente, también buscarán responder a los mercados emergentes, por lo que no es raro el nacimiento de revistas interdisciplinarias que a su vez pueden terminar siendo el germen de nuevas especialidades (Leydesdorff, Cozzens, van den Besselaar, 1993). Así, la tensión entre la fragmentación excesiva y la búsqueda de puentes hacia la unidad ha de ser visto como uno de los motores del dinamismo disciplinar de la ciencia.

El estudio de las revistas científicas, de sus relaciones entre sí y de sus influencias en el desarrollo de la ciencia, se enfrenta a un objeto complejo, lleno de matices y de aspectos interesantes, que debe ser abordado desde distintos puntos de vista. En este apartado se han tratado solamente las líneas fundamentales para la comprensión del uso más habitual del sistema de revistas científicas como fuente de indicadores bibliométricos, en particular en dos ámbitos de la mayor relevancia: el seguimiento de la estructura disciplinar y la detección de diferencias en la calidad estimada de los trabajos.

---

<sup>28</sup> En el capítulo 6 se comentan estos intentos.

El sistema de las revistas científicas posee otras muchas facetas dignas de atención, aunque no tan directamente relacionadas con los indicadores bibliométricos. Sin embargo, antes de concluir han de señalarse un par de cuestiones que aquí no deben dejarse completamente de lado. La primera atañe a la amplitud de la zona (geográfica, política, cultural) de influencia de las revistas científicas. A lo largo de la presente exposición acerca de las relaciones de competencia entre las revistas y sus consecuencias, ha podido sin duda recibirse la impresión de que la ciencia es una empresa que trasciende las fronteras que en otros planos separan a los seres humanos. Es verdad que la institución de la ciencia disfruta, en muchos aspectos, de una autonomía elevada respecto a los distintos poderes nacionales, y también lo es que los científicos pertenecen en cierto sentido a una comunidad mundial en la que importa poco el origen de cada uno. La idea de que la ciencia es universal, independiente de cualquier tipo de particularismos, ha sido constante desde su origen moderno. Esto es un hecho para una buena parte de la ciencia que se hace en el mundo, como lo atestigua la existencia de un extenso núcleo de revistas internacionales (más bien habría que llamarlas "anacionales") a las que tiene acceso virtualmente toda la comunidad científica. La gran difusión, el alto nivel de calidad y la naturaleza de los temas que tratan son las razones del interés general que despiertan estas revistas. Este conjunto de revistas es aproximadamente lo que se ha llamado *corriente principal* (*main stream*) de la ciencia (Carpenter, Narin, 1980b; Garfield, 1972). Es muy complicado estimar la proporción de la ciencia que se hace en el mundo que no aparece en esa corriente principal, especialmente porque los límites externos de la ciencia son borrosos y esto exigiría tomar decisiones controvertidas acerca de si algo es o no científico.

Pero, además, aunque la ciencia sea, en sus aspectos generales, de interés universal, no puede excluirse el que muchos de sus desarrollos y aplicaciones, sin dejar de ser científicos, tengan un carácter local o un interés muy particular. Esta es la razón de que puedan existir muchas revistas dignas de cierta consideración aunque no pertenezcan a esa corriente principal. No cabe, por tanto, despreciar sin más la producción científica que aparece en este tipo de revistas. La cuestión práctica está en determinar si la corriente principal proporciona una muestra satisfactoria del sistema científico que se quiera estudiar. Si la respuesta es negativa, habrá que tomar en consideración las revistas de carácter local que sean interesantes, sin olvidar que las diferencias de calidad entre

ellas pueden ser significativas. Algunas áreas, por la naturaleza de sus contenidos, están más afectadas por este problema que otras. Por otra parte, cabe esperar que las fuerzas de la competencia sean menos potentes en el nivel local que en el internacional, ya que este tipo de revistas suelen tener garantizado un sostén institucional. El peligro de que estas revistas caigan bajo la influencia de círculos de poder o se conviertan en cotos reservados es, por el contrario, superior.

Ésta es, precisamente, la segunda cuestión colateral: el peligro de que surja una estructura de poder que controle los canales de publicación científica. Una vez que existen canales consolidados (con credibilidad y tradición), éstos adquieren de hecho el carácter de oficiales, y resulta prácticamente imposible obtener reconocimiento a través de otros diferentes o no "oficiales". Esto hace, como se ve enseguida, que el control de esos canales establecidos tenga un valor estratégico de primer orden, debido a la enorme influencia que proporciona en el panorama científico. Con estas premisas, es fácil temer el riesgo de que se formen grupos que hagan valer sus intereses particulares en el proceso de publicación. Si existen redes de presión serán, como es de esperar, entidades ocultas, informales, difusas, puesto que sus componentes no reconocerían nunca formar parte de tales organizaciones manipuladoras. Pero es difícil imaginar que en la ciencia puedan sostenerse conspiraciones de ese calibre, estructuradas y estables.

En resumen, las revistas científicas son los más potentes estructuradores de la comunicación científica. Por su carácter activo y por las relaciones que se establecen entre ellas, desempeñan un papel primordial en el control de calidad de lo que se publica en los canales científicos y en la diferenciación de las disciplinas. Todo ello las convierte en un objeto de la mayor relevancia para el análisis bibliométrico: en primer lugar, delimitan el conjunto de los documentos a los que está justificado atender; en segundo lugar, proporcionan un criterio para la distribución temática y la definición de marcos de referencia adecuados; por último, permiten realizar estimaciones sobre la calidad de agregados de documentos.

### *3.2. La estructura del documento científico*

Los documentos científicos, aquellos escritos mediante los que tiene lugar la comunicación formal de los resultados alcanzados en una investigación, han adquirido una configuración estable, tras un proceso de evolución relativamente rápido. No cabe la menor duda acerca de la existencia de una forma literaria propia de la ciencia que, con muy ligeras variantes, es fielmente respetada por todos los autores que publican sus trabajos de modo oficial. De hecho, el ajustarse a esa forma es en la práctica un requisito para que sean publicados.

Los varios tipos de escrito científico comparten una misma estructura y estilo, aunque no todas las partes de esa estructura están siempre presentes en ellos. El tipo de documento científico por excelencia es el artículo científico, que representa un informe acabado sobre algún aspecto de una investigación, una pieza completa ofrecida para el debate o la consideración de los otros científicos.

El análisis bibliométrico de la literatura científica tiene como fundamento último el que la presencia del documento en el proceso científico no es algo contingente; pero depende de un modo más inmediato de la constancia de los rasgos de esa literatura que pueden tomarse como base de información estadísticamente significativa. Varias de las partes que constituyen la estructura del documento científico han recibido una considerable atención en los análisis bibliométricos, aunque frecuentemente con una justificación separada para cada una de ellas. Abordar un tratamiento más unitario de la configuración de esos documentos puede servir para explotar mejor las posibilidades de análisis que ofrecen a los estudiosos de la ciencia. Éste es el propósito del trabajo de Mullins, Snizek y Oheler (1988):

"Creemos que es posible caracterizar la tarea intelectual de un área en un cierto momento como un compuesto de las formas tomadas por las varias partes de los escritos científicos. Argüimos que el análisis sociológico del documento científico puede revelar al menos una parte del cuadro acerca de qué cosas se necesitan para hacer trabajo científico." (Mullins, Snizek, Oheler, 1988: 82)

Este es el punto de partida para un ágil recorrido en el que se comentan las distintas partes del documento científico en clave histórica y funcional, dedicando especial atención a las posibilidades interpretativas de cada sección. Atienden a una muestra de esas partes, que categorizan del

siguiente modo: marcas de identificación (que incluyen título, lista de autores, afiliación institucional, agradecimientos, palabras clave y resumen); cuerpo del texto (se ocupan de los siguientes aspectos: formato, estilo discursivo, y uso léxico); ayudas visuales (ilustraciones, gráficos, ecuaciones...). Las referencias, que han sido objeto de numerosos estudios, no son tratadas en ese trabajo.

Siguiendo aproximadamente el esquema propuesto por estos autores, a continuación se intentará complementar su aportación desde un punto de vista ligeramente diferente. Todas las partes del documento tienen una función informacional más o menos explícita, pero habitualmente clara; aquí se insistirá más en la función *retórica* que cada una puede desempeñar. Los usos retóricos del lenguaje científico han sido tratados desde distintos puntos de vista. Ha sido frecuente señalar su contraste con la imagen tradicional de la ciencia, como algo máximamente objetivo, alejado de intereses particulares y comprometido con la precisión, con la contrastabilidad o la sobriedad expresiva libre de valoraciones. Formas concretas del estilo científico han sido consideradas como figuras retóricas que esconden más de lo que muestran. Aquí se intentará adoptar una postura menos comprometida. Se asumirá (dejando de lado sentidos más recientes relacionados con el engaño o la petulancia) la definición tradicional, que la entiende como el "arte de convencer por medios no conclusivos". El documento científico se verá aquí como un medio especialmente adecuado para resaltar el interés de los *resultados científicos* de los que informa. A continuación se comentarán las diferentes partes que conforman la estructura típica de estos documentos, insistiendo en el modo en que cada una puede ser empleada por los autores para *persuadir* de que sus resultados merecen ser atendidos.

### *Título*

El título es una frase que encabeza y presenta el trabajo mediante una descripción breve de su contenido. Era ya habitual en los primeros documentos científicos modernos, a finales del s. xvii. Se espera que sea una especie de resumen que sirva para que el posible lector pueda determinar si le interesa o no ese documento: sitúa el trabajo en un foco temático, ayuda a clasificarlo de algún modo y mejora la eficiencia de la comunicación científica general, evitando pérdidas de tiempo a los investigadores ajenos al tema y facilitando la localización a los

investigadores relacionados con él. La elección de un título adecuado es algo que depende de las habilidades comunicativas de los autores y que éstos no siempre llevan a cabo con completa fortuna, por lo que los editores de las revistas pueden sugerir cambios con la finalidad de resaltar lo más significativo del trabajo. El título es un medio poderoso del que disponen los autores para destacar lo *nuevo* y lo *relevante* que aportan. Así pues, además de la función informativa y de optimización del flujo comunicativo en la ciencia, los autores pueden a través de sus títulos llamar la atención sobre el núcleo de sus pretensiones de reconocimiento.

### *Lista de Autores*

La mención explícita de autoría, algo casi inherente a la ciencia actual, fue sin embargo una costumbre que comenzó a consolidarse sólo desde principios del s. XVIII; anteriormente, a menudo se escribía y publicaba anónimamente. Marca una tensión entre dos polos en la ciencia moderna, el individual y el comunitario: la francesa *Academie Royale des Sciences* no permitía que aparecieran los nombres en los trabajos de sus miembros por concebir que la empresa científica sería más productiva sobre una base cooperativa y común. En el esquema mertoniano, la autoría es el signo de la propiedad intelectual del investigador y permite la reivindicación de prioridad, a cambio de poner sus resultados a disposición de la comunidad científica. La lista de autores establece una relación en dos sentidos entre los autores y su aportación: por un lado, los autores se hacen responsables, avalan en suma, el trabajo que presentan; por el otro, buscan reforzar su consideración como científicos a través de la importancia de ese trabajo. Es la credibilidad de autores y trabajo lo que está en juego, y por lo tanto, también el prestigio de los autores. Por otro lado, la publicidad de la autoría convierte al documento en una fuente contrastable y seria.

Para los destinatarios del trabajo, la lista de autores es una marca muy significativa para determinar su interés por el documento: la lista de nombres puede ser, en especial con investigadores afamados, un indicativo del calibre del trabajo. Este refuerzo de la atención hacia quienes están más reconocidos está en la base del *efecto Mateo*. Una consecuencia de esto es que, para los investigadores primerizos o poco conocidos, es claramente beneficioso ser co-firmante de un trabajo con un científico de elevada consideración, porque aumentan sus posibilidades de atraer la atención y de obtener reconocimiento<sup>29</sup>. La firma del trabajo puede

convertirse así en un instrumento para amplificar el reconocimiento correspondiente al trabajo en sí.

Un rasgo señalado repetidamente es el ascenso constante del número medio de firmas por documento (Price, 1963), que refleja la creciente importancia tanto del trabajo en equipo y de otros tipos de colaboración científica como la de criterios relacionados con el número de publicaciones en los procesos académicos y laborales de evaluación y selección de los investigadores.

#### *Afiliación institucional de los autores*

Hacer constar el lugar de trabajo de los autores tuvo inicialmente más que ver con la función de localización para la correspondencia generada a propósito del trabajo (Cole, Cole, 1968; Allison, Stewart, 1974). Es una consecuencia de la creciente profesionalización de la ciencia desde el final del s. XIX. En la actualidad figuran habitualmente las instituciones de todos los firmantes. Las instituciones se han convertido en entes relevante en ciencia, sometidos a unos mecanismos de reconocimiento similares a los de los autores. Por eso es una obligación del autor respecto a su institución el que ésta aparezca explícitamente mencionada en sus trabajos.

La relación de dependencia entre autor e institución en cuanto al prestigio no es fácil de establecer, aunque parece darse en los dos sentidos. Trabajar en una institución científica estable es un indicativo de una dedicación profesional a la ciencia y de haber superado ciertos controles. Para los autores poco conocidos, pertenecer a una institución prestigiosa supone un aval a los ojos de los posibles lectores. La mención de la institución de trabajo puede ser, así, un instrumento de convicción al servicio del autor, mediante el que se refuerza o se insiste en la categoría científica del investigador. Studer y Chubin (1980) encontraron que la institución podía ser un identificador más determinante que el propio autor en la recepción inmediata del trabajo.

---

<sup>29</sup> También puede ser beneficioso para el investigador de categoría superior, ya que esta práctica puede conducir a la paradójica situación en la que el responsable de un equipo o laboratorio firme sistemáticamente todos los trabajos de sus subordinados, con lo que se potencia aún más el sesgo hacia el investigador prestigioso. En todo caso, es él quien arriesga su prestigio si el trabajo no es bueno.

### *Resumen*

Facilita una comunicación rápida de los puntos esenciales del contenido del documento. Suele describir el problema y enumerar los resultados más destacados que se reivindican en el texto principal. Pueden atribuírsele las funciones del título, pero llevadas a cabo de un modo complementario, una vez que el título ha captado el interés del posible lector. Puede insistir en la novedad y en la relevancia del trabajo de un modo más preciso que el título, pero además puede referirse a aspectos relacionados con las técnicas empleadas o el rigor en la ejecución de la investigación.

El resumen es algo habitual en los documentos científicos, aunque no siempre está presente en el documento original, sino que puede ser añadido por los servicios de indización al archivar la referencia correspondiente. Debido al aumento en la difusión que puede proporcionar, es cada vez más frecuente que revistas publicadas en idiomas distintos del inglés incluyan resúmenes en este idioma para facilitar su divulgación y su inclusión en bases de datos internacionales. Esto hace que esta parte del documento se convierta en una base común para acceder a trabajos escritos en distintas lenguas.

### *Palabras clave*

Pueden ser libres o de vocabularios controlado; en este segundo caso son más estables y menos sensibles a los cambios, y pueden interferir en la descripción de la dinámica real. Su objetivo es destacar los puntos por los que el trabajo se conecta con la investigación de su disciplina. Lo más usual es que sean otorgadas por indizadores profesionales, a veces con asistencia informática. También pueden ser especificadas por los propios autores. Consisten en una serie de descriptores, simples o compuestos, cuyo número recomendado oscila entre ocho y quince.

### *Texto principal*

Es el cuerpo o núcleo esencial del escrito científico, el portador del contenido. Las otras partes aparecen como meros descriptores que informan acerca de él. La mayor parte de la atención dedicada al análisis del documento y discurso científicos ha sido dirigida precisamente hacia el texto, muchas veces tomándolo como el documento científico en sí. Ha sido tratado desde muy diversos enfoques y puntos de vista. Filósofos, sociólogos, psicólogos e historiadores se han ocupado del contenido y

de la forma del texto científico, así como de su relación con otros elementos del proceso de investigación.

El texto científico se presenta bajo la forma de relato fidedigno sobre los resultados significativos de una investigación y el modo en que se ha llegado a ellos. Habitualmente se han destacado como rasgos particulares de este discurso científico la adopción de formas argumentativas racionales, lógicas, probatorias, que apelan a evidencias y con un carácter marcadamente objetivo; también se ha resaltado su elaborada codificación y su contraste con otras clases de discurso.

Su función manifiesta es la de transmitir una determinada clase de información de modo preciso. Sin embargo, otros aspectos son los que han captado el interés de sus analistas, de modo que el texto científico ha sido interpretado como la mera apariencia o reflejo de otras funciones o procesos ocultos. Así, por ejemplo, se ha estudiado como la vía hacia el significado científico (Nagel, 1961), como resultado de patrones de intereses (Barnes, 1977, MacKenzie, 1981), como superficie de argumentos (Steed y Mullins, 1986), o como parte del proceso de construcción de la verdad (Laudan, 1977).

Son característicos del escrito científico su separación diáfana entre los datos y las interpretaciones, el planteamiento claro de una hipótesis o problema y la vinculación explícita con precedentes (esto último queda resumido en la lista de referencias). La sobriedad en la expresión, evitando los giros o figuras que no sean estrictamente necesarios, puede considerarse un ideal perseguido en estos documentos. El estilo está marcado por las oraciones pasivas, impersonales y enunciativas en presente. Precisamente esta abundancia de elementos que refuerzan la impresión de objetividad y neutralidad ha sido interpretado como un recurso retórico (Latour y Woolgar, 1979), dirigido a persuadir en último término de que el autor es creíble, es decir, a convencer. Knorr-Cetina (1981) describe detalladamente cómo se transforman y reorganizan los elementos de la investigación al componer el escrito, especialmente recontextualizando el problema y dotando a la narración de un esquema "solución a un problema" aunque durante el desarrollo concreto la contingencia y la elección oportunista hayan sido los factores dominantes.

Gusfield (1976) señala la estructura dramática del texto (tensión y resolución, identificación del bien y del mal, desarrollo de un curso de acción) que conforman las distintas secciones, aunque Knorr-Cetina

(1981) concentra ese planteamiento en la *Introducción* y considera que las otras secciones juegan el papel de apéndices a este respecto. La cuestión de si las dedicadas a los métodos y materiales se dirigen a posibilitar la repetibilidad o tienen una finalidad retórica también ha sido tratada (Gusfield, 1976; Mulkay, 1976). El texto científico también ha sido descrito como un medio de posicionamiento en redes socio-cognitivas (Callon et al. 1983; Law, 1986; Rip, 1986) es decir, de configuración estratégica del debate científico.

Las secciones en que se divide el texto de un documento científico son, con variaciones anecdóticas, las siguientes: *Introducción*, *Métodos y materiales*, *Datos*, *Resultados*, *Discusión* y *Conclusiones*. A veces pueden presentarse bajo el mismo epígrafe secciones contiguas, pero la pertenencia de cada párrafo a una u otra es clara. Esta estructura, al margen de las anteriores consideraciones sobre el estilo, la fidelidad de la narración o la preponderancia de una u otra finalidad (informar o convencer, por ejemplo) puede conectarse fácilmente con la búsqueda, por parte de los autores, de reconocimiento en los tres sentidos que se han distinguido en este trabajo. De lo que quiere *persuadir* el autor es de que merece ese reconocimiento. Todos los tipos de reconocimiento descansan en la obtención de resultados científicos. Para justificar su pretensión, el autor ha de resaltar que lo que ha conseguido es *científico*, *nuevo* y *relevante*. En cada sección se enfatiza alguno de estos aspectos; en la *Introducción*, todos ellos. Ésta contextualiza el problema construyendo una red de conexiones con la literatura precedente (concentra la mayor parte de las referencias) y enumerando las consecuencias positivas de una solución: destaca así la *relevancia* del trabajo. Al mismo tiempo define el estado previo de la cuestión, lo que sirve tanto como fondo de contraste sobre el que situar la *novedad* como para exhibir el dominio del autor en ese campo, lo cual es una condición de su *competencia* como científico y una señal de su inclusión en la comunidad científica.

Las secciones dedicadas a los *Métodos y materiales* y a los *Datos* son el lugar para defender la *cientificidad* de los resultados. Se describen los procedimientos y las bases empíricas de tal modo que queda clara la adecuación a métodos aceptados, intentando mostrar el rigor en la aplicación y la objetividad, es decir, la independencia respecto al autor. En esencia se trata de argumentar que se han tomado todas las *precauciones* necesarias para excluir otros factores ajenos a los considerados. El carácter científico del resultado se marca, pues, mediante la aplicación de

técnicas que garanticen la objetividad y el aislamiento de los posibles factores explicativos.

En *Resultados* se enumera expresamente aquello que funda la *novedad* de la aportación. La sección de *Discusión* puede tratar pormenorizadamente las justificaciones acerca de la importancia de cada factor y valorar cuál puede ser la explicación más plausible, dotando de sentido a los resultados. De aquí se derivan las *Conclusiones*, en donde se resumen los puntos principales insistiendo en la *novedad* y aludiendo a la *relevancia* de la aportación al conectarlo con el problema planteado al principio, defender que se trata de una alternativa preferible o tratar las consecuencias positivas o ventajas que conlleva.

El que el texto científico pueda ser un instrumento para la reivindicación del autor no excluye, sino todo lo contrario, otras funciones como la de informar o convencer. Éstas son también condiciones ineludibles en la estrategia por la obtención de reconocimiento, pues son el modo de defender su competencia científica, su capacidad para obtener cosas nuevas y relevantes mediante la aplicación de métodos científicos. Al informar, expone a la crítica y elimina la posibilidad de sospechas por ocultación; evidentemente también es una condición decir *qué* resultados son los que considera nuevos y relevantes. Si consigue convencer, alcanza el grado más alto de justificación sobre el valor de la aportación, y en éste se incluye el que sea científica, nueva y relevante.

#### *Gráficos, cuadros, tablas, ecuaciones, etc.*

Forman parte de la codificación y formalización del lenguaje científico. A través de ellos se muestran de forma ordenada los datos más significativos, resaltando o abstrayendo determinados rasgos. No todo el conocimiento científico se transmite en forma proposicional, mediante lenguajes naturales o formales. También se transmiten imágenes y representaciones visuales, que pueden ser decisivas en cuanto a la formación de paradigmas, por ejemplo. Este tipo de representaciones constituye una base común que permite un eficaz intercambio de información entre científicos aunque hablen lenguas diversas. Representan los elementos más notorios en la constitución de un lenguaje especial de la ciencia.

### *Lista de Referencias*

Contiene la lista de todos los trabajos que han sido citados en el texto, enumerados bajo el formato abreviado establecido (referencia bibliográfica) que posibilita la identificación y localización del correspondiente trabajo. La lista de referencias es el elemento que refleja con mayor claridad la inserción del trabajo científico individual en una obra colectiva. Su doble carácter, cognitivo y social, ha dado lugar a un aluvión de estudios sobre sus funciones y su utilidad para el estudio de la ciencia a través de las conexiones que establece entre trabajos y autores<sup>30</sup>.

La costumbre académica clásica de citar obras anteriores parece estar relacionada con la invocación a las autoridades reconocidas, como recurso retórico o erudito con el que mostrar conocimiento e inducir al asentimiento, o con la exposición abreviada de ciertos contenidos, remitiendo a las obras que ya han tratado un aspecto con detalle suficiente, lo cual evita repeticiones innecesarias y agiliza la presentación (todo lo cual no sería considerado muy ajeno a la ciencia actual por parte de algunos autores). Price (1963: 112) señala que aunque ya existían los *scholía* medievales como claro precedente, el uso de las citas con una conciencia de acumulación del conocimiento es un fenómeno relativamente reciente, propio del afianzamiento de la ciencia moderna, y Merton que se ha consolidado en su forma actual en los documentos científicos sólo desde el del S XIX.

La cuestión de cómo pueden interpretarse las citas ha sido abordada desde varios puntos de vista por numerosos trabajos, que alternativamente se han ocupado de las funciones de las citas, de su significado cognitivo o de las motivaciones de los autores citantes (a veces con una confusión de estos aspectos). Ziman, dentro de su concepción del documento científico como una contribución al consenso público que es la ciencia, señala:

"Mencionar las referencias da validez a muchas de las aseveraciones que él hará en su artículo, y las entronca al consenso preexistente. El orden de este proceso, la estructura intelectual implícita en la biblioteca, el catálogo, el índice, la enciclopedia, el tratado, da sentido a la investigación del pasado y motivo a la investigación del futuro. La simple

---

<sup>30</sup> Una visión general se encuentra en Egghe y Rousseau (1990a). Luukkonen (1990) dedica su tesis doctoral a los problemas del análisis de citas en la evaluación de la investigación.

acumulación de detalles heterogéneos no basta para aportar ese orden y significado " (Ziman, 1968: 135)

Ziman entiende que las referencias articulan las contribuciones individuales en un corpus de conocimiento consensuado y común. Por otro lado, la escuela mertoniana concibe el acto de citar como una parte fundamental del sistema de recompensas de la comunidad científica: ser citado es un signo de reconocimiento de la adecuación al valor institucional de la originalidad, una especie de honor que se recibe como pago a la contribución donada desinteresadamente, y que funciona como incentivo reforzador de esta práctica (Merton, 1957; Hagstrom, 1965; Cole, Cole, 1967). Garfield (1972) sugiere que la cita a un trabajo anterior indica que éste ha sido de utilidad para la elaboración del trabajo citante.

Weinstock (1971) identifica quince funciones específicas de las referencias, entre ellas: homenaje a pioneros; reconocer trabajo relacionado; identificar metodología o equipos; proveer lectura de fondo; corregir el trabajo propio o el de otros; criticar otros trabajos; fundamentar reivindicaciones; avisar de trabajos de próxima publicación; difundir trabajos de circulación escasa; identificar trabajos originales o fuentes de ideas; negar las reivindicaciones de otros; disputar prioridades. Thorne (1977) ofrece otra lista, el reverso de la anterior, en la que denuncia las malas prácticas en el uso de referencias: dividir resultados en "unidades mínimas publicables" para ser más citado; excesivo detalle y elaboración de las citas; citar a la moda para obtener financiación, o según el gusto de los editores, por presiones de poder, por conspiración conjunta; citas clásicas a figuras eminentes, no reconocimiento de autores nuevos, citas obsoletas...

Los autores que han llevado a cabo análisis de contexto y contenido de las citas han propuesto que el reconocimiento de trabajos precedentes es rara vez la función preponderante de las citas (Moravcsik, Murugesan, 1975; Chubin, Moitra, 1975; Frost, 1979) y han desarrollado varios esquemas clasificatorios de las citas. Moravcsik y Murugesan, por ejemplo, las categorizan bajo el siguiente, en el que la cita responde a uno de los polos alternativos en una o más dimensiones: a) conceptuales / metodológicas; b) orgánicas / perfunctoras (e.d. esenciales o accesorias); c) evolutivas / yuxtaposicionales (e.d. de apoyo sobre el precedente o de alternativa); d) confirmadoras / negadoras. Además insisten en la

posible redundancia entre las referencias de un trabajo. Estos estudios, sin embargo, han tenido escasa influencia en los estudios bibliométricos de evaluación (Luukkonen, 1990) por sus limitadas muestras, diferentes esquemas y porque la existencia de varios significados cognitivos no invalida el que se cite a obras por considerarlas importantes en algún sentido (Zuckerman, 1987).

Las citas también han sido consideradas como un instrumento del proceso de persuasión por el que el autor intenta convencer al lector de la validez de sus argumentos, de la corrección de sus métodos y resultados (Gilbert, 1977; MacRoberts, MacRoberts, 1986a, 1986b). En la misma línea, Latour y Woolgar (1979) consideran que las citas son esgrimidas por el autor para fundar su propia credibilidad, conectándose con trabajos o autores que considera fiables. Knorr-Cetina (1981: 125) describe el doble uso de las citas en la búsqueda de aliados por la similitud y en la diferenciación respecto a los precedentes, es decir, en marcar el contraste, "*para establecer su trabajo como relevantemente nuevo*".

Small (1978) sugiere que las citas actúan como símbolos de conceptos (*concept symbols*): los autores asocian determinadas ideas con documentos particulares y citan éstos para referirse a esas ideas en sus propios textos. Esto está muy relacionado con la utilidad que prestan las referencias como medio para abreviar la exposición, lo cual implica una forma particular de codificación del lenguaje científico.

En los estudios de Brooks (1984, 1985), los científicos señalaron, entre siete motivos para citar, la persuasión, y el crédito positivo como los principales, mientras que el consenso social y el crédito negativo obtenían puntuaciones muy bajas. Vinkler (1987a) propone que existe un umbral mínimo que es necesario superar para que la "presión de la fuerza cognitiva" haga citar un trabajo determinado. Este umbral dependería de la relevancia profesional de ese trabajo respecto al estudio del autor citante (lo cual no aclara mucho).

Cozzens (1988) lleva a cabo un significativo intento de integración de perspectivas acerca del papel de las citas. Su propuesta es que en la ciencia, existen tres sistemas: el *retórico* o conceptual, que tiene que ver con la relevancia, la utilidad o la influencia, establece los criterios primarios; el de *distribución de recompensas*, relacionado con la calidad o la importancia, pone los criterios secundarios; y el de *comunicación*, en el que intervienen los factores concretos de publicación, como idioma,

cultura, país o área geopolítica, crea las condiciones. Todos ellos intervienen en diferentes planos, reforzándose o interfiriendo entre sí, en el proceso efectivo de citar en la ciencia. El impacto y la visibilidad, por ejemplo, dependerían de las interacciones entre ellos.

Leydesdorff y Amsterdamska consideran que si los textos científicos se tratan como intentos de integrar reivindicaciones de nuevo conocimiento en el ya compartido, las citas cumplen varias funciones argumentativas dentro de los textos citantes; éstos pueden modificar las reivindicaciones de los documentos citados y usarlas en sus propias reivindicaciones. Insisten en la multidimensionalidad de la motivación para citar incluso al mismo trabajo y a la misma frase de ese trabajo (Amsterdamska, Leydesdorff, 1988; Leydesdorff, Amsterdamska, 1990).

La impresión que produce este conjunto de aproximaciones y enfoques sobre el papel de las citas es que se trata de un problema parecido al que plantea la interpretación de un instrumento: ¿para qué sirve? (qué conjunto de cosas pueden hacerse con él, cuál es su *espacio de uso*); ¿para qué lo utiliza? (qué intención tiene el usuario) ¿cómo lo utiliza? (el modo en que lo usa); ¿qué efectos consigue? (qué consecuencias, queridas o no, directas e indirectas tiene su acción). Además, aquí está la cuestión del contexto: unos buscan el papel de las citas *en la ciencia*, otros se refieren al *documento*. Cada autor ha optado implícitamente por una de esas preguntas como la pertinente para explicar las citas. La mayor dificultad reside en que, incluso cuando se dé la concordancia entre todas estas dimensiones (es decir, que alguien quiera conseguir algo al citar, emplee la cita de modo correcto para conseguirlo y lo consiga efectivamente), no se pueden desdeñar el resto de usos o de consecuencias indirectas.

Propongo aquí que la mayor parte de las funciones o usos que se han señalado para las citas puede subsumirse en la búsqueda de reconocimiento por parte del autor citante. Al igual que las secciones del texto, las citas pueden servir al autor que cita para justificar su pretensión de reconocimiento científico de tres modos: a) delimitando la *novedad* de su contribución, al enumerar los trabajos previos que configuran el estado de la cuestión y sobre los que se establece el contraste del suyo cuando marca las diferencias; b) señalando la *relevancia* de su trabajo, mediante el entronque con problemas abiertos o pendientes; c) mostrando el *dominio* del autor sobre el contexto disciplinar, su integración como científico competente, al identificar las obras que se consideran

básicas y definidoras del consenso en esa disciplina, las que pueden tomarse como modelos ejemplares e incluso al citar sus propias obras, como prueba de su experiencia. Cualquier cita puede servir a uno o a varios de estos objetivos, que son las condiciones básicas para la obtención de reconocimiento. Esto no significa que todos los autores sean igualmente habilidosos y precisos en su uso de las citas, ni que sepan explotar las posibilidades de este recurso<sup>31</sup>; de hecho, la lista de referencias puede indicar que el autor no merece reconocimiento. Las citas accesorias, por ejemplo, no favorecen sus pretensiones, sino todo lo contrario.

Cuando un trabajo recibe un buen número de citas puede suponerse con muy poco riesgo que se debe a que ha servido a los autores que lo han citado para cumplir alguno de los anteriores propósitos. Cualquiera de ellos hace que ese trabajo sea científicamente significativo (y tiene por tanto la calidad necesaria para serlo). No significa que esté en lo cierto, ni que lo acepte quien lo cita, pero sí que le ha servido para marcar su novedad, integrar su trabajo o acreditar su dominio del tema<sup>32</sup>.

No es indispensable, por lo tanto, el que exista por parte del autor citante una intención o una motivación para valorar públicamente las contribuciones de los predecesores o colegas. Lo que se ha llamado *sistema de recompensas* puede explicarse así como una consecuencia indirecta y necesaria del mecanismo egoísta de búsqueda de reconocimiento por parte de los individuos: éstos no tienen modo de fundar su propio reconocimiento científico sin reconocer a su vez a quienes lo merecen, es decir, a quienes han tenido un papel significativo en el debate científico. Y este proceso tiene que ver con los contenidos de la ciencia,

<sup>31</sup> Polanyi (1966) considera que citar no forma parte estrictamente del proceso intelectual, sino que es una habilidad como nadar o montar en bicicleta.

<sup>32</sup> Este planteamiento permite integrar las autocitas y las citas negativas, que en otras aproximaciones se consideran *prácticas desviadas* o se aducen como objeción para el empleo de análisis de citas en evaluación. Otras objeciones son la existencia de citas atribuidas a la dependencia respecto a grupos establecidos de poder y la de que la disponibilidad o la circulación del original determinan la posibilidad de ser citado, es decir, que las citas sólo reflejan la *visibilidad* de los trabajos citados. Respecto a la primera, falta la demostración de que alcanzar una posición de poder en la ciencia pueda ser completamente ajeno a la competencia científica demostrada. La segunda oculta el reproche de que seguramente existen trabajos excelentes perdidos en humildes anaqueles de oscuras bibliotecas o en revistas poco difundidas. Pero la visibilidad no es algo independiente de su importancia para la ciencia.

independientemente de si existe o no una norma de cortesía que impulse a reconocer a los otros, una obligación (funcional para la empresa colectiva) de compensar la deuda cognoscitiva o una pauta que premia la adecuación a valores, como la *originalidad*, sostenidos por la comunidad científica. La reclamación de reconocimiento científico genera automáticamente la distribución de reconocimiento científico, porque el marco de la ciencia establecida está constituido por aportaciones previas.

### 3.3. *Las referencias bibliográficas en el análisis bibliométrico*

Las listas informatizadas de referencias bibliográficas, creadas con la finalidad principal de facilitar a los científicos la identificación de la literatura que han de tener en cuenta para su investigación, ponen a disposición de los análisis bibliométricos la información relativa a grandes cantidades de documentos científicos. En realidad, lo habitual es que estos análisis se lleven a cabo sobre los datos de los documentos proporcionados por estas bases informatizadas. No necesita una mayor justificación, por tanto, la necesidad de ocuparse de las posibilidades que abren a la bibliometría las referencias bibliográficas contenidas en esas bases de datos.

La referencia bibliográfica guarda, como es de esperar de una representación, un estrecho paralelo con la estructura del documento científico. En general, pretende ser un resumen que facilite la localización del documento original. A veces informa además del contenido del trabajo de un modo breve, para ayudar al lector a decidir si le interesa o no. Una buena referencia bibliográfica proporciona una información muy valiosa, que permite realizar análisis estadísticos ricos y variados. De hecho, contiene información no estrictamente perteneciente al documento científico, sino que lo caracteriza: fuente en que apareció, año de publicación, palabras clave, etc. Por eso es casi siempre mejor base para la bibliometría que el documento original completo.

Las distintas bases de datos que indizan los documentos primarios pueden presentar formatos propios pero, en lo esencial, difieren muy poco de uno a otra base de datos. Una de las más importantes es la base del *Science Citation Index* (SCI), creada en la década de los 60 por Eugene Garfield. En la actualidad, esta base compila cada año las referencias

bibliográficas correspondientes a cerca de seiscientos mil documentos publicados en unas tres mil quinientas revistas científicas de circulación internacional. Esta base, de carácter multidisciplinar, es la más empleada en estudios bibliométricos. Desde 1992, está disponible también la versión con *Abstracts* (resúmenes de los artículos). Es, además, la única de su género que incluye las referencias citadas por cada trabajo, y una de las pocas que listan los nombres de todos los autores del trabajo (siempre que no superen los 256, algo poco frecuente). Tomaremos el modelo de referencia completa SCI como ejemplo del máximo de información analizable en una referencia y realizaremos sobre ella las consideraciones acerca de las posibilidades de análisis que ofrece cada campo.

---



---

### Campos de la referencia bibliográfica SCI

---

<b>AU</b>	Lista de los autores del trabajo (hasta 256) (AAuthor/s)
<b>TI</b>	Título del trabajo (Title)
<b>JN</b>	Nombre de la revista o monografía; número; pgs. inicial y final (JourNal)
<b>PY</b>	Año de publicación del número (Publication Year)
<b>GA</b>	Número de acceso (Genuine Article)
<b>DT</b>	Tipo de documento (Document Type)
<b>LA</b>	Idioma de publicación (LAngeage)
<b>RF</b>	Número de referencias citadas (ReFerences)
<b>CS</b>	Lista de las instituciones de trabajo de los autores (Corporate Source)
<b>AB</b>	Resumen del trabajo (ABstract)
<b>DE</b>	Palabras clave indicadas por el autor
<b>ID</b>	Palabras clave asignadas por el indizador
<b>CR</b>	Lista codificada de referencias a otros trabajos

---



---

A continuación se muestra un ejemplo de referencia bibliográfica en el SCI (versión CD-ROM *with abstracts*). Este es el número máximo de campos que puede contener, en otras fichas puede faltar alguno.

\*\*\*\*/ISI-DIALOG J

AU Yepes G; Domingueztenreiro R; Couchman HMP

TI The Scaling Analysis as a Tool to Compare N-Body Simulations with Observations - Application to a Low-Bias Cold Dark Matter Model

JN ASTROPHYSICAL JOURNAL 401(1):40-48

PY 1992

GA KB243

DT Article

LA English

RF 59

- CS CANADIAN INST THEORET ASTROPHYS, MCLENNAN LABS, 60 ST GEORGE ST, TORONTO M5S 1A7, ONTARIO, CANADA;  
UNIV AUTONOMA MADRID, DEPT FIS TEOR C XI, E 28049 MADRID, SPAIN;  
UNIV WESTERN ONTARIO, DEPT ASTRON, LONDON N6A 3K7, ONTARIO, CANADA
- AB We show that the scaling or multifractal analysis is specially suited to quantify the clustering properties of galaxy distributions. The information provided by this method of volume-limited samples is not masked by the boundary corrections made.  
Consequently, the formalism could be a useful tool to compare the outputs of N-body simulations and galaxy catalogs.  
The multifractal statistical analysis has been applied to the output of a highly evolved, low bias ( $b(t)$  congruent to 0.8),  $\Omega = 1$  cold dark matter simulation with 128(3) particles in a 100  $h^{-1}$  Mpc cubic volume. The identification of galaxy tracers in this simulation is carried out by means of an algorithm that mimics the formation and merging of galaxies. The same statistical analysis has been performed in several samples of the 14.5 CfA galaxy catalog. From the comparison of both analyses we conclude that there exists a very good agreement in the overall clustering properties of the observational samples and the simulation, as described by the statistical method.
- DE Author keywords: Dark Matter; Galaxies, Clustering; Large-Scale Structure of Universe; Methods, Numerical
- ID KeyWords Plus: MULTIFRACTAL MEASURES; GALAXY REDSHIFTS; UNIVERSE; VELOCITY; CATALOG; SUPERCLUSTER; SEGREGATION; ATTRACTORS; 2-POINT
- CR BARROW-JD-1985-MON-NOT-R-ASTRON-SOC-V216-P17  
BERTSCHINGER-E-1990-ASTROPHYS-J-V364-P370  
BLANCHARD-A-1988-ASTRON-ASTROPHYS-V203-PL1  
BOND-JR-1987-2ND-P-CAN-C-GEN-REL-P385  
BOND-JR-1991-PHYS-REV-LETT-V66-P2179  
BONOMETTO-SA-1980-A-A-V92-P222  
BOUCHET-FR-1986-ASTROPHYS-J-V302-PL37  
[...]  
SAUNDERS-W-1991-NATURE-V349-P32  
SCHAEFFER-R-1984-ASTRON-ASTROPHYS-V134-PL15  
VALDARNINI-R-1991-MON-NOT-R-ASTRON-SOC-V251-P575  
VALDARNINI-R-1992-APJ-V394-P422  
VITTORIO-N-1991-ASTROPHYS-J-V372-PL1  
WHITE-SDM-1979-MON-NOT-R-ASTRON-SOC-V186-P145  
WRIGHT-EL-1992-APJ-V396-PL13

La referencia bibliográfica, como representación del documento científico, es la principal fuente de datos para los análisis bibliométricos. A continuación veremos una descripción de cada campo de la referencia bibliográfica, junto con una exploración de las posibilidades de análisis que ofrece cada uno de esos campos.

**Autor(es) (AU).** Identifica a los investigadores firmantes del escrito mediante su apellido e iniciales. En el caso de autores españoles, cuando indican dos apellidos aparecen fundidos en uno solo. La ley de distribución de la productividad individual que encontró Lotka (1929) se basó en este dato, que abre un rico abanico de posibilidades. El estudio de las relaciones de colaboración entre individuos y sus consecuencias ha producido numerosos trabajos, que se han centrado, por ejemplo, en la estratificación social entre los investigadores (Zuckerman, 1968;

Kretschmer, 1990), en los modos de producción en equipo (Vinkler, 1992) o en la detección de equipos de investigación (Vidal, Villarroel, 1995).

También permite explorar las diferencias entre los trabajos con diferente número de autores, o la evolución de la producción de un sistema en relación con el número medio de firmantes de los trabajos. La información contenida en este campo sirve para definir la colección de los *agentes efectivos*, es decir, los que han producido resultados en un periodo determinado. El análisis de la *demografía* de un sistema científico en términos de sus agentes efectivos (Maltrás, Quintanilla, 1992) puede iluminar rasgos esenciales para comprender su constitución o su grado de madurez.

Los problemas de homonimia y de variaciones en la grafía del mismo nombre dificultan la realización de algunos de esos estudios, que requieren un esfuerzo de comprobación minuciosa para normalizar el modo en que figuran los autores y muchas veces fuentes de información complementarias, como las memorias de las instituciones de trabajo.

**Título (TI).** Como frase concisa que resume y resalta los aspectos principales del trabajo, el título ha ejercido una considerable atracción como fuente de información para análisis sobre el contenido de los artículos. Las palabras del título pueden servir para situar temáticamente el documento, pero también como base de estudios estadísticos que aspiran a capturar las estructuras conceptuales y disciplinares de la ciencia. Del conjunto de títulos se obtiene un corpus formado por un vocabulario especial, el empleado para situar los trabajos en redes de problemas.

Se han desarrollado diferentes metodologías para los análisis de palabras compartidas (*co-words*), mediante los que se ha intentado penetrar en la microestructura de las disciplinas o encontrar patrones objetivos para definir éstas (Whittaker, 1989; Leydesdorff, 1989). El resultado más llamativo es la construcción de mapas a partir del análisis estadístico de las relaciones entre las palabras. Estos estudios prometen una espléndida base para la comprensión de la dinámica conceptual de la ciencia. Nuestro grupo ha desarrollado un método basado en el análisis de lexemas, para superar los problemas derivados de las variaciones en el vocabulario (Broncano, Maltrás, Vega, 1995).

**Revista (JN).** Recoge el nombre de la revista, monografía o actas de congreso en la que apareció publicado el trabajo. Relaciona el documento con su fuente y permite así aprovechar la información sobre esa

fuente: área a la que pertenece, nivel de prestigio, país de publicación, etc. Debido a la fuerte especialización temática de las revistas científicas, en los estudios bibliométricos se emplea habitualmente esta información para distribuir temáticamente los documentos, ante la imposibilidad práctica de analizar individualmente cada uno. Dado que la aparición de un documento en una revista particular no puede considerarse algo accidental, los datos relacionados con la importancia de la fuente (como el *factor de impacto*, que se tratará más adelante) pueden emplearse también para juzgar agregados de documentos estadísticamente significativos. Por otra parte, a través de los documentos que estas fuentes publican, se generan relaciones entre ellas, lo que las convierte en objetos interesantes para estudiar la dinámica y la estructura de la ciencia.

**Año de publicación (PY).** Indica el año en que apareció el volumen que contiene el documento. Este dato permite clasificar temporalmente la producción y realizar estudios evolutivos. Pertenece al documento y no es accidental: marca la fecha en que se hizo disponible para la comunidad científica. Otro dato interesante, en especial para microanálisis de descubrimientos, podría ser el año en que se terminó el documento o el de recepción del manuscrito, que figura en muchas revistas, pero esta información no se incluye en la bases de datos de referencias.

Algunos estudios bibliométricos emplean el año en que la referencia se indizó en la base de datos, alegando una mayor facilidad técnica y rapidez para cerrar cada año estudiado. Sin embargo, esta elección carece de significado, es arbitraria y puede dificultar la repetición de los resultados al comparar estudios de bases distintas.

Al vincular una determinada producción bibliográfica con una política científica o con un gasto de recursos, como un efecto de éstos, ha de tomarse la obvia precaución de considerar el necesario retraso del hecho de la publicación de los resultados científicos respecto a las actividades o las inversiones que fueron sus causas. Se puede estimar que, por lo general, transcurren unos dos años desde las investigaciones hasta la publicación de sus resultados, y algo más desde la puesta en marcha de una política científica hasta que sus consecuencias comienzan a ser visibles, pero esta aproximación puede variar de unos casos a otros.

**Código de identificación (GA).** Se utiliza para la petición de originales del documento. Es de escasa utilidad en los estudios bibliométricos, pero puede resolver problemas de repetición de documentos

cuando se trabaja con agregados realizados a partir de búsquedas que no garantizan la exclusividad mutua.

**Tipo de documento (DT).** El tipo de documento puede ser uno de los siguientes en esta base:

Article	Artículo
Meeting abstract	Resumen de congreso
Note	Nota
Letter	Carta
Review	Revisión
Editorial	Editorial
Discussion paper	Debate
Biography	Biografía
Software review	Revisión de programa informático
Book review	Revisión de libro

Los tipos *artículo*, *nota*, *carta* y *resumen de congreso* concentran en torno al 98% de los documentos indizados en la base SCI. En algunos estudios bibliométricos puede ser interesante discriminar la producción por tipos de documento para observar patrones característicos o realizar análisis más finos. Schubert, Glänzel y Braun (1989) consideran que sólo debe atenderse a esos cuatro tipos en los estudios basados en las citas recibidas.

**Idioma de publicación (LA).** Informa sobre el idioma en el que está escrito el documento original (existen revistas científicas multilingües). Es un factor relevante para el estudio de los patrones de comunicación. Puede estar sesgado por la cobertura de la base de datos; este dato puede servir precisamente para evaluar ese sesgo y cómo afecta a cada caso. En la base SCI, el inglés es la lengua imperante; pero no hay que olvidar que la mayor parte de las revistas internacionales, aunque estén editadas en países no angloparlantes, se publican en este idioma.

**Número de referencias (RF).** Es el número de trabajos previos a los que ha citado el documento original. Sus aplicaciones en bibliometría son muy específicas: puede servir para calcular si existen diferencias en las costumbres de citación de cada área y si éstas varían con el tiempo; también podría explorarse si existe alguna correlación con la calidad o con las citas recibidas.

**Lugar de trabajo de los autores (CS).** Permite identificar y localizar las instituciones de trabajo de los autores. Hace posible la distribución geográfica e institucional de la producción. Es un campo paralelo al de los autores: en otro nivel se pueden hacer los mismos análisis que para los autores. También sirve para estudios de colaboración y demográficos de unidades de investigación de un sistema. En la base del SCI, los nombres de países están completamente normalizados, y se siguen ciertas pautas para los nombres de instituciones; sin embargo, se requiere un esfuerzo de unificación considerable si se quieren hacer estudios institucionales, de unidades de investigación o con criterios geográficos particulares (de Bruin, Moed, 1990; Maltrás, Quintanilla, 1995).

Algunas bases sólo incluyen el lugar de trabajo del primer autor, lo que supone una limitación en su uso para análisis bibliométricos. En la base del SCI se enumeran las instituciones de todos los autores del trabajo. Como puede darse el caso de que varios autores de un documento trabajen en la misma institución o que un autor haga constar varios lugares de trabajo, sin información adicional no siempre es posible adscribir cada autor a su correspondiente lugar de trabajo. Pero puede servir también como ayuda en la elaboración de los índices de autores, para deshacer las homonimias.

**Resumen (AB).** Es un campo exclusivo de la versión *with abstracts* de esta base de datos. Permite generar un corpus del lenguaje científico con unas magníficas características, ya que parte de frases enteras que van a lo esencial del trabajo, que guardan una relación con la estructura discursiva del artículo y con una variedad léxica más cercana al uso real en la ciencia que el que pueden reflejar los descriptores basados en palabras-clave o en los títulos. Los análisis de la estructura léxica pueden encontrar en el resumen un virtuoso término medio entre la concisión excesiva del título y la ruidosa riqueza del texto completo.

**Descriptores del autor (DE).** Contiene las palabras-clave mediante las que el autor destaca lo más significativo del trabajo. No suelen pertenecer a vocabularios controlados o estructurados ni seguir normas precisas de utilización, por lo que sus uso estadístico plantea severas dificultades. Aún así pueden servir en estudios sobre ámbitos muy definidos para captar la autoimagen de los investigadores o la organización conceptual que puede derivarse de ella. Este dato no aparece en todos los documentos; algunas revistas incluso impiden que los propios autores describan el trabajo con estos términos.

**Descriptores de los indizadores (ID).** Sitúan temáticamente el trabajo mediante términos de un tesoro o vocabulario controlado, aplicados por indizadores profesionales con la finalidad de obtener una clasificación homogénea de los documentos. También existen programas de indización automática o semi-automática. Los estudios de co-palabras clásicos se basan especialmente en la información recogida en este campo, por las virtudes citadas y por la simplicidad técnica que ofrece en la colección de los datos (Callon, Courtial, Turner, 1991; Leydesdorff, 1992).

**Lista de referencias (CR).** Contiene una identificación codificada de los trabajos citados por el documento en cuestión. Es una característica distintiva del SCI, diseñada en principio para búsquedas de trabajos relacionados entre sí, pero ampliamente explotada luego en estudios sobre la estructura de la ciencia y para la evaluación. La estructura de la clave empleada por el SCI es la siguiente: primer autor, abreviatura de la revista, número de la revista, año y página inicial.

En principio sólo permite hacer estudios de citas para los primeros autores, pero si se relaciona con el documento al que hace mención, pueden aprovecharse la información de la referencia completa. Las posibilidades de análisis que abre, basados en recuentos o en las redes derivadas de las conexiones entre documentos que establecen, son innumerables; la estructura disciplinar o la social ya han sido objeto de varios estudios basados en citas.

Las posibilidades que aquí se han enumerado, aun siendo potentes y variadas, alcanzan una mayor significatividad cuando se combinan entre sí. El análisis bibliométrico ofrece, pues, una considerable riqueza, de la cual queda la mayor parte por explotar.

Parte II

---

*INDICADORES BIBLIOMÉTRICOS*



## Capítulo 4

### *Indicadores de producción*

---

Los indicadores bibliométricos de *producción científica* son medidas, basadas habitualmente en *recuentos* de publicaciones, con las que se pretende cuantificar los *resultados científicos* atribuibles bien a unos *agentes* determinados o bien a agregados significativos de esos agentes. Las publicaciones que se tienen en cuenta son documentos pertenecientes a la literatura científica, lo cual asegura su adecuación a unas determinadas características formales y de contenido, como se ha visto en los capítulos anteriores. Los agentes elementales son los investigadores, pero es más frecuente calcular indicadores de producción referidos a agregados como instituciones, regiones, países o disciplinas.

El *número de publicaciones* es el indicador de producción más sencillo, y seguramente el primer indicador bibliométrico empleado conscientemente como tal.<sup>33</sup> La base de éste y de los otros indicadores de producción es, en principio, muy simple: en circunstancias equivalentes, cuantos más documentos científicos publicados tanto mayor es la cantidad de resultados obtenidos que se pueden atribuir. Si al analizar los trabajos publicados en una selección significativa de revistas, se observara que los investigadores de cierto laboratorio son responsables de, pongamos, cincuenta artículos, mientras que los de otro laboratorio

---

<sup>33</sup> El primer trabajo conocido lleva a cabo recuentos de las publicaciones de anatomía desde 1543 a 1860 (Cole y Eales, 1917).

aparecen como firmantes de sólo cuatro, habría que inferir que es muy probable que el primer laboratorio haya producido más resultados que el segundo. Esto, evidentemente, supone y requiere que existan una serie de condiciones compartidas: que los artículos sean de la misma clase, que hayan superado controles similares, y otras por el estilo. Como puede verse, la suposición básica de los indicadores de producción es razonable, pero no puede ocultar una cierta fragilidad.

El sentido primario de los indicadores de producción, como el de otros indicadores bibliométricos, es permitir la *comparación entre un conjunto de agentes o de agregados científicos* con la finalidad de detectar diferencias relevantes que sirvan para caracterizar el comportamiento de cada uno de ellos o del sistema del que pueden formar parte. Es muy importante tener presente que los indicadores de producción sólo pueden ser interpretados comparativamente, pues esto es esencial para una correcta comprensión del uso y alcance de estos indicadores. Este carácter comparativo representa dos imposiciones fundamentales. La primera, obvia, es la necesidad de disponer de datos referidos a más de un agente o agregado. En realidad, esto no es más que otro modo de decir que los indicadores de producción no pueden proporcionar una medida absoluta de los resultados científicos, sino solamente marcas que permiten estimar las posiciones relativas de los productores. La segunda está estrechamente vinculada a ésta: los agentes deben compartir un mismo marco de referencia y ciertas características, para poder interpretar las distancias relativas entre unos agentes y otros. Por ello, el conjunto de agentes o agregados ha de estar adecuadamente construido, no basta una lista cualquiera. El procedimiento por el que se define una colección de agentes sobre el que aplicar un indicador de producción debe incluir un criterio claro y explícito que especifique cuáles son los rasgos compartidos (es decir, a los que no se puede recurrir como causantes de las diferencias que se observen); al disponer de un conjunto homogéneo de agentes se configura el marco de referencia que posibilita la interpretación de los valores del indicador.

De un modo general, por tanto, podemos decir que los indicadores de producción establecen una conexión entre una *colección de agentes científicos* y sus correspondientes *resultados* o productos de la actividad que les es propia. Los recuentos de publicaciones son el medio que permite realizar esa conexión. Aunque su resultado bruto, el número de publicaciones, puede ser directamente empleado como indicador de

producción, existen limitaciones e inconvenientes que aconsejan efectuar transformaciones sobre esos recuentos para obtener indicadores con un ámbito de aplicación más extenso y menos problemático. Los recuentos han de verse, pues, como una operación previa a la construcción de indicadores bibliométricos de producción. Esta operación ha de ser analizada con cierto detalle, pues de ella depende de un modo crítico la validez de todo lo realizado a partir de sus resultados. Además, un examen minucioso proporcionará las guías para la definición de indicadores que permitan sortear los inconvenientes de los recuentos brutos de publicaciones. En realidad, cabe ver la búsqueda de indicadores de producción precisamente como una serie de respuestas a las limitaciones de los recuentos como indicadores.

En los apartados siguientes se tratará, en primer lugar, en *qué* consiste esta conexión, cuáles son sus *fundamentos* y cuáles sus *límites*. Con ello se habrán esclarecido los fundamentos conceptuales de la *interpretación* de los indicadores bibliométricos de producción y se habrán tratado los problemas metodológicos de carácter más técnico que pueden afectar a la validez de esa interpretación. Este proceso de justificación permitirá precisar el contenido del concepto de *producción científica*, del que, a pesar de ser la referencia fundamental en este tipo de indicadores, no puede decirse que haya sido expresado con claridad. La ausencia de una definición suficientemente clara o la vaguedad que se ha asumido a este respecto es seguramente la principal causa de controversias en torno a los indicadores de producción. También se repasarán algunos de los indicadores bibliométricos de producción concretos, señalando sus aplicaciones más interesantes en el ámbito de la descripción cuantitativa de los sistemas científicos, pero también sus insuficiencias y sus límites. Entre las aplicaciones, podemos destacar el papel de estos indicadores en la construcción de marcos valorativos e incluso predictivos de la *eficiencia* y de la *productividad*. De la constatación de tales límites se derivará, por un lado, la comprensión del alcance de su significación, y por otro, la necesidad de otros indicadores complementarios. Una de las consecuencias más relevantes en este sentido es la de la distinción entre la cantidad y la calidad de la producción científica, que hunde sus raíces precisamente en el reconocimiento del carácter parcial de los indicadores de producción.

#### 4.1. Los recuentos de publicaciones

Contar el número de trabajos científicos que ha publicado un determinado laboratorio, región o país es, seguramente, el procedimiento más accesible para obtener una cuantificación con una base objetiva que describa la actividad o el peso de esos agregados en el panorama científico. Es muy difícil sustraerse a la tentadora atracción que ejerce la simplicidad de la interpretación que se intuye y la reducción a números de un fenómeno complejo. La relativa sencillez de la obtención de recuentos, especialmente desde la consolidación de bases de datos informatizadas que recogen la literatura científica, y la engañosa objetividad de sus cifras han favorecido la proliferación de estudios de variada índole que hacen uso de esos recuentos de publicaciones científicas, con un grado diverso de rigor y de fortuna. En el mejor de los casos, las limitaciones señaladas por los propios autores de los estudios bibliométricos son ignoradas, junto con las precauciones que aconsejaron, en cuanto se dan a conocer los valores de los indicadores. Esos números son, en general, o bien tomados ingenuamente, como representaciones objetivas de la realidad, o bien recibidos con suspicacias, como pretensiones espúreas de medidas objetivas.

Su aplicación no suele estar libre de polémicas, en especial porque suelen emplearse para juzgar externamente los méritos de los investigadores, el vigor de áreas temáticas, la eficacia de instituciones, etc., lo cual puede tener consecuencias bien palpables en el ámbito correspondiente. Las quejas suscitadas por este tipo de indicadores no carecen de cierta base y al menos debe admitirse que no siempre se han usado con el rigor necesario. Por ello resulta pertinente ocuparse ahora de varios problemas que pueden cuestionar la validez de los recuentos de publicaciones como proveedores de la materia prima de los indicadores bibliométricos de producción.

En primer lugar, se abordará una cuestión básica: ¿de qué modo cabe entender el resultado de un recuento, cuál es su *significado*? Si los recuentos carecieran de él, es obvio que no podrían ser empleados en el cálculo de ningún tipo de indicador. Es preciso, por tanto, dar una respuesta satisfactoria a las dificultades que puedan plantearse a este respecto.

Algo menos conceptual que lo anterior se plantea en torno al problema de la *atribución* de resultados. Por una parte, hay que justificar la

confianza en la veracidad de las autorías de las publicaciones. Por otra, hay que resolver los problemas metodológicos que origina la colaboración científica y justificar las alternativas que se presentan.

La *métrica* de los recuentos es la tercera cuestión. Aunque es de un carácter más bien técnico, afecta también a la interpretación de los valores de los recuentos. Se refiere a la naturaleza de las unidades empleadas en los recuentos: ¿son o no lineales y constantes estas unidades? Este tema tiene especial relación con la estructura de la productividad y con la variabilidad a través de disciplinas y periodos.

Por último, se comentará el alcance de los problemas relacionados con la fiabilidad de la fuente empleada, lo que en la práctica equivale a decir aquellos derivados de una cobertura insatisfactoria de las bases de datos y de sus posibles sesgos disciplinares, geográficos, etc.

#### 4.1.1. *El significado de los recuentos*

Tan natural como resultó para los pioneros en la *cienciometría* relacionar la cantidad de documentos publicados en revistas científicas con la magnitud de los resultados alcanzados, debió de ser el rechazo de esta posibilidad por parte de los que consideraban que la ciencia era un producto cultural irreducible a sus aspectos cuantitativos. Pero, más allá del estandarte que cada uno enarbole, es posible cuestionar radicalmente los recuentos de publicaciones.

El problema es: ¿qué significa la suma de dos publicaciones científicas? Es evidente que el valor para la ciencia de una cierta publicación puede ser muy superior al valor de otros cientos de publicaciones juntas, señalan repetidamente los críticos de los recuentos. Contar documentos supone, según éstos, ignorar el contenido de cada uno de ellos, cometer la injusticia de confundir cantidad con calidad o la de valorar por igual todas las publicaciones sin entrar en la consideración de los méritos de cada una. Las diferencias entre los trabajos, que pueden ser esenciales, son marginadas por los recuentos y cada publicación pasa a ser simplemente una unidad más en la suma final. Esta burda operación "no es capaz de distinguir entre la elocuencia de un genio y el ruido de vasijas vacías" (*Nature* (1970), citado por Martin e Irvine, 1980: 65). Existen notorias diferencias no sólo respecto a la calidad de las publicaciones sino también en cuanto a la cantidad de contenidos de cada una.

Por lo tanto no valen lo mismo todas, sumar un artículo bueno con uno mediocre es lo mismo, por ejemplo, que sumar dinero en dos monedas diferentes, la suma resultante es un disparate.

Pero, incluso suponiendo que todas las publicaciones fueran portadoras de un contenido equiparable, puede objetarse aún de modo más radical: ese contenido no puede sumarse. El documento es un mero vehículo de lo que importa realmente: el conocimiento científico, algo intangible que escapa a la cuantificación. No debe confundirse conocimiento e información: mucha información puede suponer un escaso conocimiento y viceversa. La información puede sumarse; el conocimiento, no. El valor científico de las publicaciones reside en el valor del conocimiento, en su importancia para el progreso científico.

Las consecuencias del trabajo de Kuhn también parecen proporcionar una base para abundar en esta posición crítica de los recuentos. Una vez rota la ingenua concepción del progreso científico como un desarrollo lineal y acumulativo, resulta complicado imaginar la adición de conocimientos cuando suponen una ruptura conceptual radical, es decir, cuando son inconmensurables entre sí.

Puede agruparse este conjunto de objeciones diciendo que se dirigen a uno de los puntos críticos en cuanto a la posibilidad de que los recuentos tengan algún significado: la *aditividad*. Los recuentos son esencialmente una operación de suma, y sin aditividad son simplemente absurdos. Es necesario especificar qué es lo que se suma, por qué puede sumarse y cómo debe entenderse el resultado.

Otro grupo de dificultades gira en torno a la *representatividad* del documento científico respecto a la contribución a la ciencia. Lo que se pone en duda es el carácter necesario y recíproco de la relación entre el descubrimiento científico y su publicación formal: es posible que no todas las publicaciones tengan contenido científico, o que no todas las contribuciones aparezcan en esos canales formales. Unos advierten, por ejemplo, que las publicaciones de los científicos no siempre contienen nuevos conocimientos: algunas pueden simplemente repetir cosas ya sabidas con ligeras modificaciones, otras pueden contener errores. Por lo tanto, al contar publicaciones sin más control (es decir, sin filtros cualitativos) se *colarían* algunas intrascendentes o equivocadas que producirían sesgos en el recuento final. Otras veces se objeta que los documentos publicados por los científicos contienen en realidad sólo una pequeña parte del conocimiento científico difundido, especialmente

porque entre los científicos existen canales informales de comunicación a través de los que fluiría lo realmente sustantivo. Otros argumentos, además de éstos, ponen en duda el carácter necesario de la relación entre la consecución de los resultados y su publicación formal. Quizá muchos científicos se estén guardando para sí mismos, para sus empresas o para una reducida élite de colegas los resultados de sus investigaciones. O se haya marginado a otros que no consiguen publicar sus trabajos, seguramente adelantados a su época, por la incompreensión de los revisores.

Estos obstáculos están estrechamente vinculados con la motivación de la conducta difusora de los investigadores. A partir del cuestionamiento de la imagen mertoniana, se ha puesto un énfasis creciente en la posibilidad de que las razones que pueden llevar a los científicos a publicar se deban a factores ajenos a los asépticos ideales científicos, como presiones sociales, políticas, de carrera, costumbres institucionales o disciplinares. Este problema se ha agravado en las últimas décadas al cobrar cada vez mayor importancia los motivos curriculares o profesionales desde que se ha extendido el uso de los recuentos como criterios de valoración de los resultados de los investigadores o las instituciones científicas. También puede suponerse que se dé lo contrario: que no exista un incentivo suficiente para poner a disposición común un producto valioso obtenido con el esfuerzo o la inversión propia. Todo esto, unido a una posible duda sobre la eficacia del sistema de control de calidad que han de superar las publicaciones, ataca directamente a las dos direcciones del vínculo que se supone entre contribuciones y literatura en la ciencia.

Junto a las razones aducidas para negar sentido a los recuentos de publicaciones, puede que existan también sentimientos casi viscerales implicados en este rechazo. A muchos científicos podría resultarles repulsivo que los mismos métodos que ellos han aplicado con éxito para comprender mejor el mundo natural, puedan usarse sobre esa misma actividad científica, complejísimo producto de una larga trayectoria cultural y que requiere tanto esfuerzo de los individuos humanos más capaces para ser asimilada. Intentar reducir los frutos de la actividad más noble, creativa y propiamente humana a una secuencia de números raya el sacrilegio. La ciencia ha sido vista frecuentemente como una actividad distintiva, elevada, compleja y creativa, cuya descripción en términos matemáticos y mecanicistas dejaría de lado sus aspectos más esenciales<sup>34</sup>.

Es necesario clarificar estas disputas si se quiere continuar adelante, hay que separar argumentos de malestares, en suma, delimitar con precisión cuál es la cuestión, explorar hasta dónde pueden llegar las pretensiones de quienes encuentran útil contar documentos científicos y señalar las implicaciones que esta afición pueda tener. Lo que queda tras la enumeración de las principales dificultades conceptuales con que se enfrenta el intento de considerar significativos a los recuentos de publicaciones científicas es, al menos, una formulación concisa del problema. *Los recuentos son una operación aditiva que supone que todas las publicaciones son equiparablemente portadoras de lo valioso para la ciencia.* Los críticos pueden rechazar la posibilidad de que esa suposición sea cierta por alguna de las siguientes razones, no incompatibles entre sí: las publicaciones no son necesariamente portadoras de lo valioso para la ciencia, o no lo son en modo equiparable, o lo valioso no puede sumarse, pues lo valioso para la ciencia es el conocimiento que la hace progresar.

Los recuentos de publicaciones científicas se han venido utilizando de manera habitual, principalmente con el argumento de que, aunque no sean medidas perfectas de la producción científica, sí constituyen una buena aproximación estimativa. Las bases de esta confianza se encuentran en el control de calidad de las mejores revistas científicas, en considerar que el canal que éstas constituyen es el principal difusor del conocimiento científico, en aceptar que las normas mertonianas, después de todo, sí juegan un papel en el buen funcionamiento de la ciencia, en que la ciencia se autorregula para asegurarse la pervivencia... En suma, en que la conjunción de una serie de factores diversos permite admitir objeciones como las planteadas sin que ello implique la desacreditación definitiva de los recuentos como indicadores de la producción.

Esa confianza en los recuentos es el resultado de una percepción intuitiva de la validez de éstos, sin que se base en una fundamentación sólida. Prueba de ello es la vaguedad reinante en cuanto al significado de la noción básica en este terreno, aquella a la que hacen referencia los resultados de la operación de contar documentos, la noción de *producción científica*. A pesar de ser usada con profusión, todavía no se ha especificado con exactitud su contenido ni aclarado su relación con la operación de los recuentos. Avanzar en esta tarea pendiente es lo que nos proponemos ahora.

---

<sup>34</sup> Puede considerarse paradigmático de esta corriente el editorial de *Nature* (1968) criticando a Price.

Los recuentos no se ven afectados por las diferencias de contenido entre los trabajos publicados, los tratan a todos por igual. Es ineludible justificar este hecho para poder defender que los recuentos tengan algún sentido. Una posible solución sería admitir que no todas las publicaciones tienen el mismo valor, pero replicar que al trabajar con agregados suficientemente grandes las diferencias dejan de ser estadísticamente significativas; en otras palabras, que cuando nos encontramos ante dos montones de artículos escogidos aleatoriamente (o suficientemente grandes), encontraremos en ellos la misma proporción de artículos excelentes, buenos, dudosos, mediocres e intrascendentes. Como los recuentos son operaciones estadísticas realizadas sobre muestras significativas, el resultado final no se verá afectado por las diferencias particulares entre los trabajos, ya que puede interpretarse como una suma de documentos de tipo medio.

Pero esta solución no daría cuenta, por ejemplo, de las diferencias productivas que se observan entre las disciplinas: si a partir de un cierto número podemos pensar que todos los artículos son del *tipo medio*, ¿por qué razón publican más los médicos que los matemáticos? Además, en los recuentos tampoco se cumple la condición de selección aleatoria, ya que estamos intentando averiguar la producción de agentes determinados previamente, cuyas diferencias esperamos descubrir: no podemos por lo tanto presuponer que sus trabajos son iguales, podría darse el caso de que un agente produzca sólo artículos buenos y otro sólo malos. La *presunción de inocencia* de los grandes números no ayuda aquí a detectar los rasgos peculiares de cada agente.

Es posible otra respuesta, que encuentra un sentido en el que puede decirse que todas las publicaciones *sí* tienen el mismo valor. Recurre al funcionamiento del sistema de control de las publicaciones por parte de la comunidad científica. No todos los manuscritos enviados a las revistas científicas son publicados: los editores de las revistas científicas hacen examinar esos trabajos por investigadores de contrastada competencia en el campo correspondiente, quienes se encargan de valorar el interés y el cumplimiento del rigor científicos para rechazar aquellos trabajos que no alcancen el nivel exigido. Esto garantiza que todos los trabajos finalmente publicados superen un mínimo, que sean científicamente interesantes. En otras palabras, que contengan al menos lo que podemos llamar *unidad elemental de novedad y de relevancia* científicas<sup>35</sup>. Para comprender bien esta noción hay que advertir que esto no

significa que sólo se publiquen trabajos incontestables, que este filtro sirva para depurar completamente los trabajos que contienen errores. Los trabajos publicados con errores o conclusiones equivocadas pueden ser considerados también resultados científicos, ya que la novedad y la relevancia no es lo mismo que certeza. No se debe olvidar que una parte de la actividad científica está dedicada a comprobar la fiabilidad de las aseveraciones de otros científicos; a veces, los resultados que luego se demuestran equivocados pueden abrir nuevas vías a la investigación.

La novedad y la relevancia no son necesariamente conocimiento científico, sino que son más bien la condición de que éste pueda surgir. Las conjeturas bien planteadas e integradas en una corriente de problemas, los datos parciales que pueden suponer apoyos todavía provisionales, las hipótesis, la noticia y descripción de nuevos fenómenos, etc., son los elementos básicos que alimentan el proceso del que puede surgir el conocimiento científico, entendido éste como el conjunto de aseveraciones que se consideran suficientemente fiables y fundadas y que muestran un grado aceptable de coherencia entre sí. Puede decirse, sin adquirir ningún compromiso epistemológico particular, que *el conocimiento científico se construye a partir de los resultados científicos*. Simplemente se trata de una cuestión semántica, que separa dos fases de la producción de conocimiento científico: resultados, en los que es posible considerar y atender a cosas aún no ratificadas por completo, y conocimientos, que están destilados coherentemente de esos resultados. Los resultados científicos son aquello que merece la atención de la comunidad científica; son, en algún grado, nuevos y relevantes para la ciencia. El error de atribuir a todas las publicaciones al menos una unidad de novedad y relevancia no puede ser muy grande, por lo que los recuentos pueden entenderse como sumas de esas unidades.

El planteamiento de que las publicaciones científicas se caracterizan por ser nuevas y relevantes abre un camino para definir un sentido en el

---

<sup>35</sup> Existe un interesante precedente, muy relacionado con esta noción: "Una reacción casi instintiva contra los recuentos sin sentido es convenir que cada trabajo representa, al menos, un quantum de información científica útil y que algunas contribuciones concretas pueden desbordar de tal forma ese valor que, por sólo una de ellas, un autor puede ser valorado por encima de los científicos prolíficos, con un centenar o incluso con un millar de publicaciones ordinarias" (Price, 1963: 109). Desgraciadamente, Price no desarrolló esta idea ni la dotó de un marco que la justificara. Recuérdese además que constituía el núcleo de la anterior presentación de lo que es un *resultado científico*.

que las publicaciones son equivalentes y sumables entre sí, pero también para justificar que son intercambiamente portadoras de algo valioso para la ciencia. Desde luego, es demasiado pedir que el valor de una publicación resida en haber cambiado el rumbo de la ciencia. Parece más razonable y ajustado a la realidad de la actividad científica pensar que la ciencia consiste en el esfuerzo cotidiano de muchos científicos, aunque no todos ellos (sino sólo unos cientos entre millones) asciendan al olimpo de los héroes del conocimiento. En todo caso, detectar las contribuciones sobresalientes no es precisamente el fin que puede proponerse quien aborda un estudio de tipo estadístico. Lo que interesa en un análisis de esa clase es extraer los rasgos generales más significativos para conocer mejor el funcionamiento de la ciencia en cuanto sistema o de alguno de sus subsistemas.

De este modo, los recuentos pueden interpretarse como sumas de unidades elementales de novedad y relevancia. Es al resultado de estas sumas a lo que denominamos *producción científica*. Esta noción difiere significativamente de aquella otra a la que pueda llamarse *contribución científica*. La confusión entre ambas nociones es, seguramente, una de las principales causas de discordancias respecto a la validez de los recuentos y de los indicadores derivados de ellos, por considerar los críticos que los recuentos de publicaciones no pueden representar la importancia o significación de la contribución científica debida a cada investigador o agregado.

En cambio, la noción de unidad elemental de novedad y relevancia, a pesar de tener su origen en el funcionamiento del sistema de publicación, abre una posibilidad ajena a la valoración interna de las contribuciones científicas. En ésta, el juicio acerca de la magnitud de las contribuciones entremezcla de manera cambiante y contextual criterios diversos, en los que es difícil separar los aspectos de cantidad y los de calidad. La distinción entre cantidad y calidad parece irrelevante en la práctica cotidiana de los científicos quienes, en caso de valorar la obra completa de un colega, lo hacen de modo global y aplicando muchos criterios al mismo tiempo.

Por el contrario, en el terreno que nos ocupa (el de la evaluación externa) es necesario distinguir claramente las dos dimensiones de los resultados científicos. Mediante la categoría de unidad de novedad y relevancia, puede reinterpretarse desde un punto de vista externo a la ciencia la actividad evaluadora de los trabajos de los colegas por parte de

los científicos. Según la concepción que hemos delineado en los apartados anteriores, algo es un resultado científico cuando la comunidad científica lo admite como un conocimiento nuevo y relevante (en algún grado). Cada vez que se produzca este reconocimiento podemos decir que estamos ante un resultado científico. Ahora bien, el juicio objetivo de la comunidad científica acerca de si la calidad de ese resultado es mayor o menor puede depender de diferentes consideraciones; aunque no tiene por qué ser explícito ni unánime, sí parece estar relacionado con el *grado* de novedad y de relevancia que se aprecia en los resultados respecto a ciertas dimensiones (no siempre todas ni las mismas). La comunidad científica puede juzgar qué grado de novedad o de relevancia se ha alcanzado considerando las virtudes formales del trabajo (potencia explicativa, descriptiva o problematizadora, simplicidad, elegancia, originalidad, exhaustividad, precisión u otras que pueden definirse en el momento), sus repercusiones en el esquema aceptado en el momento o en los problemas abiertos, u otros aspectos. Pero es necesario tener en cuenta que los procesos de valoración interna de la ciencia parecen referirse más bien a las contribuciones (a los conocimientos científicos) que a los resultados. La evaluación interna de los resultados no persigue otorgar una calificación; forma parte del proceso en el que se comprueba la trascendencia y potencialidades de los resultados para conformar nuevo conocimiento científico.

Lo que todo esto significa es que el esfuerzo por dotar de significado a los recuentos de publicaciones científicas conduce a aceptar la parcialidad de su contenido, es decir, sitúa ante la necesidad (y la posibilidad) de distinguir entre la cantidad y la calidad de los resultados. *Los recuentos de publicaciones pueden servir de base para medir la cantidad de los resultados, y esto implica admitir que existe otra dimensión que no miden (la calidad)*. Qué se entienda por calidad en la ciencia se tratará en los capítulos siguientes.

#### 4.1.2. La atribución de los resultados

Como ya se ha señalado en el comentario sobre las partes de los documentos científicos y de las referencias bibliográficas de éstos, existe un vínculo explícito y obvio entre los investigadores y sus publicaciones: la lista de autores (como también la de las instituciones en las que

han desarrollado su trabajo) figura claramente especificada en estos documentos. De este modo, la operación de asignar los resultados que, por las buenas razones ya tratadas, se supone contiene una determinada publicación científica, al agente que los ha producido es algo tan sencillo como emplear los datos de autoría que aparecen en esa publicación o en su referencia bibliográfica para saber a qué suma, a qué casilla, por así decirlo, hay que añadir otra unidad.

La atribución de los resultados a sus agentes productores encuentra su fundamento, en efecto, en el sistema de publicaciones. La búsqueda de reconocimiento, como ya se mostró en el primer capítulo, es algo ineludible para los agentes científicos que quieran seguir siéndolo, y la vía para conseguirlo es la publicación oficial. En ese contexto, el sistema de competencia honesta, controlado en último término por los mismos competidores, ofrece una garantía suficiente de que lo publicado contenga resultados científicos, lo que permite asignar unos resultados a los autores de las publicaciones. Implica además una voluntad de los agentes por maximizar el número, el contenido y la difusión de sus publicaciones, puesto que no hay razones para desdeñar el reconocimiento que puedan obtener. Esto último significa que los resultados que los propios investigadores consideren meritorios tendrán una probabilidad muy alta de acabar siendo publicados. Todo lo anterior alimenta también la confianza de que quienes figuran como autores sean efectivamente los agentes responsables de haber logrado esos resultados.

Llegados a este punto, la aparente trivialidad de atribuir unos resultados científicos a los investigadores que los han producido se descubre como un ejercicio que requiere asumir unos cuantos supuestos y que presenta varias dificultades en su aplicación práctica, sobre todo cuando son varios los nombres que encabezan un documento científico publicado. Por ejemplo, y haciendo uso de una sospecha precavida, puede ponerse en duda la *autenticidad de la autoría* de las publicaciones: ¿cuál es la garantía que tenemos de que los firmantes de un trabajo sean todos ellos y solamente ellos los responsables de la investigación y de los resultados expuestos? Lo cierto es que resultaría muy complicado purgar la mentira de quien apareciera en el documento sin haber participado en su elaboración, y más complicado aún sería detectar las omisiones a quienes, habiendo colaborado, no figuraran luego como autores.

Los recuentos, que se limitan a obtener sumas y asignárselas a los agentes correspondientes, parecerían entonces extender los efectos de

ese tipo de injusticias que se acaba de señalar. La responsabilidad de esto recaería sobre los firmantes, puesto que su firma es también una declaración pública de autoría. A los medidores externos no les queda más remedio que confiar en que los individuos sabrán guardar sus intereses propios no compartiendo reconocimiento injustificadamente y denunciando sus exclusiones, aunque es difícil olvidar que pueden darse situaciones de complicidad beneficiosa para varios mentirosos. En todo caso, los recuentos reflejarán una situación que es, con anterioridad, ya *de hecho* en la comunidad científica.

La posibilidad de que el fenómeno de la autoría falsa o incompleta tenga lugar despierta enseguida recelos, pues constituiría una posible fuente de sesgos en los indicadores de producción basados en recuentos bibliométricos, sesgos que obrarían en detrimento de la validez de estos indicadores y oscurecerían su interpretación. Sin embargo, dadas las dificultades prácticas para validar cada mención de autoría, los indicadores de bibliométricos de producción pueden erigirse, por el contrario, en instrumentos para minimizar los efectos de esas falsas atribuciones. Es fácil imaginar criterios que, sobre la base de un número suficiente de datos, permitan discriminar entre los comportamientos productivos consistentes y aquellos otros ocasionales o incidentales; así, se presenta más asequible ocultar el hecho de una aparición injustificada o de una omisión en un documento particular que sostener la mentira sistemática y contumaz necesaria para adulterar significativamente el indicador de producción. La posibilidad de que se produzca algún tipo de fraude en la mención de autoría no es considerado un factor peligroso para la validez de los recuentos de publicaciones.

No ocurre lo mismo con el hecho de la *colaboración científica*, que se refleja en las publicaciones mediante la coautoría de los trabajos, y que sí plantea problemas metodológicos acerca del proceso de atribución de los resultados porque admite, como se verá enseguida, varias interpretaciones y alternativas técnicas. Antes de pasar a ocuparnos de la colaboración en relación a los recuentos, no está de más señalar que existe una cierta relación entre la cuestión acerca de la exactitud en cuanto a la autoría y la de la colaboración: las costumbres y algunas condiciones externas (ambas variables a lo largo del tiempo, a través de las disciplinas y entre unidades geo-políticas o institucionales) pueden determinar los criterios para que un investigador figure o no como autor o, en otras palabras, cuál es la contribución mínima necesaria para

aparecer en la lista de autores. El claro aumento tanto de la proporción de documentos científicos confirmados como del número de coautores por documento puede haberse visto influido no sólo por la transformación de la actividad científica misma sino también por un cambio en las costumbres de firma, impulsado quizá por una necesidad creciente de reconocimiento de los investigadores.

La atribución de los resultados científicos puede ser, en efecto, problemática debido a la *colaboración* entre investigadores. Es bien conocido que, desde principios de este siglo, se ha mantenido una clara pauta de crecimiento en el número de autores en cada publicación científica, hecho que es sin duda el fiel reflejo de un modo de investigar que se está generalizando. El trabajo en equipo o la colaboración a distancia a través de las nutridas redes de intercambio de información entre científicos hacen que un resultado concreto pueda ser el producto del esfuerzo conjunto de varios agentes. El peso que haya tenido en ese logro la contribución de cada uno de ellos puede ser diferente, y la proporción, cambiar de un caso a otro. Por otro lado, la colaboración puede modificar la capacidad productiva de cada parte, dando quizá una ventaja a los colaboradores frente a los que no colaboran. Por éstos y otros motivos, urge dar una solución a esta cuestión: ¿cómo atribuir los resultados en el caso de varios firmantes?<sup>36</sup> El hecho simple de la colaboración admite varias interpretaciones, además de poder referirse a distintas situaciones, lo cual genera alternativas metodológicas entre las que es necesario escoger.

#### 4.1.3. *Métodos de fraccionamiento*

Existen varios procedimientos para asignar los resultados a sus productores. La preferencia por cualquiera de ellos puede encontrar argumentos razonables, pero es necesario hacerlos explícitos, en especial porque cada uno conlleva una concepción acerca de qué sea la colaboración. En determinados ámbitos, las diferencias entre lo que arroja una u otra técnica pueden ser escasas y no significativas, por lo que en estos casos se impone, naturalmente, la de mayor sencillez.

---

<sup>36</sup> Se ha explorado igualmente si los artículos con más autores eran más citados. La hipótesis que subyace a esta interpretación sería que quizá la coautoría produce una mayor contribución a la ciencia (Lawani, 1984).

Estos procedimientos pueden aplicarse tanto en el nivel de los investigadores individuales como en el nivel de los distintos agregados que puedan hacerse a partir de ellos (institucionales, geo-políticos, temáticos o las combinaciones de éstos). Todas las alternativas disponibles para el nivel de los individuos son aplicables también en el de los agregados, pero en éste último son posibles además otros métodos, debido a las relaciones que pueden producirse entre los distintos niveles. Comencemos por el nivel de los individuos, explorando las opciones que tenemos para dar una respuesta a la pregunta: ¿cómo podemos asignar crédito a cada uno de los coautores de una publicación científica? La alternativa básica es fraccionar o no el resultado. La variedad se produce en los modos posibles de fraccionamiento, ya que son concebibles diferentes maneras de repartir el resultado entre sus coautores.

*Asignación completa.* A cada coautor se le atribuye el resultado completo, del mismo modo que si lo hubiera producido en solitario. No se fracciona el resultado. Asume que no hay por qué penalizar a los colaboradores, ya que la aportación que hace cada autor puede ser constante, sin importar que se colabore o no; esto implica que se valoran más los resultados en colaboración, ya que un trabajo se cuenta tantas veces como autores tenga. Quizá cabe interpretar también que toma cada autor como un agente imprescindible para que el resultado sea tal cual es.

Una consecuencia de este método es que da lugar a recuentos redundantes, es decir, en los que el mismo resultado se contabiliza varias veces. Si las pautas de colaboración cambian significativamente, este hecho puede producir recuentos equívocos, sesgados por las variaciones en la importancia relativa de los documentos en colaboración o en el número de coautorías. Esto es un obstáculo para emplear este método en la valoración de la producción de agregados y de la evolución de esa producción.

*Asignación exclusiva al primer autor.* Todo el crédito se le asigna al primer autor, y ninguno al resto de los coautores. De este modo se evita el recuento múltiple sin fraccionar el resultado. La principal ventaja de esta alternativa es su simplicidad técnica, ya que disminuye el número de operaciones necesarias para completar el recuento. Parece implicar la preponderancia del papel del primer autor, pero al aplicar este método no se asume necesariamente que la aportación de los otros coautores carezca de importancia: se espera que, para un número suficiente de datos, las posibilidades de aparecer en primer lugar sean

similares al menos para todos los autores importantes. En el nivel de los individuos, dado el gran número de apariciones esporádicas o secundarias, este método define un subconjunto en la población total de agentes (el de los primeros autores) cuyo estudio puede ser de gran interés. Sin embargo, los factores que pueden distorsionar estos recuentos pueden ser variables y estar fuera de control, lo que puede producir grandes injusticias para algunos autores.

Es interesante señalar que éste es el método empleado por el *Institute for Scientific Information* (ISI) para asignar el crédito procedente de las citas a un trabajo: sólo el primer autor de la publicación citada contabiliza en su haber las citas recibidas por esa publicación.

**Fracción igual para cada coautor.** El resultado se divide en tantas partes iguales como coautores lo hayan producido y todos reciben idéntica fracción. Este método otorga el mismo peso a todos los cofirmantes de un documento y considera anecdótica la posición ocupada en la lista de autores. También toma como básico el hecho de que todos los resultados son iguales con independencia del número de productores que cada uno haya tenido, por lo que supone que la contribución aportada por cada coautor será menor cuanto mayor sea el número de colaboradores. Con éste método, al igual que en los siguientes en los que se fracciona el resultado, la suma de todos los recuentos parciales es el recuento total de documentos diferentes, ya que no hay redundancia en el procedimiento.

**Más al primero que al resto.** Otro grupo de posibilidades surge bajo el presupuesto de que la investigación en colaboración se realiza en equipos dirigidos por un investigador principal y que es éste investigador el que aparece encabezando la lista de autores. Se distinguen así dos niveles en la actividad que ha dado lugar a un determinado resultado: el de la responsabilidad y el de la colaboración. Para aplicar esta idea a un recuento, el resultado se divide en fracciones de modo que una sea mayor y el resto iguales entre sí. La mayor se asigna al primer autor por considerarlo el responsable directo de la contribución, y a sus ayudantes se les reconoce su participación asignándoles a cada uno una fracción menor. Un ejemplo, entre los infinitos posibles, es otorgar la mitad al primer autor y repartir la otra mitad entre el resto de cofirmantes. En este caso, si sólo fueran dos autores no se consideraría un equipo dirigido sino un tándem igualitario, y recibirían los dos la misma fracción.

Pueden imaginarse otras fórmulas que asignen al primer autor un peso, aunque superior al resto, decreciente en función del número de autores.

*Fracción decreciente por orden de firma.* Cabe también atribuir un sentido jerárquico a la lista de autores, de tal manera que la ordenación exprese la importancia de la aportación de cada autor al resultado final; con ello se justifica asignar una fracción diferente a cada autor, menor cuanto más retrasado aparezca en el lista. También aquí existen múltiples fórmulas que permiten aplicar este requisito; en unas, la diferencia entre las fracciones sucesivas será constante, y al especificar esa diferencia se define la pendiente del descenso lineal entre las fracciones; en otras, la diferencia entre las fracciones puede ser variable, como por ejemplo la expresada por una función cuadrática inversa.

Dotar de sentido al orden de firma comporta un cierto riesgo: los criterios para establecer ese orden pueden haber sido arbitrarios o ajenos a la importancia de la contribución, como por ejemplo reflejar el orden alfabético de los apellidos (Zuckerman, 1968).

De las técnicas anteriores, las que distribuyen fraccionadamente una unidad entre los autores admiten una variante interesante: que en lugar de repartir una unidad exacta entre los responsables de un trabajo, les otorgue fracciones cuya suma total exceda esa unidad. Esta posibilidad se apoyaría en la opinión de que se debería valorar a las publicaciones en colaboración por encima de aquellas con un sólo autor. Existen ciertos indicios en favor de esta opinión, como la detección de diferencias entre documentos en colaboración y en solitario en cuanto al número de citas recibidas (Presser, 1980; Lawani, 1984). Si se aplicara esta corrección, habría que determinar cuánto se sobrevaloran los trabajos en colaboración, y estudiar si puede aplicarse un multiplicador constante para todas las áreas y periodos. Esta variante debe ser tenida en cuenta y merecería una reflexión detallada en cuanto a su viabilidad práctica.

Las anteriores son las alternativas básicas entre las que debemos elegir al realizar un recuento bibliométrico. Dependen, por un lado, de la variedad de modos en que la colaboración puede tener lugar, y por otro, de la disponibilidad de datos para aplicar ciertos criterios o imaginar una situación u otra. La fidelidad máxima a la realidad exigiría el uso de un procedimiento distinto para cada caso particular, o que los mismos autores especificaran la fracción del trabajo que se les debe atribuir a cada uno. Pero no hay indicios seguros en los que fundar la preferencia por uno u otro método en cada caso concreto, y tampoco es de

esperar que una injerencia externa cambie las costumbres de los científicos en sus publicaciones añadiendo un esfuerzo con una rentabilidad dudosa para ellos. Hay que tener presente que los recuentos, como el resto de operaciones que originen indicadores bibliométricos, deben buscar la significación estadística, por lo que ciertas diferencias particulares carecen de importancia al tratar con conjuntos suficientes de datos. De entre los métodos presentados, hay que elegir el que más se adapte al contexto de estudio y a los intereses que lo guían. Pueden existir motivos para pensar que en un determinado grupo de individuos es más frecuente o más probable un tipo de colaboración que otro, lo que aconsejaría el uso del método de recuento correspondiente. Pero también puede ser útil el uso complementario de varios de estos métodos como medio de profundizar en aspectos diferentes de la producción<sup>37</sup>. En el caso de que no hubiera diferencias significativas entre varios métodos, queda entonces justificada la preferencia por el más sencillo de ellos.

El uso de los indicadores de producción ofrece una enorme utilidad y despierta un singular interés cuando se obtienen los valores correspondientes a los agregados institucionales o geo-políticos. Cualquiera de los métodos anteriores puede emplearse, en teoría, sobre los individuos para obtener el recuento de los agregados correspondientes: basta con sumar los recuentos individuales de todos aquellos investigadores que pertenezcan, por ejemplo, a la misma institución, a instituciones del mismo tipo, o a tal región o país, etc. Se trata simplemente de relacionar los datos de autoría con los de las instituciones que se mencionan como lugares de trabajo de los autores. Sin embargo, en la práctica no suele saberse con certeza qué autores pertenecen a cada institución, por ejemplo en aquellos casos en los que el número de autores es diferente del de instituciones debido a que varios autores son compañeros de trabajo en la misma institución; también puede darse el caso de que un autor haga constar más de un lugar de trabajo. La información sobre los vínculos laborales de los autores puede estar a menudo presente en los documentos originales, pero casi siempre se pierde en las referencias bibliográficas que de ellos se conservan en las bases de datos informatizadas. Como son estas referencias la materia inmediata del análisis

---

<sup>37</sup> Éste es también un modo exploratorio de acercarse al fenómeno de la colaboración; más adelante, al discutir los indicadores de colaboración (capítulo 7), se aplica un enfoque relacional para estudiar en detalle la estructura formada por los agentes científicos.

bibliométrico, resulta prácticamente imposible obtener los recuentos de los agregados a partir de una mera suma de sus componentes individuales. Por este motivo, a la hora de efectuar los recuentos de los agregados, se parte directamente de la lista de lugares de trabajo, prescindiendo del número de firmantes de la publicación que trabajan en cada uno.

Los datos de esos lugares de trabajo permiten varios niveles de agregación: cada mención hace referencia a una unidad básica, que suele ser el *departamento* (o su equivalente funcional), pero los datos que la identifican señalan el organismo del que depende y la ciudad o país en al que está radicada esa unidad. Esto abre la vía hacia la asignación de los resultados a agregados situados en los niveles superiores.

Al lado de los métodos disponibles para calcular los recuentos, expuestos anteriormente para el caso de investigadores individuales, cuando nos ocupamos de un nivel superior al de las menciones o apariciones institucionales se presenta la opción de tener o no en cuenta el número de apariciones de unidades dependientes de cada agregado, lo que da lugar a un abanico de procedimientos de recuento. Esto se comprenderá mejor con un ejemplo. Supongamos que los autores de un determinado artículo trabajan en siete instituciones distintas, pertenecientes a tres países; cuatro instituciones están en el primer país, dos en el segundo, y una sola en el tercero. Si queremos aplicar el recuento fraccionado igualitario en el nivel de los países, podemos asignar  $1/3$  a cada país, porque son tres los países colaboradores; pero también podemos conceder  $4/7$  al primero,  $2/7$  al segundo y  $1/7$  al tercero, ya que la contribución de cada uno puede considerarse proporcional a los medios (instituciones) que cada país ha aportado. En el segundo caso, decimos que se ha tenido en cuenta el número de apariciones de nivel inferior. Cuando son varios los niveles inferiores, podría también elegirse entre ellos. En el ejemplo anterior, podría haberse considerado el *tipo institucional* como un nivel relevante, y haber asignado a cada país una fracción proporcional al número de tipos institucionales diferentes que hubiera movilizado.

Otra vez son la finalidad del estudio y la valoración del coste respecto a la trascendencia de las diferencias detectadas los criterios más importantes para optar por uno u otro método. El análisis pormenorizado de ciertos mecanismos productivos puede requerir un método de recuento más complicado técnicamente, pero en otros casos

el detalle excesivo puede ser innecesario para obtener una visión general válida. En el nivel de los sistemas nacionales de investigación, los recuentos obtenidos por los diferentes métodos pueden manifestar divergencias dignas de consideración y cuya ignorancia puede conducir a una valoración incorrecta de la situación.

#### 4.1.4. *La métrica de los recuentos*

En los apartados anteriores se ha encontrado un sentido coherente a las sumas obtenidas mediante los recuentos de publicaciones científicas: cada una de estas publicaciones contiene al menos una *unidad elemental de novedad y relevancia*, que son la condición de los resultados científicos; esta noción permite sostener la ineludible aditividad de los recuentos. Pero aún falta saber el modo en que esa unidad puede definir una escala y de qué tipo será.

El problema es determinar el tipo de métrica que se deriva de los recuentos de publicaciones. Nuestra experiencia más cotidiana nos pone en contacto con diferentes medidas, como la distancia que separa dos ciudades, el peso de un montón de patatas o de naranjas, etc. Las escalas más habituales, como las anteriores, están basadas en unidades que producen incrementos constantes e independientes de las circunstancias. Si tenemos dos kilogramos de patatas, harán falta otros dos para tener el doble; si tenemos cien kilos, necesitaremos otros cien para doblarlos. Añadir un kilo produce el mismo incremento de peso en cualquiera de los dos casos. Dos kilogramos de patatas pesan exactamente lo mismo que dos kilogramos de naranjas, por la mañana y por la tarde. Un mismo incremento de la masa de fruta producirá siempre (al menos en nuestra frutería y en ausencia de cataclismos geológicos o fraude en las básculas) el mismo incremento en unidades de peso. Las unidades de las escalas que usamos diariamente son en ese sentido permanentes, fijas: valen lo mismo en cualquier parte de la escala y en cualquier circunstancia. Estas escalas se denominan *lineales*: sus incrementos son constantes, su ecuación dibuja una recta, normalmente ascendente, que expresa la constancia de los efectos que tienen los cambios en lo medido sobre las unidades que arrojan la medida.

No todas las escalas que pueden llegar a afectarnos son tan "naturales". Supongamos que nuestro simpático vecino de al lado, amante de la

música clásica y que acaba de actualizar su equipo, nos pide permiso para escuchar sus arias preferidas con una intensidad sonora de "sólo tres decibelios más" que de costumbre. Lo que está preguntando, independientemente de cuál fuera el volumen habitual, es si podemos soportar el *doble* que antes. La escala de los decibelios no es lineal, sino logarítmica, y con cada tres decibelios se duplica la intensidad. Si antes no sospechábamos que le gustaba la ópera, no importará que suba esos tres decibelios de nada. Si, por el contrario, teníamos fuertes indicios de su pasión, y ahora no queremos cohibir su espiritual entretenimiento, tendremos que pensar en proteger los vidrios de nuestras ventanas con cinta adhesiva.

Dependiendo del caso, puede ser preferible una escala lineal o una no lineal. Pero para aplicar una u otra, puede ser necesario "calibrar" el indicador que utiliza tal escala: un indicador funciona en cierto sentido como un instrumento de medida, el cual convierte ciertos estados o ciertas alteraciones del medio en señales comprensibles. En otras palabras, puede requerirse algún tipo de transformación de los datos a los que se tiene acceso directo, para convertir las variaciones detectadas en la entrada en magnitudes de salida que sean interpretables. Por lo tanto, ahora hay que ver cómo parecen comportarse las unidades elementales de novedad y relevancia en relación con los resultados científicos, para determinar el modo en que pueden emplearse los recuentos para construir indicadores de producción.

La cuestión tiene que ver con la naturaleza de las unidades empleadas por los recuentos: si son fijas o varían a lo largo de la escala; si son absolutas o relativas al contexto. En cuanto a lo primero, la distribución individual de la productividad, medida en publicaciones/autor y descrita por la conocida *ley de Lotka*, puede sugerir una aditividad no lineal en los recuentos. Habitualmente, la notable asimetría en la productividad se interpreta como un proceso de *ventaja acumulativa*. En cuanto a lo segundo, existe el convencimiento general de que no se pueden comparar los recuentos brutos entre áreas, ya que cada una parece tener una productividad característica: las unidades no son transportables sin más, son relativas al contexto de origen, es decir, al marco disciplinar.

#### 4.1.4.1. *La estructura de la productividad científica*

En 1926, Alfred Lotka puso de manifiesto una de las regularidades más notables y que más atención han atraído en el ámbito de los estudios cuantitativos de la ciencia. Encontró que la distribución de las publicaciones entre sus autores responde a un patrón similar al de la productividad biológica, y en concreto, que esa distribución se ajusta con una elevada precisión a una ecuación cuadrática inversa: el número de autores que han publicado exactamente  $n$  trabajos es inversamente proporcional a  $n^2$ .

Este enunciado, que expresa el contenido de lo que se conoce como *ley de Lotka*, ha sido empíricamente corroborado en innumerables ocasiones, mostrando una sorprendente fortaleza y exactitud con independencia de áreas, épocas o países<sup>38</sup>. Su único requisito, además de los obvios de contar con una población de autores suficiente y de que la recopilación de sus publicaciones sea exhaustiva, es que abarque un periodo suficiente como para que los autores más productivos tengan tiempo de llegar a su nivel.

Esta ley ofrece una evidente utilidad para realizar predicciones. Por ejemplo, como indica Price (1963: 86), permite saber que por cada 100 autores con un sólo trabajo, habrá 25 con dos, 11 con tres, y así sucesivamente. Empleada de otras formas, la ley sirve para estimar la población total de autores a partir del número de trabajos, la distribución por tramos de la producción, etc.

Pero especialmente es de interés profundo la significación del hecho que describe, que es el de la enorme y sistemática desigualdad entre los autores en cuanto a su eficacia publicadora. Tomemos el caso hipotético de una población de autores, de los cuales 1000 han publicado un único

---

<sup>38</sup> La forma de la ley de Lotka (1926) es la siguiente:  $a = k \cdot n^{-\alpha}$ . Se ha encontrado que el parámetro  $\alpha$  tiene cierta dependencia contextual, principalmente entre disciplinas, y su valor varía ligeramente en torno a 2. Esta ley, sin embargo, pierde precisión en cuanto a los grandes productores, que parecen disminuir a un ritmo superior al inverso del cuadrado. Para salvar este problema, Price (1963: 90) propuso una modificación válida tanto para pequeños como grandes productores, que expresa de manera acumulada esa distribución:  $A = u \cdot k / (p(u+p))$ , donde  $A$  es el número de autores que publican al menos  $p$  trabajos, y  $u$  es el umbral que separa los pequeños de los grandes productores, fijado por Price en 15 trabajos para una vida completa. Numerosos trabajos se han ocupado de esta ley; al menos debe señalarse aquí que la ley de Pareto sobre la distribución de la renta ha de considerarse un claro precedente.

trabajo. Esto significa que el conjunto total ascendería a unos 1600 autores aproximadamente, con una producción conjunta de unas 4000 publicaciones. Lo llamativo es que bastaría solamente con el 3,5% de los autores más productivos para sumar una cuarta parte de la producción total; con el 12% más productivo se alcanza la mitad; y no es necesario el 40% de los autores para superar las tres cuartas partes de las publicaciones de todo el grupo. Los 1000 autores con un sólo trabajo, aunque son más del 60% de los autores, aportan solamente una cuarta parte de las publicaciones del grupo (es decir, lo mismo que el 3,5% de autores más productivos). En el extremo opuesto a los de una sola publicación, el autor más prolífico estaría en torno a los treinta trabajos. Todo esto da una idea clara de las grandes diferencias que existen entre los autores, de la enorme asimetría en la distribución de las publicaciones.

Otro modo de ver al ley de Lotka es como una función de probabilidad de la productividad. El hecho más notorio entonces es que la diferencia de probabilidad que existe entre pertenecer a niveles productivos contiguos disminuye a medida que se publican más trabajos. Para ser más concretos: en el ejemplo anterior, publicar dos trabajos es la mitad de probable que publicar uno solo, es decir, la diferencia de probabilidad entre publicar un trabajo y publicar dos, es de dos a uno. Sin embargo, la diferencia entre 4 y 5 trabajos publicados es de 1,25 a 1, y entre 13 y 14, de 1,08 a 1. Para encontrar la misma diferencia de probabilidad que existe entre 1 y 2 trabajos publicados, tenemos que comparar la probabilidad de 4 y de 8 trabajos o la de 8 y 16, que son también ambas de 2 a 1.

La comparación entre las probabilidades de intervalos contiguos de producción es un modo formal de expresar algo señalado muchas veces: la progresiva facilidad que parece encontrarse para publicar un nuevo trabajo a medida que se publica más. Esto nos introduce en un segundo aspecto de la distribución descubierta por Lotka: además de describir un estado, puede verse también como la consecuencia de un proceso particular cuya característica fundamental es el progresivo aumento de la productividad, si se mide ésta por el número de publicaciones. La cuestión es, entonces, encontrar una explicación adecuada de ese estado, determinar las causas que producen una distribución tan estable, una asimetría de una invariabilidad tan sólida.

La explicación que ha gozado de un mayor éxito es la de que esta desigualdad es el resultado de un proceso de *ventaja acumulativa*, en el

que "el más rico enriquece a un ritmo que hace al pobre relativamente más pobre" (Merton, 1968). Esta noción se ha aplicado algunas veces meramente para describir la aceleración de la productividad; aquí se comentará, en primer lugar, entendida como una explicación causal de este proceso: la hipótesis de la ventaja acumulativa sostiene que existen mecanismos identificables (y señala varios concretos) que hacen que quien obtiene resultados vea aumentada su probabilidad de obtener más resultados con una cierta proporcionalidad a los resultados obtenidos, de modo que su productividad aumentará aceleradamente.

En la sociología de la ciencia, esta hipótesis ha ocupado un lugar central en los estudios acerca de la estratificación social de los científicos, tema íntimamente ligado a las desigualdades en producción y en reconocimiento. En resumen, lo que viene a decirse es que aquellos científicos que producen resultados interesantes ven reforzada su capacidad para producir más resultados gracias a que el reconocimiento que obtienen les da acceso a más recursos para la investigación: mejores laboratorios, medios, becas, colegas estimulantes, estudiantes destacados, etc. También sienten la presión de la expectativa que crean y disfrutan de los ánimos que reciben para dedicar más tiempo a su investigación. Así, será más probable que produzcan nuevos resultados, los cuales le traerán más reconocimiento y más facilidades, que podrán convertir de nuevo en más resultados. Quienes gozan de prestigio también tienen más posibilidades de ser escuchados y de que sus resultados sean admitidos aun cuando no están del todo claros. Por el contrario, quienes no reciben reconocimiento por su investigación verán reducidas sus posibilidades futuras de producir resultados. Como además la distribución de reconocimiento en la ciencia no parece ser proporcional sino favorecer a quienes tienen más prestigio adquirido (*efecto Mateo*) la productividad y el reconocimiento se refuerzan mutuamente y se multiplica el efecto de acumulación de ventaja, es decir, una diferencia creciente de recursos y de incentivos entre los autores, lo cual explica las diferencias productivas observadas.

Tomada en este sentido, la hipótesis de la ventaja acumulativa recurre básicamente a mecanismos derivados de un determinado modo de organización social, el propio de la ciencia, que *redistribuye* las facilidades necesarias para su actividad característica de un modo que hace crecer las diferencias entre los investigadores dependiendo directamente de los resultados que produzcan. Estos mecanismos actúan tanto en el

nivel material como en el psicológico y en el intelectual en forma de recursos e incentivos, cuya cantidad determina la posibilidad de alcanzar nuevos logros.

La hipótesis de la ventaja acumulativa, en el sentido sociológico que se acaba de delinear, tiene un innegable atractivo y se ve apoyada por abundantes indicios. Allison, Long, Krauze (1982: 615) atribuyen su amplia aceptación al avance que supone en la explicación de las enormes desigualdades en cuanto a productividad y reconocimiento, a que es tan intuitivamente plausible que para muchos simplemente *debe* ser cierta, y a que resulta igualmente atractiva para los admiradores y para los críticos del sistema de estratificación en la ciencia. Quizá habría que añadir que ofrece una base legitimadora de la ciencia, ya que permite ver ésta como un sistema autorregulado que tiende a maximizar el aprovechamiento de los recursos que consume, concentrándolos en quienes mejor los emplean. En este sentido cabría interpretar a Cole y Cole (1973) y a Zuckerman (1977).

Puede generalizarse la hipótesis explicativa de la ventaja acumulativa como la expresión de un sistema dinámico retroalimentado. Este tipo de sistemas se caracteriza porque los resultados de una determinada acción tienen efectos reforzadores (retroalimentación positiva) o inhibidores (retroalimentación negativa) sobre la intensidad subsiguiente de la acción que fue su causa. Un caso típico de sistema con retroalimentación negativa es el de una cisterna que corta la entrada de agua cuando ésta alcanza un nivel predeterminado. Por el contrario, una reacción nuclear en cadena es un ejemplo de retroalimentación positiva: la energía liberada por las fisiones produce nuevas fisiones, que liberarán más energía, y así sucesivamente. La hipótesis causal de la ventaja acumulativa en la ciencia incluye mecanismos positivos y negativos: los logros refuerzan la posibilidad de nuevos logros, mientras que los fracasos pueden ser inhibidores; la acción conjunta explicaría las llamativas diferencias observadas. Como proceso, la ventaja acumulativa produciría, al menos entre ciertos márgenes, una aceleración desigual de la productividad entre los individuos, lo que da lugar a diferencias crecientes entre ellos<sup>39</sup>.

---

<sup>39</sup> La ventaja acumulativa es un principio aplicable no sólo en el nivel individual de la productividad científica, también encuentra sentido en la descripción del desarrollo general de la ciencia. Pueden encontrarse antecedentes ya en el enunciado por Engels (1844) de la *ley del desarrollo acelerado de la ciencia*: "La ciencia progresa proporcional-

Desde un punto de vista formal, puede decirse que la ventaja acumulativa sería un caso particular del tipo de procesos gobernados por el principio "el éxito genera éxito", formulado por Simon (1955) y más tarde simplificado por Price (1976). En una formulación adaptada a los procesos de producción de información (Egghe, Rousseau, 1990a: 297 ss.) su enunciado es el siguiente:

"Cuanto más ítemes ha producido una fuente, más alta es la probabilidad de que esa fuente produzca otro ítem; pero siempre hay una pequeña probabilidad de que una fuente sin ítemes produzca su primer ítem" (Egghe, Rousseau, 1990a: 297).

De este principio, a través de un tratamiento formalizado de base probabilista, puede derivarse una "aproximación razonable de la ley de Lotka", la Función de la Distribución de la Ventaja Acumulativa. Además de este resultado, es interesante señalar el modo en que se aplican modelizaciones de urnas para el estudio de este tipo de procesos: un proceso de ventaja acumulativa equivaldría a un procedimiento de extracción de bolas tal que cada acierto (extracción de bola blanca) sería seguido por la introducción de un cierto número de bolas blancas en la urna, de modo que la extracción siguiente tendría más posibilidades de encontrar una bola blanca. Los errores pueden penalizarse introduciendo bolas negras o impidiendo nuevas extracciones. Esto permitiría ver la ventaja acumulativa esencialmente como un mecanismo retroalimentado de redistribución de probabilidades.

La hipótesis de la ventaja acumulativa puede tomarse, por lo tanto, de dos maneras: como una descripción formalizada de un tipo de dinámica que da lugar a distribuciones asimétricas, o como una explicación material que incluye los mecanismos concretos requeridos para ese tipo de dinámica. La versión explicativa que puede reconstruirse a partir de la literatura recurre básicamente a los mecanismos derivados de la *organización social* de la ciencia sostenida por la escuela mertoniana; entre esos mecanismos también se incluyen algunos que podrían denominarse psicológicos o psicosociales. Las consecuencias de estos mecanismos inciden especialmente en la redistribución de las facilidades requeridas para la investigación.

---

mente a la masa del conocimiento acumulado por la generación precedente, es decir, en las condiciones más corrientes, también en progresión geométrica". Engels, "Umriss zu einer Kritik der Nationalökonomie", citado por Piñero (1973).

¿Hasta qué punto puede considerarse satisfactoria la hipótesis de la ventaja acumulativa como explicación de las desigualdades observadas en la ciencia? Por un lado estaría la cuestión de si resulta o no adecuada; por otro, la de si hace innecesarias o excluye otras posibles explicaciones, es decir si es completa o, en caso de que no lo fuera, su grado de compatibilidad con otras alternativas. La hipótesis de la ventaja acumulativa podría ser inadecuada porque sus requisitos formales o las consecuencias que de ellos se derivan no se cumplieran en el proceso productivo de la ciencia. También es posible pensar que algunos de los mecanismos concretos que se señalan pueden no producir los efectos que se les atribuye, o que la lista de mecanismos no sea completa, o que intervengan en el proceso otros niveles explicativos además del de la organización social.

Pero la primera pregunta ha de ser: ¿por qué es necesaria una explicación de las desigualdades? En otras palabras: ¿qué presupuestos son los responsables de que sorprenda una distribución de las publicaciones como la descrita por Lotka, y cuáles hacen plausible la suposición de un mecanismo de acumulación de ventaja?

La marcada asimetría de la distribución de las publicaciones puede resultar inesperada porque sea claramente *discordante con la distribución del talento científico* que se supone entre la población de investigadores. Muchas veces parece implícita la opinión de que las diferencias de capacidad científica son menores que lo reflejado por la ley de Lotka, es decir, que las diferencias productivas no reflejan diferencias intrínsecas entre los científicos. Se detecta una falta de proporcionalidad entre ambas distribuciones porque la producción crecería más rápidamente que lo que se espera de la capacidad. En este caso, conjeturar un mecanismo amplificador de las diferencias, como lo es el de la ventaja acumulativa, es una solución simple y elegante, que además cuenta con correlatos concretos. La discordancia es más aguda cuando al reparto de la capacidad científica se le atribuye una distribución gaussiana: parece intuitivo esperar que, en ausencia de fuerzas externas significativas y en igualdad de condiciones, la producción siguiera también un patrón normal, es decir, que fuera proporcional al de las capacidades. Por lo tanto debería existir una fuerza que explique cómo se llega a esa asimetría a partir de la simetría de la distribución normal.

La primera consideración ante este problema es la de constatar la dificultad conceptual de obtener una medida de la capacidad científica,

entendida como causa de la productividad, que sea completamente independiente de la estructura que se ha observado en las publicaciones: existe un constante peligro de circularidad en la medida. Tampoco hay que olvidar las críticas a los procedimientos de medida de la inteligencia que presuponen una distribución normal y, claro, la obtienen. En estos casos es el tipo de escala elegido la causa de la distribución observada.

Pero la cuestión puede plantearse de un modo más profundo. De lo que se trata realmente es de saber cuáles son los factores que determinan la cantidad de publicaciones: si son las aptitudes individuales, el trabajo invertido, los recursos utilizados, la posición en la comunidad científica, el entorno intelectual, etc. o cualquier combinación de éstos o de otros. Depende del factor que se señale o del alcance que se le atribuya a cada uno el tipo de explicación necesaria para la asimetría en la producción. En la hipótesis de la ventaja acumulativa, el énfasis se dirige a los recursos y a la posición en la comunidad (especialmente como causa principal de la disponibilidad de recursos, entendidos en el sentido amplio que incluye también las ventajas en cuanto a la información, contexto, etc.). Las diferencias en la productividad se hacen depender primordialmente de las diferencias en cuanto a los recursos disponibles. Asimismo, las variaciones de productividad también se atribuyen a las variaciones en los recursos. Puesto que no se excluye que la aptitud individual esté en la base de la productividad, parece deducirse que la aptitud se supone aproximadamente constante en cada investigador.

Todos estos presupuestos implícitamente asumidos por la hipótesis de la ventaja acumulativa cuentan con alternativas igualmente razonables. La productividad puede estar determinada por un complejo variable de factores que incluya los ya citados, por ejemplo. En ese complejo, la carencia de unos factores puede ser salvada con la sobreabundancia en otros. Además, la cantidad de recursos disponibles puede ser más determinante en unas áreas que en otras: la necesidad de recursos materiales parece más ineludible en algunas disciplinas experimentales, pero no igualmente determinante en otras. En cuanto a la distribución de las aptitudes científicas, puede esperarse que, incluso cuando entre la población total respondiera a una curva normal, la población de científicos estaría seleccionada aleatoriamente a partir de un mínimo, con lo que no sería de extrañar un patrón congruente con la distribución de Lotka: muchos investigadores que superan el "mínimo científico" por poco y pocos que lo superan ampliamente. La

explicación del crecimiento en la productividad individual podría tener algo que ver con el proceso de aprendizaje: con la práctica, los individuos mejoran su eficiencia; los investigadores también pueden aumentar su productividad en cuanto a resultados y a publicaciones. Las clásicas curvas de aprendizaje, que reflejan un pronunciado aumento inicial en la tasa de aciertos (o una disminución en el tiempo requerido por la tarea) seguido de una estabilización, son bastante acordes con las curvas cronológicas de la productividad individual de los científicos. Aunque no hay que menospreciar el efecto positivo de la autoridad o la fama anterior en los subsiguientes intentos de publicación, lo cual formaría parte de la explicación convencional de la ventaja acumulativa, también contribuiría a esa ventaja el proceso individual de aprendizaje: se mejora el estilo, se mecaniza la escritura, se eligen los temas y los canales adecuados, se desarrolla el olfato para detectar cuáles son los puntos esenciales, etc.

Si la productividad estuviera determinada por un complejo de factores, las pequeñas diferencias en cada uno de ellos podrían sumarse y dar lugar a diferencias productivas que con el tiempo originarían diferencias de producción cada vez mayores en términos absolutos. Si las diferencias son de partida, puede no ser necesaria la ventaja acumulativa para explicar la asimetría en la distribución de la productividad. Claro que entonces habría que identificar los mecanismos que dotan a algunos de pequeñas ventajas en la mayor parte de los factores productivos; estos mecanismos quizá se encuentren en los procesos de formación y selección de los investigadores, que pueden marcar las diferencias en la "parrilla de salida".

Otro ámbito de relevancia para el análisis de la productividad científica ha de ser el de la *naturaleza del proceso de descubrimiento*. En la hipótesis de la ventaja acumulativa se concibe que el descubrimiento es un proceso en cierto sentido aleatorio, describable en términos de probabilidades, en el que las diferencias de éxito pueden explicarse justificando las diferencias en las probabilidades de cada individuo. La ventaja acumulativa sería el mecanismo que redistribuye las probabilidades y causa esas diferencias. En la modelización formal de la hipótesis se asume claramente un esquema probabilista, como ya se ha visto. Price (1963: 113 ss.), para explicar los descubrimientos múltiples de los que Merton se había ocupado, encuentra que éstos se ajustan a la distribución de Poisson que resultaría si mil individuos intentaran, con los ojos

vendados, recoger mil manzanas que colgaran de sus árboles. Este ejemplo es aducido como una prueba a favor de que los descubrimientos múltiples no son en absoluto sorprendentes sino, al contrario, lo esperable; lo que aquí interesa de él es que puede interpretarse como un particular tipo de imagen sobre el descubrimiento científico.

Pero es posible concebir que el descubrimiento científico no sea un proceso aleatorio del tipo de la *cosecha a ciegas*. Puestos a buscar metáforas, podría compararse más bien con el de la búsqueda de oro: algunos prefieren la garantía de las pequeñas pepitas de los sedimentos del río, otros persiguen las vetas o los filones de la montaña. Para los afortunados (o de especial olfato) de esta última clase, un sólo descubrimiento les supone una extracción más abundante y continuada. De igual modo, parece que muchas veces en la ciencia los resultados llegan todos juntos, relacionados unos con otros, como si el descubridor hubiera hallado una veta o un filón. Si la metáfora *minera* fuera más adecuada que la *recolectora*, resulta bastante natural que se produzcan acusadas diferencias productivas, incluso con mucha igualdad en las capacidades o en los recursos, por lo que no sería completamente necesario acudir a la ventaja acumulativa para aclararlas.

Tras esta somera revisión de las posibles causas explicativas de la asimetría en la producción, conviene explorar un tercer dominio, además de los ya tratados de los factores productivos y de la naturaleza del descubrimiento: el relacionado estrictamente con la medida de la producción. Cuando se dice que la ley de Lotka es la ley sobre la distribución de la *productividad científica*, se está asumiendo que el número de publicaciones es la medida de la producción; y lo mismo cuando se interpreta como muestra de la desigualdad productiva entre los investigadores. La hipótesis de la ventaja acumulativa, en tanto que aplicada al ámbito de la producción científica, adquiere el mismo compromiso. Sin embargo, en el contexto del presente trabajo, el de la fundamentación de las medidas bibliométricas, no extrañará una aproximación más crítica y cautelosa en este punto. Precisamente este capítulo se ha dedicado a indagar sobre el sentido de los recuentos de publicaciones, por lo que no cabe obviar la cuestión sin más.

Existe, desde luego, una posición menos ingenua acerca de los recuentos de publicaciones en relación con las diferencias reflejadas por la ley de Lotka:

Creo que se debe simplemente a que el número de publicaciones no es una medida aditiva lineal de la productividad de la forma requerida por las distribuciones de Gauss. Nuestro punto de abordaje no es la media aritmética del mayor y del menor productor, sino más bien la media geométrica. Intuitivamente se nota que la diferencia que hay entre tres y seis trabajos es más parecida a la que hay entre 30 y 60 que la existente entre 30 y 33. Por todo ello, es razonable pensar que se trata de algo parecido a la ley de Fechner o Weber en psicología experimental, puesto que la medida adecuada no corresponde a la magnitud del estímulo, sino a su logaritmo; tiene que haber iguales intervalos de esfuerzo correspondientes a iguales porciones del número de publicaciones (Price, 1963: 92.)<sup>40</sup>.

Seguidamente, Price propone la definición de lo que puede considerarse una medida de producción científica, en lugar del recuento bruto de trabajos: la solidez  $s$  de un autor es el logaritmo del número de trabajos que publica a lo largo de su vida. Esta transformación de la escala tiene una consecuencia apreciada por Price, que es la de hacer que por cada unidad de incremento de  $s$  el número de autores se divida por un factor casi constante. Esto es lo que sucede también en los extremos de una distribución normal, como por ejemplo la resultante de aplicar la escala AGCT de inteligencia. En esta escala, el 100 divide en dos a la población, y los cuartiles son 80 y 120; en las puntuaciones por encima de 140 (como en las inferiores a 60) el número de casos se divide por un factor de 10 por cada 10 puntos de la escala. Así puede establecer una correspondencia entre la solidez científica y la parte superior de la población ordenada por nivel de inteligencia según una escala normal (lo cual encuentra su fundamento en resultados de ciertos tests que sitúan la inteligencia media de los doctores en un abanico entre 130 y 140 AGCT según las disciplinas).

La ley de Pareto puede, en consecuencia, ser considerada como el mero resultado de combinar una distribución razonable de las aptitudes con una medida basada en la ley de Fechner de la efectividad de dichas aptitudes. En el caso de la productividad científica encontramos una explicación teórica similar de la forma que tiene la ley empírica. La única diferencia entre la distribución de dinero y la de trabajos científicos, o

---

<sup>40</sup> En una nota a pie de página señala Price que el carácter logarítmico de las distribuciones de la productividad científica había sido sugerido antes por W. Shokley (1957).

la distribución más generalizada de que Zipf encontró para explicar casi todas las distribuciones naturales de cosas clasificadas en orden de tamaño, es que para la ciencia existe un límite superior, definido por la cantidad de textos que un autor puede publicar a lo largo de su vida (Price, 1963: 94).

El planteamiento de Price es muy sensato, aunque puede dejar la impresión de adolecer de una base excesivamente intuitiva. Tiene la virtud de no dejarse arrastrar por lo llamativo de la desigualdad publicadora de los científicos y de buscar una explicación que convierta a ese hecho en algo natural, después de una aproximación crítica sobre el significado de la medida bruta de las publicaciones (lo cual se echa de menos en otras teorizaciones sobre la distribución de la productividad). Sin embargo, no ofrece una explicación suficiente de que la capacidad publicadora siga un patrón de efectividad como el de la ley de Fechner; simplemente señala que el ajuste es razonablemente bueno. Quizá sea esta la razón de que su posición en este tema no haya tenido un impacto comparable al que han gozado otras aportaciones contenidas en su famoso ensayo *Little Science, Big Science* (traducido al castellano bajo el título *Hacia una ciencia de la ciencia*).

Una de las nociones introducidas en el presente trabajo, la de *unidad mínima de novedad y relevancia*, puede reforzar el planteamiento esbozado por Price. Efectivamente, "se nota intuitivamente" que los recuentos de publicaciones no pueden ser empleados como medida lineal de la producción: enseguida se sospecha de esas diferencias crecientes y demasiado grandes. Pero si consideramos que el proceso de publicación exige a cada trabajo al menos una unidad mínima de novedad y relevancia científicas, puede encontrarse una explicación razonable de estas diferencias: *las unidades mínimas de novedad y relevancia que exhiben los trabajos de un mismo autor* (seguramente también las de un mismo equipo) *no son siempre completamente disjuntas, se pueden solapar parcialmente*. Esto no representa, en modo alguno, un mal funcionamiento del sistema de publicación. En realidad, lo injustificado es esperar lo contrario, y eso es lo que implícitamente se asume al aceptar ingenuamente los recuentos de publicaciones como reflejo de la producción científica.

Las sucesivas publicaciones de un autor, en efecto, pueden compartir una fracción básica de los resultados y diferenciarse en los detalles del tratamiento, de la aplicación que se sugiera o de las conclusiones que se

extraigan. Precisamente esa fracción compartida puede ser la más difícil de obtener, la que más trabajo, tiempo y medios requiere. Sin embargo, a veces es muy interesante exponer con detalle las implicaciones de los propios trabajos precedentes, reforzar las hipótesis expuestas con anterioridad, o integrar esos resultados en posiciones teóricas más generales. La investigación científica no es un proceso atemporal; sus resultados emergen de un desarrollo concreto en el que esforzadamente se van aclarando las diversas facetas y del que prevalecen edificaciones apoyadas en numerosos contrafuertes. *El error es tomar cada publicación concreta como un acto independiente en el que tiene lugar una aportación definitiva a la ciencia.* Cada publicación tiene sentido como parte de un proceso del que se espera un corpus estable de conocimientos estructurados; este corpus partirá de esos elementos, pero no se reduce a una mera suma de ellos.

Por eso, desde el punto de vista de un autor (o incluso del de un equipo) no hay una relación biunívoca entre los resultados alcanzados y las publicaciones. Al comprender el hecho particular de la publicación científica como una mera parte de un proceso más general, se incluye más fácilmente en ese proceso lo que son las trayectorias investigadoras, tanto las individuales como las de los grupos de trabajo. La serie de publicaciones de éstos son puntos que marcan el camino seguido, pero no son su destino; sus contribuciones a la ciencia pueden rastrearse siguiendo esas marcas, pero pertenecen a un nivel más abstracto. En el ámbito individual, la delimitación y definición de una aportación puede abarcar varias publicaciones, lo que explica una buena parte de la progresiva facilidad para publicar. Esta es una buena razón para tomar los recuentos de publicaciones como un indicador de producción individual solamente después de aplicar una transformación de tipo logarítmico.

Para terminar con la no-linealidad de los recuentos como medidas de la producción científica, quizá convenga una pequeña aclaración. Enseguida se advierte la semejanza entre las consecuencias de adoptar la metáfora del descubrimiento científico como *extracción minera* y las de asumir que las publicaciones no son representaciones disjuntas de los resultados: en ambos casos parece surgir el *efecto filón*. Sin embargo, ambas hipótesis no son lo mismo. La primera se refiere a la naturaleza del descubrimiento, que se imagina de modo que un descubrimiento puede proporcionar en bloque resultados encadenados, pero diferentes entre sí. El solapamiento de las publicaciones, en cambio, tiene que ver

con el funcionamiento del sistema de publicación, que hace que fracciones del mismo resultado básico puedan ser compartidas por diferentes publicaciones del mismo autor o equipo. El autor del descubrimiento tiene una significativa ventaja para explotar la veta, pero no tiene garantizada la exclusividad, pues otros colegas pueden ser más perspicaces y adelantársele legítimamente. Por el contrario, el "derecho de solapamiento" sólo lo disfruta quien obtuvo los primeros resultados de la serie, por así decirlo.

En resumen, al enfrentarse con el análisis de la estructura de la productividad científica, aparece en primer lugar un hecho que sorprende al menos por su constancia: la asimétrica distribución de las publicaciones señalada por Lotka. La explicación más socorrida es la hipótesis de la ventaja acumulativa, que ha sido empleada recurriendo a varios mecanismos relacionados con la organización social de la comunidad científica. A lo largo del presente apartado se han cuestionado los compromisos básicos e implícitos en que se apoya la plausibilidad de la hipótesis de la ventaja acumulativa. La debilidad de estos presupuestos ha mostrado que otras alternativas explicativas no pueden ser rechazadas sin más, y especialmente que existen varios niveles distintos que han de ser tenidos en cuenta para explicar la distribución de la producción de publicaciones.

Lo más prudente es, seguramente, admitir que todas estas explicaciones de la asimetría en la publicación son en buena medida compatibles y por lo tanto parciales; probablemente ninguna sea suficiente para explicar tanto la distribución resultante como el proceso que da lugar a ella. La ventaja acumulativa podría, más bien, actuar temporalmente, sobre todo al principio de las carreras individuales, de modo que aumentaría la probabilidad de que aquellos que destacan se sitúen en los mejores lugares para la investigación, por ejemplo. Esto conllevaría también una ventaja inicial en cuanto a los recursos disponibles y al entorno intelectual. Pero, aunque no se puede negar la influencia de la cantidad de recursos, es bien cierto que no todas las disciplinas y temas son igualmente dependientes de ellos; por otra parte, los efectos de los recursos pueden ser progresivamente menores, como indica la ley de rendimientos decrecientes. Tampoco está clara la conexión entre el éxito y la aumento inmediato de los recursos disponibles; más bien parece que el aumento de producción se produce, una vez alcanzado un éxito, antes de que los recursos conseguidos gracias a él hayan podido tener

efecto. El aprendizaje es otro factor que ha de considerarse en la conducta publicadora de los investigadores. Que el descubrimiento científico proporciona muchas veces fructíferos filones es una sensación bastante generalizada, al menos mucho más que la que considera el descubrimiento como un proceso aleatorio que encuentra resultados independientes entre sí. Pero, sobre todo, es muy cuestionable la interpretación de los recuentos de publicaciones como indicadores lineales de la producción científica. La transformación de los recuentos brutos mediante una escala de tipo logarítmico no se apoya meramente en algo intuitivo, sino que es coherente con una concepción del proceso de publicación en el que las publicaciones particulares no son actos independientes, sino momentos necesarios para la construcción del conocimiento científico en un nivel superior que trasciende la mera suma de las aportaciones contenidas en los trabajos.

La aplicación de una transformación logarítmica sobre los recuentos brutos no siempre es necesaria. Es muy recomendable para valorar la producción de individuos o de equipos bien definidos, pero puede prescindirse de ella al comparar la producción de grandes agregados, como los países, ya que se supone que todos ellos tendrán una proporción parecida de grandes y pequeños productores. En todo caso, la mayor utilidad de esta discusión es el papel que juega en la delimitación del sentido de los recuentos, en particular, y en la aclaración de los fundamentos de los indicadores bibliométricos, ya que es una muestra de cómo la interpretación de estos indicadores no puede hacerse *a priori*, sino más bien después de un proceso de depuración crítica de los aspectos cuantificables que se tienen a mano.

#### 4.1.4.2. *Las variaciones entre disciplinas*

Otro rasgo sorprendente de los recuentos de publicaciones es que muestran a las claras significativas diferencias entre las disciplinas científicas en cuanto al número de publicaciones. En una primera aproximación, esto sugeriría la existencia de patrones de productividad muy dispares, dependientes del contexto disciplinar en el que se desarrolla la investigación. El hecho es que, aunque dentro de cada disciplina existen desigualdades como la descrita por la ley de Lotka, al comparar los

niveles publicadores de cada área se encuentran notables diferencias en el número medio de trabajos publicados por investigador, por ejemplo.

La existencia de estas diferencias invita a la reflexión sobre cuál pueda ser su origen. Es una cuestión de gran trascendencia en lo que se refiere al problema de la fundamentación de los indicadores bibliométricos. Uno de los usos que se espera de estos indicadores es el de evaluar comparativamente la investigación realizada en áreas diversas, por lo que ha de conocerse con certeza dónde están los límites de este tipo de usos. Conviene, por lo tanto, preguntarse por cuáles pueden ser las explicaciones que justifiquen lo que se presenta como notorias diferencias productivas entre las áreas científicas.

Puestos ante una serie de datos sobre las publicaciones de, por ejemplo, un país determinado, correspondientes a áreas diversas de las que se dispone además de información sobre el número de investigadores (lo que permite juzgar los datos con independencia del tamaño de la comunidad investigadora que trabaja en cada área) podría suponerse que las diferencias en el número de publicaciones reflejan diferencias correlativas de capacidad o de esfuerzo por parte de esos investigadores. Una primera explicación sería, entonces, que los investigadores pertenecientes a las disciplinas que más trabajos publican son más competentes que los de las que lo hacen en menor medida, es decir, la diferencia productiva se atribuiría a una mayor inteligencia, mejor preparación, superior esfuerzo, etc. de los investigadores de las áreas que más publican. Esta suposición no se puede rechazar sin más; Price, en otro contexto, (1963: 95-96) reproduce las cifras obtenidas por Harmon (1961) en las que se muestran diferencias notables en cuanto al nivel medio de inteligencia AGCT entre los doctores de algunas áreas, que sitúan a los físicos y matemáticos en el nivel superior (140 y 138 respectivamente) y a los biólogos y pedagogos en el inferior (126 y 123). Sin embargo, además de resultar un tanto extravagante, esta explicación sería inadecuada por la falta de correspondencia ya que, por poner un ejemplo entre otros posibles, las matemáticas son un área con pocas publicaciones mientras que la biología, por el contrario, tiene una elevada producción literaria.

Es posible buscar la explicación de las diferencias entre áreas apelando a la distinta naturaleza del objeto de cada disciplina. Quizá sea más fácil obtener resultados en unas disciplinas que en otras, quizá sus misterios son menos impenetrables, menos profundos o esquivos. También, siguiendo el mismo hilo, puede suponerse que la frontera de la

investigación esté más avanzada en algunas disciplinas, con lo que la dificultad para encontrar algo nuevo sea mayor en ellas. Sin embargo, recurrir a esta clase de justificaciones, además de apoyarse en una base oscura y casi imposible de poner a prueba, implica asumir una imagen de la tarea investigadora y del progreso científico de una muy dudosa consistencia interna. De ninguna manera puede extraerse de la historia de la ciencia la conclusión de que exista por sí mismo un ámbito de lo desconocido que disminuya como consecuencia del desarrollo científico: más bien parece que con cada nuevo descubrimiento haya aumentado el número de misterios; ni que el flujo de descubrimientos se haya ralentizado a medida que se conocen más cosas. Las consideraciones en este terreno o bien tienen un excesivo componente especulativo o bien se sitúan en controversias que no carecen de vitalidad.

Otro tipo de explicaciones de la diferencia publicadora entre las áreas puede acudir a factores que, en buena medida, pueden ser considerados coyunturales, en el sentido de que su manifestación en cada área es más accidental que necesaria. Por ejemplo, podría ser que los métodos en uso fueran más eficaces en unas áreas que en otras. La elección de los métodos que se emplean en cada área no puede seguir criterios de eficacia en abstracto, sino de conveniencia o de posibilidad de aplicación a la investigación en curso; sin embargo, no puede descartarse completamente que la difusión de nuevos métodos más potentes, aplicables en varias áreas, sea desigual entre ellas. De la misma familia, aunque más robusta, sería la aserción de que en cada área predomina un particular tipo de investigación, y cada uno de esos tipos lleva asociada de manera natural una productividad determinada. Es cierto que una orientación experimental puede ofrecer resultados más abundantes que una teórica, e incluso puede haber diferencias dentro de cada una de estas líneas. También, según los periodos, pueden agudizarse estas diferencias o bien cambiar las tornas, gozando la teoría de rachas fructíferas frente a lo experimental. Lo que sí puede decirse es que no es fácil, en la práctica, aplicar este tipo de distinciones sin toparse con actividades en las que las fronteras son borrosas o carecen de sentido. Pueden imaginarse otras hipótesis de este estilo: la diferencia podría residir en la generalidad de las disciplinas, de modo que aquellas más generales acabarían recogiendo algunos trabajos generados por investigadores de otras más particulares, por ejemplo. Pero en todos los casos es complicado pasar

de estos planteamientos intuitivos a comprobaciones prácticas, sobre todo porque se basan en conceptos un tanto nebulosos.

Los medios de los que se ha podido disponer es un factor al que ineludiblemente se ha de atender al preguntarse por las causas de la productividad. Las diferencias productivas entre las áreas científicas podrían estar causadas porque existan marcadas desigualdades en la cantidad de los recursos empleados por cada área. Ésta es, desde luego, una cuestión central en el terreno de la política de la ciencia, cuyas decisiones últimas, las más cercanas a la acción, consisten muchas veces en determinar la asignación de recursos dedicados a cada área. Pero lo primero que se encuentra al analizar los rendimientos de las disciplinas, si se estiman mediante el cálculo del "costo por publicación", es precisamente una firme constancia en las diferencias a través de las áreas. Esto indica que no se puede esperar que el número de publicaciones sea proporcional a los recursos empleados por cada disciplina, ya que cada una parece tener un costo inherente, propio, relacionado al menos parcialmente con la particularidad de sus métodos y problemas, de los instrumentos que requiere, etc. Ahora bien, si se logra una aproximación válida del costo medio en cada área, podrían afinarse más las evaluaciones y llegar a determinar si cada área ha empleado relativamente más o menos medios, por comparación con el patrón correspondiente, y verificar si el exceso o la carencia relativos explica las diferencias productivas. Esta es una línea cuya exploración y desarrollo podría conducir a resultados interesantes en el ámbito de la gestión de la ciencia. En todo, caso, no puede dejarse de lado el hecho de que la relación entre recursos y publicaciones está mediada por otros factores, por lo que es arriesgado llevar demasiado lejos esta simplificación.

Todos los intentos que puedan inscribirse en los tipos generales de explicación que se acaban de comentar tienen algo en común: asumir que el número de publicaciones es algo que tiene el mismo significado con independencia del área al que pertenezcan dichas publicaciones. Es decir, supone una aceptación acrítica de los recuentos como medida. Está claro que las unidades más preferibles y más útiles son aquellas independientes del contexto, las que no necesitan de traducción o transformación para conservar su significado; pero el deseo no configura la realidad. Las publicaciones no son unidades de esa clase. De nuevo, el problema procede de una concepción simplificada del proceso de publicación en la que cada documento es considerado el objeto de un acto

autónomo e independiente. Por el contrario, cada publicación adquiere sentido en el contexto de la disciplina a la que pertenece, formado por todas las publicaciones anteriores sobre la materia, por la comunidad de investigadores que la cultivan, etc. Es sensato pensar que el nivel de exigencia en los diferentes aspectos, el alcance del contenido, su forma de fundamentación, encuentran una concreción ligeramente diferente en cada recinto disciplinar, con lo que el esfuerzo necesario para producir un artículo puede variar de un campo a otro.

Las disciplinas constituyen los marcos, relativamente autónomos, en los que se edifica el consenso sobre qué sea una unidad elemental de novedad y de relevancia. La independencia entre unas y otras encuentra su causa en el relativo aislamiento derivado de la especialización. Además de las tradiciones o costumbres que se generan en cada área, las diferencias en la población de investigadores especializados, la competencia que existe entre ellos, la candencia de ciertos problemas, el número de revistas del área y cosas así definen dinámicamente qué se considera una unidad de novedad y de relevancia. La consecuencia de esto es que los patrones que se aplican en cada una, así como las expectativas generadas, son variables, con lo que se hace imposible la comparación directa de los recuentos brutos entre disciplinas. Más aún: también pone en cuestión el que se puedan comparar datos de la misma área correspondientes a periodos distantes entre sí. Si los recuentos de publicaciones adquieren sentido al concebirlos como sumas de unidades elementales de novedad y relevancia, es obvio que lo perderían si esas unidades se hubieran generado en marcos diferentes y autónomos.

Las unidades elementales de novedad y relevancia se parecen en este sentido a los pasos como unidades de distancia. En situaciones en las que es apreciado un cierto grado de precisión, es necesario contar con información suplementaria para poder emplear unidades como éstas, es decir, saber quién dio los pasos, pues no son lo mismo cincuenta pasos de un niño que los cincuenta pasos de un joven con prisa o los de un abuelo paseando. En las áreas científicas ocurre algo parecido, que cada una tiene su paso, si puede decirse así. Cuando se persigue una estimación general, puede admitirse que todos los pasos sean aproximadamente iguales; no así cuando se realiza un análisis de grano fino.

La existencia de disparidades en cuanto al número de publicaciones de cada área científica, una vez neutralizadas las diferencias de tamaño de las áreas, es algo que se conoce desde los primeros momentos de los

estudios cuantitativos de la ciencia. Han sido asumidas como diferencias naturales, inherentes a las disciplinas, como un hecho significativo, pero al que no se le ha prestado demasiada atención: no han abundado las justificaciones de las diferencias ni tampoco se han tenido en cuenta al presentar las producciones literarias de agregados en los que intervienen áreas diversas, como regiones o países. En niveles de agregación grandes, es posible que esto no suponga un peligro apreciable; pero si se comparan agregados de tamaño medio o pequeño, como universidades, cuyas especializaciones temáticas son poco semejantes, pueden producirse sesgos y equívocos notables. Un ejemplo: si un país está especializado en un área con un ritmo de publicación lento, y tiene un déficit relativo en un área de elevada fluidez publicadora, su peso en el concierto científico internacional quedará infravalorado, si las publicaciones de todas las áreas se toman como una unidad sin normalización previa. Este peligro se acentúa si además la base de datos empleada para el estudio tiene diferentes grados de cobertura para cada área.

#### *4.1.5. La cobertura de las bases de datos*

El principal problema práctico de los recuentos se refiere a la fiabilidad de las fuentes empleadas. Es impensable realizar un estudio bibliométrico de cierta envergadura manejando directamente las fuentes originales, es decir, las revistas y monografías científicas. Hoy en día, se han consolidado diversas bases de datos informatizadas que recogen la bibliografía de la ciencia, y son éstas las fuentes empleadas para llevar a cabo los estudios bibliométricos debido a las evidentes ventajas que comportan. Sin embargo, todavía no existe una base que recoja absolutamente todas las publicaciones firmadas por algún científico, y es dudoso que esté disponible en los próximos años: todas las bases bibliográficas son necesariamente selectivas en algún sentido. Esto da lugar a que puedan no ser adecuadas para cualquier tipo de estudio, lo cual es aprovechado por algunos críticos para enfatizar esta debilidad y atacar la validez de los recuentos por la dificultad práctica de conseguir recuentos completos de las obras producidas por tal o cual conjunto de investigadores.

La cobertura de las bases bibliográficas puede ser problemática debido a varias razones. La queja más radical posible procede de la

aspiración a una base completa, que recoja todos los escritos científicos. No es difícil demostrar que una muestra adecuada es suficiente para obtener datos fiables para un estudio bibliométrico, por lo que la completitud no es necesaria. Incluso podría ser perjudicial: es mejor contar con una muestra selectiva que garantice un cierto nivel y que disminuya la probabilidad de la repetición ocasionada por publicaciones en canales o circuitos científicos diferentes. Muchos trabajos son publicados bajo diferentes formas en canales distintos, como una evolución natural del trabajo, que en sucesivos perfeccionamientos y revisiones alcanza una presentación más definitiva en una revista internacional, por ejemplo. Por lo tanto, la incompletitud más que una limitación puede verse como una condición para que los recuentos sean fiables: es preferible una muestra que suponga una selección sobre criterios conocidos que el conjunto total de las publicaciones.

Un problema mayor de la cobertura de las bases de datos es el de los sesgos, especialmente los geográficos y lingüísticos. La consecuencia de estos sesgos es que algunas áreas culturales estén infrarrepresentadas respecto a otras, lo que da lugar a una imagen falsa del conjunto. Al emplear una base de datos es preciso tener en cuenta estos posibles sesgos y tratar con precaución los datos absolutos. Los riesgos de la interpretación disminuyen si, suponiendo que las inclinaciones se mantienen aproximadamente constantes a largo de un periodo, se consideran preferentemente los cambios relativos, la evolución, en lugar de las representaciones aisladas de una serie cronológica.

En las bases de datos multidisciplinarias existe, además, el problema de la representación relativa de cada área. No es fácil determinar si todas las áreas tienen una cobertura similar, y más difícil aún si los otros tipos de sesgos antes comentados son iguales para todas las áreas. Si el grado de cobertura es muy diferente en cada área, los recuentos brutos de agregados de áreas podrán reflejar diferencias irreales. Este problema ha de ser tenido en cuenta, aunque su solución es muy complicada: para obtener una representación más ajustada, sería necesario contar con datos fiables acerca del tamaño relativo de cada área (independientemente del de sus publicaciones) como por ejemplo de su población de investigadores; además se debería saber la productividad media de cada área, y estimar un parámetro que ponderara los recuentos de áreas diversas. Esas condiciones no se satisfacen con facilidad, por lo que al menos es

aconsejable ser precavido en la interpretación de agregados que incluyan la producción de áreas diversas.

#### 4.2. *El concepto de producción científica*

Hemos visto que la producción científica puede entenderse como la suma de unidades elementales de novedad y de relevancia. El concepto de producción científica se refiere a la *cantidad* de los resultados y no a esa otra dimensión, la *calidad*, que puede suponerse a falta de ulterior precisión relacionada con el valor final de éstos para la ciencia. La producción científica es un concepto ajeno a la práctica científica, en la cual se valoran integradamente numerosos aspectos. En ese contexto, es significativo hablar de contribución, pero ésta pertenece a un nivel distinto del de los resultados científicos. Aunque no debe llevarse demasiado lejos, puede ser ilustrativa la metáfora que compara la producción científica con la mena que se obtiene de una extracción minera: contiene mineral valioso, pero no se sabe exactamente cuánto. Cuanta más mena se extraiga, sin embargo, mayor será la probabilidad de obtener una mayor cantidad del mineral valioso.

El concepto de producción científica como la suma de los resultados alcanzados por un grupo o un individuo no es independiente de la existencia de los indicadores de producción. En cierto sentido se podría decir que la producción científica es exactamente lo que miden los recuentos de publicaciones una vez purgados convenientemente: este concepto se ha configurado a partir del proceso de refinamiento de las anomalías encontradas en los indicadores disponibles (recuentos), en el proceso de interpretar adecuadamente unos valores numéricos cuyo sentido se intuía. Esta es una cuestión común a todos los indicadores bibliométricos, que no pueden escapar de las limitaciones impuestas por su procedencia: la literatura científica no tiene como función principal servir como base para la evaluación externa de la ciencia. Su papel central en ese sistema la convierte en una vía insustituible y llena de virtualidades para este fin, pero eso no significa que tenga que ajustarse a una serie de conceptos predefinidos en otros ámbitos.

### 4.3. Indicadores bibliométricos de la producción científica y sus aplicaciones.

**Número de documentos.** Consiste, obviamente, en contar las publicaciones científicas pertenecientes a un autor, institución, región, país, etc. Es la aplicación directa de los recuentos brutos de documentos como indicador bibliométrico. A pesar de sus inconvenientes, es ampliamente empleado, debido a la sencillez de su obtención y seguramente también a su aparente objetividad y a la facilidad de su interpretación: lo que se cuenta son publicaciones, objetos concretos, y las cifras dicen simplemente cuántas hay. Ya se ha visto, sin embargo, que estas cifras pueden verse afectadas por muchos factores que invalidan una interpretación tan ingenua. En realidad, los números de publicaciones deben considerarse más bien como la base para el cálculo de otros indicadores de producción, y su utilización como indicadores debe restringirse a los contextos en los que esté garantizada la comparabilidad por la ausencia de sesgos significativos.

**Equivalentes a documentos completos.** Son el resultado de efectuar recuentos con algún tipo de fraccionamiento. Responden a la dificultad planteada por la existencia de resultados obtenidos mediante la colaboración, hecho que se refleja en la firma conjunta de los documentos. Existe un abanico de posibilidades para efectuar una atribución fraccionada; de éstos, algunos pueden ser más adecuados que otros en cada caso particular. La mayoría de ellos hace que la suma de las producciones de todos los agregados sea exactamente la producción del sistema que se esté considerando. Esto aminora la confusión que pueden generar las repeticiones en los recuentos, es decir, la contabilización múltiple del mismo documento debido a la colaboración, así como los efectos ilusorios de aumento de producción debidos a aumentos en la colaboración. Sin embargo, no neutraliza otros sesgos, como los originados por la diferente productividad en cada área.

**Solidez.** Consiste en realizar una transformación de tipo logarítmico sobre los números obtenidos por alguno de los anteriores tipos de recuento. Resulta adecuado para valorar la producción bibliográfica de individuos o pequeños equipos, nivel en el que las diferencias en publicaciones no parecen reflejar diferencias proporcionales en cuanto a resultados científicos. Fue propuesto por Price (1963) al interpretar la distribución de la productividad señalada por Lotka. Asume el

presupuesto de que los recuentos de publicaciones no son un reflejo lineal de la producción de los autores, razón por la cual es necesaria la transformación logarítmica. Este procedimiento aminora la influencia de las diferencias entre áreas, por lo que también mejora la comparabilidad de datos relativos a disciplinas diversas.

**Percentil productivo.** En lugar de emplear las cifras obtenidas de los recuentos, puede sacarse provecho de la información en cuanto al orden que establecen. Un modo de normalizar las listas de producción es expresar la posición de un individuo diciendo, en lugar del número de documentos o fracciones que le corresponden, el porcentaje de investigadores a los que supera, es decir, que están por detrás de él en cuanto a número de publicaciones. Emplear la posición relativa en lugar de los recuentos permite realizar comparaciones entre áreas salvando las diferencias productivas entre ellas. Hay que tener en cuenta que, debido a la asimetría en la producción, ascender posiciones cuesta menos artículos en la parte baja que en la alta. Este indicador sería adecuado para poblaciones de investigadores perfectamente identificadas, y esto no siempre es posible en la práctica.

**Porcentajes en el área.** Una sencilla normalización de los recuentos, sean completos o fraccionados, consiste en calcular el porcentaje que representan respecto al marco disciplinar. Es una buena aproximación del peso en cada área y salva el problema de las diferencias productivas. Cuando hay que comparar unidades o instituciones de áreas diferentes, es preferible a los recuentos directos. También refleja mejor la evolución, pues es menos sensible a los cambios de costumbres o de cobertura de las bases de datos. Una precaución: puede ser engañoso al compararlos con el peso en los recursos, porque la producción no es necesariamente lineal respecto a los recursos empleados, sino que existen factores que determinan una curva decreciente en los rendimientos: la ciencia puede ser progresivamente más cara (véase más adelante el comentario sobre la eficiencia).

Pueden distinguirse varios indicadores basados en el cálculo del porcentaje que representa un recuento particular respecto al recuento general: podemos llamar *contribución* al que emplea algún recuento fraccionado sin repetición; *presencia* al que utiliza la asignación completa, es decir, el recuento sin fraccionamiento y con repetición, y *encabezados* al porcentaje de documentos en los que se es primer autor o se figura en primer lugar, es decir, el recuento sin fraccionamiento y sin

repetición. Como ejemplo, en la siguiente tabla se muestran las diferencias en estos indicadores para la producción científica española en el área de Físicas recogida en la base de datos SCI (Maltrás, Quintanilla, 1995). Las diferencias en las tasas de crecimiento de cada indicador reflejan cambios en la colaboración internacional de España.

	1986-87	1988-89	1990-91	% variación
presencia	1,37	1,61	1,82	32,85
contribución	1,1	1,27	1,36	23,64

También pueden aplicarse estos indicadores a la producción completa, pero deben ser tomados como estimaciones aproximadas debido a las diferencias productivas entre áreas y a la cobertura desigual de las bases de datos.

**Especialización temática.** Comparar la distribución de la producción por áreas de una institución o país con un marco general adecuado puede ser útil para detectar fortalezas o debilidades relativas. Esta comparación puede realizarse de varios modos. Un modo común es el de calcular la fracción entre los porcentajes que representa un área o disciplina en una unidad determinada y en el total considerado: si es inferior a 1, la especialización en esa área es negativa, si es superior, es positiva, es decir, refleja un peso mayor de lo esperado. Esta idea puede aplicarse de distinto modo: en lugar de calcular una simple fracción del tipo observado respecto esperado ( $O/E$ ), pueden calcularse otros indicadores más expresivos:  $(O-E) / E$ , cuyo valor, multiplicado por 100, indica el porcentaje que supone el defecto, si es negativo, o el exceso, si es positivo, de lo observado en relación a lo esperado; o también:  $\max(O, E) / \min(O, E)$  (con signo positivo si  $O > E$  y viceversa, es decir, multiplicado por  $|O-E| / (O-E)$ ), que indica el número de veces que hay que multiplicar, si es negativo, o dividir, si es positivo, el valor observado para obtener el esperado.

### Un comentario sobre la eficiencia de los sistemas científicos.

Los indicadores de producción científica juegan un papel central en el enfoque que describe los sistemas científicos sobre la base de un esquema entrada-salida. Si las entradas son, generalmente, los recursos consumidos en la investigación, las salidas se suelen medir en términos

de publicaciones científicas, como indicadores de los resultados o de los conocimientos.

La noción de *eficiencia* sirve para valorar el intercambio que un sistema ha tenido con su entorno, mediante la comparación entre ciertas magnitudes de salida y las correspondientes de entrada. En general, expresa la proporción que ha supuesto *lo que se ha obtenido* del funcionamiento de un sistema determinado respecto a *lo que ese sistema ha empleado* para producirlo, es decir, a su consumo:

$$\text{eficiencia} = \frac{\text{Producido}}{\text{consumido}}$$

Esta noción se usa en varios ámbitos. La eficiencia termodinámica es la proporción de la energía que un sistema entrega respecto a la que consume. Su aplicación en ingeniería se refiere a las máquinas y motores: un motor o máquina es más eficiente cuanto más se acerque a 1 la proporción entre el trabajo ejecutado y el aplicado. El principio de entropía indica que ninguna máquina puede conseguir un rendimiento igual a la unidad, ya que el funcionamiento propio de la máquina consume energía, en la transformación se pierde energía.

La eficiencia económica tiene que ver con el costo de producción: la eficiencia aumenta cuando se producen más unidades con el mismo costo o cuando cuesta menos producir lo mismo, es decir, cuando se reduce el costo por unidad producida.

Quintanilla (1988) propone una generalización de la noción de eficiencia, aplicable a la evaluación de los sistemas técnicos, basada en una relación entre los objetivos ( $O$ ; aquello que se quiere conseguir) y los resultados ( $R$ ; lo que efectivamente se consigue):

$$E = \frac{(O \cap R)}{(O \cup R)}$$

Mediante esta definición se expresa la importancia relativa de los resultados *no buscados*: una tecnología es tanto más eficiente cuanto se consiguen más objetivos de los propuestos con un mínimo de consecuencias colaterales. La eficiencia máxima es, por lo tanto, la coincidencia plena de los objetivos y los resultados obtenidos. Obsérvese que el numerador es una medida de la *eficacia*: ésta es máxima cuando se consiguen todos los objetivos propuestos, es decir, cuando  $(O \cap R) = O$ .

De las anteriores variantes, la que parece ofrecer una implementación más sencilla es la de rendimiento económico. Al considerar qué cosas pueden tomarse como entradas y como salidas dignas de atención al describir un sistema científico, pueden plantearse varias posibilidades. Lo habitual es, al exponer los modelos generales de la ciencia como sistema *input-output* (p.e. Martin, Irvine, 1980: 64) incluir entre las entradas tanto los recursos materiales empleados y el esfuerzo del personal implicado como los conocimientos de los que se parte. Entre las salidas, se destaca el nuevo conocimiento producido, y a veces las aplicaciones más directas en la industria o en la sociedad, o los beneficios económicos generados. Podría intentarse, por ejemplo, medir la eficiencia científica solamente en términos de producción y consumo de conocimiento, como lo más característico de estos sistemas; pero no se vislumbra un modo práctico de llevar a cabo esa idea. Por un lado no se sabría exactamente qué conocimientos se han empleado; por otro, al ser estos conocimientos en principio públicos, accesibles a todos, no serviría para establecer diferencias entre unos sistemas y otros.

Es mucho más sencillo limitarse a valorar el *costo* de las actividades científicas y compararlo con una medida de los *resultados*. La eficiencia de un sistema científico, por lo tanto, puede establecerse claramente como la proporción que representan los *resultados* obtenidos respecto a los *recursos* empleados. Será más eficiente cuanto menor sea el costo por unidad de resultados producida. Una medida satisfactoria de los resultados puede obtenerse mediante los indicadores bibliométricos de producción que se acaban de comentar. En cuanto al costo de las actividades, aunque plantea aún ciertas dificultades, existe un amplio consenso internacional en cuanto a las normas que se han de aplicar, gracias al esfuerzo conjunto de expertos de diversos países en la elaboración de las recomendaciones de la OCDE, que constituyen el llamado *Manual de Frascati* (la última revisión es de 1993). Así, parece que disponemos de una medida apta para comparar distintos sistemas científicos.

Pero todavía hay que notar que la eficiencia así definida no depende solamente de la capacidad y el esfuerzo de los agentes. Por un lado se reconoce que existen factores contextuales que delimitan la amplitud de los márgenes del rendimiento. Especialmente importantes son las diferencias encontradas entre disciplinas científicas, cuyo coste por unidad de resultado alcanzado puede variar de un modo significativo de una a otra: conseguir un resultado es más caro en unos campos que en otros,

y esta diferencia se deriva del tipo de problemas, métodos y técnicas implicados y del instrumental necesario en cada caso. También pueden ser influencias decisivas el lugar de trabajo (el tipo de institución) o el momento histórico. Al señalar estas disparidades se quiere alertar sobre el peligro de un uso inapropiado de los valores del rendimiento: su empleo para evaluar los agentes tiene que ser consciente de estos factores contextuales, que deben ser neutralizados en lo posible. Es inadecuado comparar las medidas brutas de eficiencia de dos sistemas cuya especialización temática es muy diferente; y arriesgado hacerlo cuando esa especialización se desconoce.

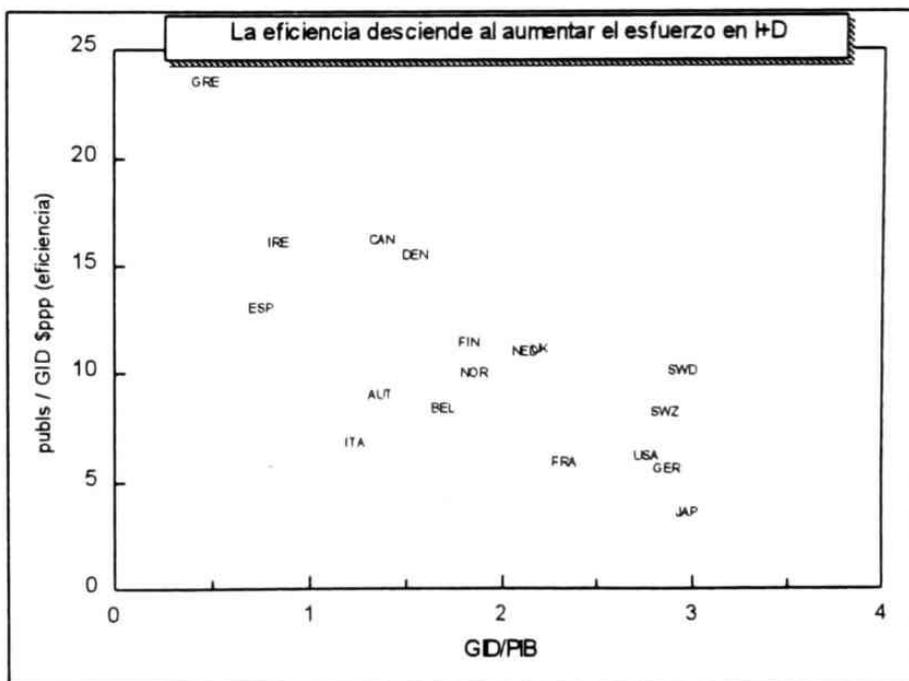
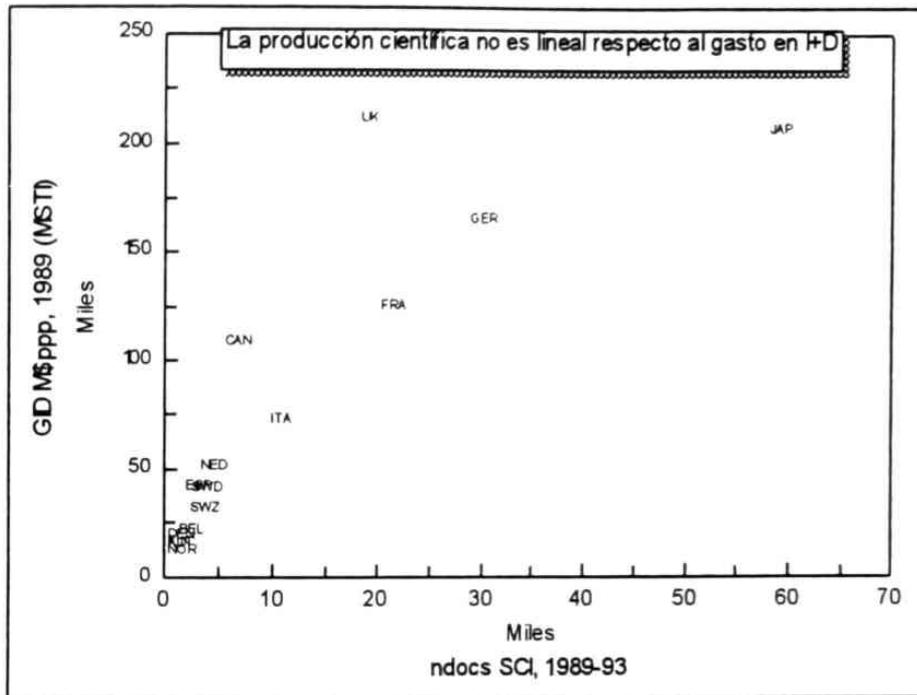
Además de estos factores *contextuales*, han de considerarse otros que pertenecen más bien a la *estructura* misma de la eficiencia, no sólo de los científicos, sino también de otras actividades humanas. Seguramente la proposición más conocida acerca de la eficiencia es la *ley de rendimientos decrecientes*, cuya primera enunciación clara se atribuye al economista clásico David Ricardo (aunque había sido ya anticipada implícitamente por algunos predecesores). Lo que esta ley viene a decir es que, a partir de un cierto punto, el aumento de los factores productivos no va acompañado de un aumento proporcional, sino progresivamente menor, de lo producido. Esto significa que existe un límite desde el que la elevación de los factores productivos provoca una disminución del rendimiento, ya que la producción no crece en la misma medida<sup>41</sup>. En la ciencia parece ocurrir exactamente lo mismo cuando se estudian, por ejemplo, los sistemas nacionales. Los países con sistemas de investigación de un mayor tamaño relativo obtienen rendimientos científicos menores, o dicho de otro modo, parece costarles una inversión progresivamente mayor cada unidad de crecimiento en producción científica. La razón que puede explicar este comportamiento es bastante simple: la ciencia es más cara cuanto más avanzada; cada nuevo descubrimiento requiere material e instrumentos más complejos y caros, personal más formado y especializado. Pero esto se olvida cuando se manejan indicadores de rendimiento de un modo poco reflexivo, y pueden llegar a extraerse conclusiones casi opuestas a las sensatamente deducibles, tomando como indicio de una "mayor eficiencia en ciencia" lo que más bien es síntoma de "menos desarrollado en ciencia"<sup>42</sup>. Para reducir este

<sup>41</sup> En una situación tecnológicamente estacionaria, es decir, mientras no se produzcan cambios tecnológicos que alteren la productividad.

<sup>42</sup> Afirmaciones como las de Pestaña (1996: 10) acerca del *satisfactorio* rendimiento del

inconveniente al evaluar el rendimiento de los agentes, es aconsejable tener en cuenta alguna medida complementaria acerca del nivel relativo de recursos disponibles por éstos. Lo más importante aquí es que la comparación entre las medidas de eficiencia no debe realizarse directamente entre ellas, sino respecto al valor esperado por una *función del rendimiento* que estime el valor esperado en cada estadio de un determinado modelo de sistema científico.

Los siguientes gráficos ejemplifican las anteriores consideraciones. A partir de los datos de gastos en I+D de los países desarrollados en 1989, ofrecidos por los *Main Science and Technology Indicators* de la OCDE, y de las publicaciones científicas de esos países recogidas en el SCI durante el período 1989-93). En términos generales, ya que existen sin duda otros factores que hay que considerar (tipo de sistema científico, especialización, etc.) muestran con poco espacio para la duda la existencia de una función de rendimiento que es necesario considerar para valorar los valores de la eficiencia de los sistemas científicos.



Otro aspecto interesante en la interpretación de las medidas de rendimiento es el de la *masa crítica*, es decir, la necesidad de un número

mínimo de recursos y de personas dedicadas a algo para empezar a obtener rendimientos aceptables. Esto se traduce en rendimientos significativamente inferiores a "lo normal", o incluso nulos, cuando se está por debajo de la masa crítica correspondiente. Se intuye desde hace bastante tiempo que éste es el caso de la ciencia, que para que un grupo de científicos se constituya en una comunidad científica (aunque sea local) tiene que estar formado por un número suficiente de investigadores y personal de apoyo, y disponer además de unos recursos y una infraestructura proporcionados. Si esto fuera así, los indicadores de eficiencia podrían ser en ocasiones de gran utilidad para detectar el grado de constitución de una comunidad, y en estos casos debería evitarse su uso en la evaluación puesto que su estimación estaría sesgada precisamente por estas condiciones estructurales de los agentes.

Otorgar un papel central a la noción de rendimiento al estudiar cuantitativamente un sistema científico supone una concepción *maquinal* de la ciencia. Concebir el sistema científico como un motor mecánico puede ser útil en ocasiones, pero no tenemos ninguna razón para pensar que esta imagen agote la descripción de la ciencia ni que sea la única disponible<sup>43</sup>. En todo caso, hay que mostrarse precavidos al emplear los indicadores de producción para evaluar la *eficiencia*, ya que existen como hemos visto una serie de factores ajenos al empeño de los agentes que lo pueden determinar, y ser a la vez más ambiciosos, pues este uso tradicional no agota las posibilidades de estos indicadores que pueden ofrecer pistas decisivas para saber algo más acerca de la estructura de la ciencia.

---

<sup>43</sup> Como simples sugerencias, un modelo orgánico o uno ecosistémico podrían abrir perspectivas interesantes en el estudio de la ciencia. Gell-Mann (1994) considera a la ciencia como un caso de "sistema complejo adaptativo".

## Capítulo 5

### *El problema acerca del concepto de calidad científica*

---

#### *5.1. La calidad científica y los indicadores bibliométricos*

La búsqueda de medidas satisfactorias acerca de la calidad de la ciencia producida por los científicos es una de las tareas a las que se ha dedicado un mayor esfuerzo en el ámbito de los estudios cuantitativos sobre la ciencia. Es fácil comprender el interés que ha guiado ese afán. No es necesario, por lo tanto, insistir en el enorme beneficio que supondría la capacidad de distinguir qué ciencia es "mejor" o de saber quién realiza las contribuciones "superiores", más interesantes, o con más futuro, de un modo relativamente barato y con un elevado grado de objetividad, a salvo de sesgos interesados o de preferencias subjetivas. Los recursos requeridos por la actividad científica proceden de sistemas externos a la ciencia; esos indicadores de la calidad permitirían aumentar tanto la justicia de la distribución de los fondos, como la eficacia en el uso de esos recursos.

La necesidad de disponer de indicadores bibliométricos *de calidad* puede verse también como un efecto de las limitaciones intrínsecas de los indicadores bibliométricos de producción para valorar de un modo completo los resultados de la investigación de los científicos, capaz de discriminar, al menos en cierto grado, la importancia o el interés de

unas contribuciones y otras. Los indicadores bibliométricos de calidad recuperan algunos de los elementos que inevitablemente soslayan los indicadores de producción en el proceso de reducción que conlleva la cuantificación. Algunas de las diferencias que particularizan a los trabajos son así tenidas en cuenta como indicios para construir otras medidas acerca de grupos de documentos.

Impulsa la búsqueda de indicadores de calidad, concretamente, la carencia de una respuesta adecuada ante la duda que se plantea, por un lado, al enfrentarse a dos medidas de producción idénticas: ¿debemos valorar del mismo modo dos producciones científicas cuantitativamente iguales, o pueden representar por el contrario contribuciones a la ciencia de importancia o interés muy diferentes? En este caso se presenta de manera obvia que pueden existir características diferenciadoras entre los dos conjuntos a los que se refieren los indicadores de producción. Pero el mismo problema se presenta al pretender comparar producciones que difieren mucho en cuanto a su tamaño, como por ejemplo las de dos países en una misma área temática: es concebible que deban existir algunos aspectos sobre los que establecer esa comparación independiente de la magnitud de ambas, que permitan decir en algún sentido que ésta es *mejor* que la otra. Tanto la necesidad de diferenciar lo que es cuantitativamente igual como la de lograr comparaciones que trasciendan las diferencias de dimensión delimitan el sentido de los indicadores bibliométricos de calidad, sentido que puede resumirse en que la valoración de distintos conjuntos de resultados científicos debe considerar la cantidad, pero también su calidad respectiva; o, en otras palabras, que cantidad y calidad requieren medidas que son, en principio, independientes; parece plausible pensar además que son complementarias en cuanto medidas de los resultados científicos.

De este modo, si al referirnos a la producción científica de un agente o un agregado decimos que es *mayor* o *menor* que la de otro, al mismo tiempo podremos hablar de su calidad y decir si ésta es *superior* o *inferior*. Pero los indicadores bibliométricos de producción y calidad no agotan, juntos, la descripción de las características de los trabajos individuales que componen el conjunto al que se refieren esos indicadores; éstos sólo sirven para apoyar una valoración del conjunto de trabajos como tal.

El empleo del término *calidad* para designar otra dimensión que no es captada por los indicadores de producción o cantidad requiere una

explicación previa. Recurrir a ese término en el ámbito de los indicadores bibliométricos encuentra una justificación precisamente en que su vaguedad es adecuada para expresar la amplitud de significados que en ellos puede adoptar, no completamente ajenos a los que se despiertan al hablar de la *calidad interna*. Por este motivo, puede resultar útil ocuparse de los usos y significados del término "calidad" en el ámbito de la ciencia, cuando menos para aclarar las relaciones que puedan existir entre algunos de esos significados y las interpretaciones de los indicadores a los que aplicamos el rótulo "de calidad".

Como todo producto de una actividad humana, la ciencia puede estar mejor o peor hecha, lo que en general significa que puede ajustarse más o menos a una serie de criterios, normas o expectativas, o que su "funcionamiento" ofrezca unos frutos considerados más o menos valiosos. Esto implica que, incluso dentro de la especificidad de la ciencia, puede hablarse de la calidad de un trabajo científico concreto desde una multitud de perspectivas e intereses, lo que conlleva una variedad de usos, contextos y significados del concepto de *calidad científica*.

El problema de definir qué sea la calidad en la ciencia sólo puede afrontarse entonces de un modo formal, ya que ésta parece asumir una amplia diversidad de sentidos dependiendo de cada caso concreto. Por lo que hace a lo que nos viene ocupando, conviene tratar desde un punto de vista general de qué modo puede concebirse la calidad en la ciencia y las implicaciones que esa concepción pueda tener en cuanto a sus efectos, en especial sus trazas o pistas medibles, lo que nos introducirá en las cuestiones acerca de la interpretación y el uso de los indicadores que mantengan con ella una cierta relación. En esta aproximación general a la delimitación del ámbito de la noción de calidad en la ciencia será inevitable mostrar los solapamientos que a veces se han producido entre esta noción y otras muy íntimamente conectadas, como son las de contribución, importancia, progreso o aplicabilidad, entre otras; puestos ante la dificultad de ofrecer una lista extensiva de los rasgos o sentidos de la calidad científica, sabremos apreciar la vía que se nos ofrece para delimitar de modo más bien intensional esta noción mediante su contraposición con los conceptos antes mencionados y otros próximos. Aclarar las relaciones que constituyen esa red de conceptos ayudará a concebir de una manera más precisa el papel que cada uno de ellos puede jugar en nuestra valoración externa de los resultados científicos. En el presente capítulo, por lo tanto, se comentarán los aspectos más

destacables de los principales hitos en el esfuerzo por la medida de la calidad.

En el siguiente, se reflexionará sobre el modo de articular los principales puntos de ese debate, con el propósito de poner de manifiesto las condiciones de constitución de un concepto sobre la calidad científica. Una propuesta en el modo de entender la calidad de un modo coherente con las posibilidades de los indicadores será el corolario. Por último, se tratarán algunas cuestiones, más bien de índole técnica, relacionadas con la posibilidad práctica de construir determinados indicadores de calidad, en particular aquellos que persiguen responder a los problemas reales planteados por la política científica.

## 5.2. La noción de "calidad" en los estudios cuantitativos de la ciencia

La preocupación por la calidad en la ciencia no está separada, como ya se ha adelantado, de las medidas acerca de la cantidad de las contribuciones científicas. De hecho, podemos situar, quizá más que simbólicamente, el comienzo del protagonismo de la noción de *calidad científica* en las advertencias (Price, 1963: 81 y 109) que Price considera necesario realizar en el curso de sus análisis cuantitativos de la estructura y el desarrollo de la ciencia contemporánea: en ellas reconoce implícitamente que la medida de las contribuciones a la ciencia no puede reducirse a los recuentos de publicaciones. En coherencia con su prefacio en esa misma obra, en el que compara el enfoque estadístico del estudio cuantitativo de la ciencia con el que sigue la termodinámica al estudiar el comportamiento de un gas (del que interesa el comportamiento global y no el individual de cada molécula), Price llama la atención sobre las diferencias que existen entre los trabajos individuales en cuanto a su importancia para la ciencia. Sin embargo, una vez admitidas esas diferencias, Price señala en las mismas páginas que existe una correlación "razonablemente buena" entre la *productividad* de los investigadores (medida en número de publicaciones) y la *eminencia* que cada uno ha alcanzado (estimada por la inclusión en listas seleccionadas de científicos).

Lo verdaderamente significativo es que, aunque él no emplea directamente el término, la dicotomía se impuso pronto, y se extendió como algo natural la necesidad bien de complementar en el estudio estadístico

de la ciencia las medidas de *cantidad* con otras de *calidad*, bien la de reforzar la hipótesis de que las segundas eran innecesarias.

Lo que sucedió en los años siguientes merecería una reconstrucción histórica más detallada que la que aquí conviene abordar. Para nuestros fines, será suficiente con resaltar los rasgos más relevantes del proceso por el que, a través de una compleja interacción de factores y circunstancias, se generó una importante confusión en torno al concepto de *calidad científica*. Una de las circunstancias más influyentes fue, seguramente, la conjunción de la euforia inicial de los emergentes estudios cuantitativos sobre la ciencia con la disponibilidad, gracias a la entonces reciente constitución del repertorio bibliográfico *Science Citation Index*, de datos sobre las citas recibidas por cada trabajo. Puestos al lado de los recuentos de publicaciones, los recuentos de citas se presentaban como el complemento perfecto de aquellos, como el reflejo evidente y directo de una evaluación democrática y desinteresada de los trabajos científicos por parte de los especialistas que los realmente utilizaban. De este modo, se originó una serie de trabajos dedicados a comprobar la validez de los recuentos de citas como indicadores de calidad, relacionándolos con otras medidas independientes, o por el contrario otros que señalaban algunos inconvenientes en el uso de tales recuentos para determinar la calidad de los trabajos.

Desde la ventaja que nos da la perspectiva temporal podemos pensar que, en aquellos años, la calidad había sido en realidad definida tan sólo negativamente, es decir, como aquello que marca diferencias que pueden escapar a la mera cuantificación, aunque quizá entonces esto no era muy conscientemente asumido por los investigadores. Más bien parece que se pensaba que existía un consenso implícito sobre qué era la calidad científica, que cualquier científico competente podría distinguir en su especialidad los trabajos buenos de los malos o clasificar uno como mejor que otro. Cabe decir sobre este punto que desde el principio se había asumido que la calidad científica debía estar estrechamente ligada a la importancia de la contribución, en especial siguiendo el modelo de la producción individual. Aunque la naturaleza de tal ligazón no estaba especificada con detalle, ni tampoco estaba clara la noción de *importancia científica* de la contribución, la suposición de que básicamente se estaba de acuerdo permitía contentarse con el uso intuitivo que podía derivarse de tal indefinición. Como consecuencia de este hecho, no se sentía la necesidad de ofrecer una definición explícita, y la existencia de

usos variados de la noción de *calidad científica* con diferentes matices podía pasar inadvertida.

La sucesión de trabajos acerca de los análisis de citas y su interpretación conformó un proceso de depuración en el que tuvieron lugar algunas propuestas de conceptos y de distinciones en relación a la contribución científica. Este proceso permitió pasar de defender o atacar las citas como medida de la dimensión cualitativa a delimitar progresivamente el contenido de lo que las citas podían medir. Veamos una selección de trabajos que puede servir para ilustrar los principales hitos y propuestas en la discusión acerca de la medida de la calidad en la ciencia.

### 5.2.1. *Los pioneros: calidad como prestigio*

El estudio llevado a cabo por Kenneth E. Clark (1957) acerca de los psicólogos se convirtió pronto en una referencia clásica en el debate. Clark, tras preguntar a un panel de expertos por los psicólogos que habían realizado las contribuciones más significativas, comparó este resultado con otras medidas independientes. La que mostraba una correlación más alta era precisamente el recuento de citas a la obra, por lo que concluyó que era el mejor indicador disponible acerca del valor de la investigación realizada por psicólogos (Cole, Cole, 1967; 1968; Cole, 1970; Cole, Cole, 1971). Podemos tomar este trabajo, además de como iniciador de una línea de pensamiento en la que parece asumirse que la calidad del trabajo realizado determina la valoración social o la eminencia alcanzada por los científicos entre sus colegas, como el principio de una serie de esfuerzos encaminados a defender la validez de los recuentos de citas como indicadores de la calidad. La calidad es igual a la significación o importancia alcanzada en la ciencia por la obra, lo que implica entre otras cosas que su determinación depende de un juicio *a posteriori*, cuando la obra ha sido suficientemente difundida entre los especialistas correspondientes y tras el tiempo suficiente para "digerirla".

En la misma línea se sitúan las aportaciones de los hermanos Jonathan y Stephen Cole (Cole, Cole, 1971). Como miembros de la escuela de sociología de la ciencia que arranca de Merton, conciben el hábito de citar como uno de los mecanismos fundamentales por los que la

comunidad científica otorga reconocimiento a aquellos investigadores que han realizado contribuciones significativas. Su hipótesis supone que este reconocimiento tiene un grado suficiente de justicia y que se basa en el juicio que merecen los trabajos publicados por el autor (puestos por él a disposición de la comunidad); la conclusión es que, incluso admitiendo la existencia de algunos sesgos y factores distorsionantes, el número de citas es un reflejo satisfactorio, aunque aproximado, de la calidad de los trabajos. Puede decirse que se quedaron a un paso de operacionalizar el concepto de *calidad científica* y definirlo como "aquello que es medido por el número de citas" (lo cual no habría sido nada descabellado, porque en aquel momento la calidad no estaba definida de ninguna manera). Pero seguramente pesó la tradición, y en un principio al hablar de la *calidad* se referían de modo genérico a la *significación científica*. Más tarde (Cole, Cole, 1973), ante la lista de factores que podrían desviar al sistema formal de citación de su funcionamiento ideal como distribuidor de reconocimiento basado en los méritos científicos, aportaron la interesante distinción entre la *calidad absoluta* y la *calidad socialmente determinada*<sup>44</sup>.

La separación de estas dos dimensiones en la calidad implica, al mismo tiempo, la existencia de rasgos objetivos que determinan si un trabajo tiene o no calidad y la posibilidad de que éstos no sean reconocidos adecuadamente por la comunidad científica. Los recuentos de citas sirven, según los Cole, para medir lo que denominan *calidad socialmente determinada*. Aunque es patente la influencia que tuvo el empeño por probar la validez de los recuentos de citas como indicador de calidad, lo cierto es que su distinción apunta uno de los problemas básicos en cuanto a los juicios sobre la calidad científica, el conflicto entre los polos objetivo-subjetivo (en este caso, lo subjetivo sería la valoración social, la de los colegas, que dependería de la capacidad de éstos para captar los rasgos objetivos y de algunas circunstancias particulares que podrían afectar a esta capacidad).

También inscrito en la escuela mertoniana y relacionado con los esfuerzos por definir la calidad, pero en otra línea de pensamiento, se sitúa el trabajo de J.M. Chase sobre los criterios que se emplean como filtro en la publicación científica (Chase, 1978). Además de postular como necesarios y básicos los imperativos institucionales de Merton,

<sup>44</sup> Ya habían escrito (Cole, Cole, 1972: 369) que el número de citas puede ser considerado en general un auténtico indicador de la *influencia*.

recopila en una lista diez criterios sobre los que se juzga la calidad científica del escrito antes de ser publicado: *rigor lógico*; *reproducibilidad* de las técnicas de investigación; *claridad y concisión* en la redacción; *originalidad*; *precisión matemática*; *cobertura de la literatura* significativa; *compatibilidad con la ética* aceptada en la disciplina; *significación teórica*; *pertinencia* con la investigación del momento en la disciplina; y *aplicabilidad* en problemas prácticos en el área. Así reúne algunos de los esfuerzos anteriores por especificar el contenido de lo que se llamaba genéricamente *calidad científica*.

Como cualquier enumeración, ésta puede ser más o menos completa o adecuada, y suscitar controversias o sugerencias en relación con la inclusión, la exclusión o con la ordenación de criterios; lo interesante, sin embargo, es que representa un punto de vista acerca de la naturaleza de la calidad en la ciencia. Esta aproximación abre varias posibilidades: permite concebir que la calidad está compuesta por una serie de virtudes o rasgos que pueden poseerse en distinto grado; o, por otro lado, que todas o algunas de ellas son condiciones necesarias y/o suficientes para atribuir calidad científica a un escrito. Aunque no se concrete cuáles han de ser los procedimientos para aplicar estos criterios con vistas a ordenar una serie de trabajos por orden de calidad (hay que recordar que el trabajo de Chase está orientado a definir el mínimo aceptable), ni tampoco la justificación para la inclusión o exclusión de criterios (que parecen presentarse como atemporales), puede imaginarse la posibilidad de llegar a una convención en cuanto a escalas y ponderación del peso de cada criterio. Este tipo de concepción no presupone la variabilidad en cuanto al grado de calidad, pero tampoco la impide. Además, funda la posibilidad de un juicio sobre la calidad *a priori* de los efectos que pueda tener el trabajo en la ciencia, lo cual no implica necesariamente que considere irrelevantes estos efectos en cuanto al juicio sobre la calidad.

### 5.2.2. Esfuerzos de precisión conceptual

En un similar contexto problemático sobre la medición cuantitativa de la contribución científica, M.J. Moravcsik propuso (Moravcsik, 1973) la separación semántica de tres cantidades: *actividad*, *productividad* y *progreso*. En sus palabras,

Bajo circunstancias inusualmente favorables, las tres pueden coincidir, pero la mayor parte del tiempo no lo harán (Moravcsik, 1973: 268).

Su objetivo declarado es precisar conceptualmente algunos aspectos en relación con la medida cuantitativa del *crecimiento de la ciencia*, en particular con la intención de delimitar la aplicabilidad de los recuentos de publicaciones y de citas para este fin. Una vez aceptada la anterior tricotomía, se tratan las restricciones impuestas por una serie de objeciones en el uso de ambos tipos de recuentos.

Así, usa el término *actividad* para referirse al trabajo (como conjunto de tareas) que se realiza durante el desarrollo de la investigación. *Productividad* es, dice, la distancia recorrida hacia el destino, es decir, el avance efectivo que han permitido los resultados alcanzados mediante la actividad científica. Por último, llama *progreso* a la proporción que representa el avance conseguido respecto al camino total hasta el destino. Tanto la productividad como el progreso son difíciles de medir con meros recuentos de publicaciones o citas, que suelen reflejar más bien la actividad.

Moravcsik afirma que, si concebimos la ciencia como un sistema *entrada-salida*, la *actividad* estará relacionada con la entrada (*input*), mientras que la productividad y el progreso tienen que ver con la salida (*output*). Para él, aunque la mayor parte de las cuestiones relativas a la elección de prioridades científicas pertenecen a la entrada de recursos materiales y humanos, las consecuencias más importantes de la investigación científica surgen en la salida. Por lo tanto, existe una buena razón para basar las decisiones que afectan a la entrada de los recursos en medidas acerca de la salida. Por ello resulta deseable disponer de indicadores satisfactorios de producción y de progreso, que superen las limitaciones y sesgos de los meros recuentos de publicaciones y de citas. Moravcsik repasa algunas de las dificultades que ponen en cuestión la validez de esos recuentos como medidas del progreso científico: diferencias entre las áreas debido a diversos factores (patrones de comunicación, número de investigadores, colegios invisibles, cuidado y esfuerzo requerido en las publicaciones...); fenómenos como la obliteración, que hace que lo muy conocido deje de citarse; trabajos con errores; diferencias en las prácticas editoriales; efectos geográficos, etc. Además, sugiere que

[...] la naturaleza y el papel de las publicaciones científicas ha cambiado tan drásticamente durante las últimas décadas que un análisis temporal del desarrollo de las ciencias en términos de simples recuentos de publicaciones o citas puede estar seriamente desencaminada. (Moravcsik, 1973: 273)

Su objetivo principal es la medida del progreso científico, aunque no logra decir exactamente en qué consiste éste para él. Sin embargo, sí parece que lo considera una cuestión interna de la ciencia, al enfatizar que quizá las medidas de citas y de publicaciones puedan pertenecer solamente a la mecánica de la actividad de comunicación de los científicos y no puedan decirnos nada útil sobre el desarrollo de las ciencias (1973: 275).

Si bien puede intuirse una cierta afinidad entre la anterior distinción de tres conceptos y la que separa la cantidad de la calidad, un examen más detenido hace que se aparezca como un tanto complicado el encajar ambas terminologías; de hecho, Moravcsik sólo se refiere de pasada a la segunda de ellas (p.269). Una tentativa de casarlas puede partir de la aceptación del supuesto de que la actividad científica genera resultados, algunos de los cuales son denominados *productividad científica* porque contribuyen al avance del conocimiento; cabría entonces pensar que el conjunto de los resultados (los que producen avance y los que no) sería la cantidad, mientras que la calidad podría quizá entenderse o bien como la parte útil, es decir, la productividad, o bien como la fracción que representa ésta respecto al total de los resultados. En el primer caso, si no nos conformamos con esa definición meramente formal, se requeriría una medida absoluta para cuantificar el avance; en el segundo caso, aquellos resultados que tuvieran "menos desperdicio" serían de más calidad, con lo que esta noción se acercaría a lo que podemos llamar la *eficiencia*<sup>45</sup> de la actividad científica.

Otro modo de establecer una correspondencia entre ambas clasificaciones, entre la tricotomía de Moravcsik y la dicotomía cantidad-calidad, podría consistir en tomar el *progreso científico* realizado como

---

<sup>45</sup> En este contexto resulta especialmente clarificador aplicar la definición que Quintanilla (1989) da para la eficiencia de los sistemas técnicos:  $E = |O \cap R| / |O \cup R|$ , (ratio entre la intersección de objetivos y resultados y la unión de objetivos y resultados) especialmente porque resalta el compromiso de Moravcsik con que la ciencia tenga un fin concreto y definible.

medida de la *calidad*, con lo que la *productividad* sería entonces la *cantidad*. Los resultados que no contribuyen al avance se ignoran en este caso por completo. Esta opción se presenta bastante atractiva, puesto que liga las diferencias de calidad al progreso que cada trabajo ha producido; es bien cierto que la conexión entre calidad de los trabajos y progreso científico se adivina íntima, aunque por el momento sea de modo un tanto confuso.

Otra posible interpretación consistiría en asumir que la calidad se refiera no al resultado, sino a la actividad. La calidad científica sería, entonces, lo bien que se ha realizado el trabajo de investigación. En este caso sería necesario aclarar las relaciones causales que existen entre la calidad y el éxito, es decir, si la calidad es una condición necesaria, suficiente o solamente aconsejable para producir progreso en la ciencia.

La primera impresión que provoca esta variedad interpretativa es que la distinción propuesta por Moravcsik oscurecería, más de lo que ya estaba, la relación entre la cantidad y la calidad. Comienzan a aparecer otras magnitudes, como la de *esfuerzo no productivo*, o la de *progreso*, que complican la aparente simplicidad de considerar que pueden encontrarse diferencias entre los productos sobre la base de su calidad. Esto bien puede valer para despertar la duda acerca de la pertinencia de la distinción cantidad-calidad en la práctica científica cotidiana. Quizá esa práctica esté guiada por sus propios objetivos y valores, y no necesite recurrir a la dicotomía clásica para describir adecuadamente sus logros o su desarrollo. Esta hipótesis cobra cierto vigor si a la anterior sospecha unimos la condición de Moravcsik como físico practicante y competente, reconocido en su ámbito. Por otra parte, también se aviva en la memoria el recuerdo de que la necesidad de separar conceptualmente la cantidad de la calidad proviene de las dificultades para interpretar adecuadamente las incipientes medidas que se tenían a mano, léase los recuentos de publicaciones y de citas. El artículo de Moravcsik se dedica precisamente a mostrar las limitaciones de ese tipo de recuentos para reflejar el crecimiento de la ciencia, e ignora deliberadamente la distinción que separa la cantidad de la calidad porque propone otra alternativa a ella. La disponibilidad de objetos o propiedades cuantificables no es una razón suficiente para que las medidas que puedan construirse sobre ellos tengan algún significado, y menos aún para que ese significado coincida justamente con el que deseamos medir.

La reserva que aconseja el párrafo anterior es más que razonable: nos previene de los peligros que conlleva el dejarse guiar por un optimismo ingenuo en cuanto a las posibilidades de medida, y trae de nuevo a escena el hecho, descuidado a veces, de que la ciencia no es ella misma un objeto científico en el mismo sentido en que lo pueden ser los objetos naturales, sino el producto de un sistema social. La colección de conceptos y de taxonomías que empleemos en la descripción de la ciencia puede depender fuertemente de nuestros intereses como observadores y de las limitaciones de los indicadores que podamos construir.

De todos modos, la distinción propuesta por Moravcsik no es menos oscura que la que separa la cantidad de la calidad. Sin un esfuerzo interpretativo un tanto libre, no es nada fácil entender lo que significa exactamente. Su metáfora topológica (al definir el progreso y la actividad científica en términos de distancias y avances en una superficie) no es accidental: refleja una concepción del progreso ciertamente problemática y seguramente demasiado simple. Una consecuencia de aceptar su explicación, es que conduce a admitir que alguien habrá progresado mucho sólo porque se haya marcado metas poco ambiciosas. Además, su noción del progreso está regida por el punto de destino. Pero éste no siempre se puede determinar con precisión: ¿qué pasa si surgen resultados inesperados y cambia el rumbo? El conocimiento exacto de la distancia que separa el estado actual del punto de llegada es imposible a no ser que se sepa con certeza dónde está el destino, y en ese caso carecería de interés porque ese punto ya formaría parte del conocimiento actual, es decir, no sería ningún avance llegar de nuevo a él. Esto implicaría que solamente puede medirse el progreso que han supuesto los resultados pasados que han sido completamente asimilados, de los que no cabe esperar ya más consecuencias.

Dentro de esta misma familia de concepciones sobre qué sea el progreso de la ciencia, al lado de esta metáfora del camino hacia un destino está la metáfora del explorador, es decir, de la ampliación del espacio conocido, que implica casi siempre una reducción del espacio ignoto. Pero este tipo de imágenes del progreso son muy poco compatibles con la posibilidad de profundos cambios en los conceptos y en los marcos teóricos, posibilidad insoslayable a juzgar por la fortaleza que las propuestas teóricas comprometidas con ella han gozado durante décadas a lo largo de este nuestro siglo.

El concepto de progreso científico es el centro de una discusión compleja en la filosofía de la ciencia desde hace varias décadas<sup>46</sup>. Por poner un ejemplo, uno de los criterios más socorridos, el de la solución de problemas pendientes, no siempre es satisfactorio: cuantas más preguntas se solucionen más preguntas surgirán. Es posible que el desarrollo del conocimiento sólo pueda medirse por comparación con el pasado, no por lo cerca que estamos de "no se sabe dónde por definición". Si somos capaces de preferir el estado actual respecto a cualquiera de los precedentes, entonces se ha avanzado; pero la reconstrucción histórica puede mostrar caminos serpenteantes y pasos atrás provisionales, necesarios para llegar a dónde se está ahora. Quizá una parte del progreso es precisamente desarrollar esos criterios que nos permiten sentirlo. El acicate de estar "ya cerca" es muy bueno para estimular los científicos, pero se lleva sintiendo desde hace ya varios siglos; parece que la última verdad fundamental, que la ley básica está al alcance de la mano. Hasta hace bien poco, por ejemplo, se oía a algunos físicos que la unificación de todas las fuerzas elementales estaba próxima.

La propuesta de Moravcsik, por lo tanto, encuentra su principal debilidad en su dependencia total de una noción de progreso pobremente definida. Su definición hace virtualmente imposible disponer de medidas para esas categorías, por lo que no es de extrañar que ni los recuentos de citas ni los de publicaciones puedan cumplir esa misión.

A cambio, su separación del esfuerzo y sus resultados del progreso efectivo puede interpretarse sin duda como que la actividad científica puede generar resultados que no contribuyan al avance. Es posible que su posición parezca estar más cerca de considerarlos como ruido inútil, como desperdicio de trabajo y de recursos. Pero su distinción, como se verá más adelante, puede tomarse también como el inicio del camino hacia la valoración científica de los resultados incluso aunque no produzcan progreso directamente.

Bajo una notable influencia de esta propuesta de Moravcsik, surgió uno de los trabajos más importantes en la serie de intentos por sistematizar la interpretación de los indicadores bibliométricos basados en recuentos de publicaciones y de citas, el que debemos a B.R. Martin y J. Irvine (Martin, Irvine, 1980). Proponen la aproximación de los *indicadores parciales convergentes*, como solución a las limitaciones que

---

<sup>46</sup> Kitcher (1993) presenta una excelente panorámica de este debate.

presenta cada indicador individual como reflejo del *progreso* en la ciencia. En la raíz del planteamiento está tanto la constatación de que existen diferencias en la contribución a la ciencia de cada publicación como la posibilidad de que, además, esas diferencias no sean reconocidas con justicia, ya sea por exceso o por defecto. La aportación al *progreso* de la ciencia es el criterio básico para juzgar el valor de las publicaciones científicas.

En ese trabajo, Martin e Irvine llevan a cabo una revisión de los principales intentos por justificar las medidas de publicaciones y de citas. Es significativo su esfuerzo por plantear el problema, con una interesante clarificación conceptual, y por integrar sintéticamente algunas de las aportaciones precedentes. Además, llevan a cabo un estudio empírico que les permite poner a prueba su propuesta.

Toman como básica la distinción de Moravcsik que se acaba de comentar, aunque su exposición puede considerarse más una reformulación. La *actividad* está relacionada con el consumo de recursos; la *producción* de la investigación es la medida en que este consumo crea un cuerpo de *resultados científicos*; por último, el *progreso* es el conjunto de contribuciones sustantivas al conocimiento científico, tal y como es juzgado por otros científicos. Aquí aparece más clara la noción de producción como la suma de resultados en la que están incluidos también los que no constituyen una "contribución sustantiva".

Pero sobre esta consideración básica ofrecen una interesante aportación. Consideran necesaria la distinción de tres conceptos al hablar de la investigación descrita en un documento científico: *calidad*, *importancia* e *impacto* (Martin, Irvine, 1980: 70). El primero, *calidad*, es una propiedad del escrito y de la investigación que éste describe. Tiene que ver con lo bien que se hizo la investigación, es decir, si está libre de errores obvios, si tiene originalidad, claridad, etc. Sin embargo, es más relativa que absoluta, está cognitiva y socialmente determinada: los juicios dependen de los intereses que se tengan o de la finalidad que se espera de la investigación. La *importancia* se refiere a la influencia potencial de la publicación, es decir, la influencia que ejercería en el avance del conocimiento científico *si la comunicación fuera perfecta* entre los investigadores, podemos decir, si todos ellos tuvieran en la práctica acceso a los materiales que les interesan y éstos estuvieran presentados en un modo perfectamente comprensible, fueran claros y explícitos. Por último, el *impacto* es la influencia que ha tenido de hecho en el resto de las investigaciones.

Depende de su *importancia*, pero también de una serie de factores que desvían el perfecto funcionamiento del sistema de comunicación científica, tales como sesgos sociales y psicológicos de los científicos, aislamientos geográficos y culturales, grado de difusión de la fuente, etc.

Llegan a la conclusión de que no es posible obtener ninguna medida absoluta o directa de la calidad, la importancia o el impacto. Los indicadores bibliométricos, como los recuentos de citas o de publicaciones, sólo pueden ser *indicadores parciales* de las anteriores dimensiones. La razón principal es que existen muchos factores, aleatorios y sistemáticos, relacionados con el sistema de comunicación de la ciencia, que influyen en esas medidas bibliométricas además de las características propias de los trabajos escritos. Este reconocimiento permite, por ejemplo, eliminar muchas controversias en torno a la validez de los recuentos de citas como indicadores de la *calidad* o la *importancia*: las citas pueden ser tomadas sólo como un indicador *parcial* del *impacto*. Las debilidades atribuidas a los análisis de citas (retraso en citar trabajos importantes, errores que reciben muchas citas, obliteración, trabajos excelentes en campos pequeños, etc.) reciben una explicación adecuada si las citas se entienden como reflejo del impacto y se relacionan más bien con la organización de la ciencia y de su sistema de comunicación.

Consideran, asimismo, que los recuentos de publicaciones no son *medidas* del *progreso*, pero sí son *indicadores parciales* de éste. Esto es así porque la publicación puede estar determinada parcialmente por el progreso alcanzado por un investigador o un grupo, pero también por otros factores como localización institucional, presiones sociales y políticas, etc. De esto último parece deducirse que, en su opinión, cabe esperar que todas las publicaciones contribuyan al progreso en alguna medida, aunque no todas lo mismo: la producción siempre da lugar a algún progreso, por pequeño que éste sea. Afirman que puede tenerse una cierta confianza en los recuentos relativos a grupos seleccionados por compartir una serie de características comunes, de modo que los efectos de los "otros factores" en la publicación quedarían minimizados.

La propuesta de Martin e Irvine integra varios tipos de indicadores parciales, incluyendo otros no bibliométricos como la evaluación por pares, en un sistema de indicadores que sustentan mutuamente su validez, es decir, se aceptan mientras no se contradigan entre sí y apunten en la misma dirección. Se trata de una propuesta bien trabada y

razonable, guiada especialmente por la prudencia interpretativa de los indicadores disponibles.

Como comentario general, puede decirse que su planteamiento parte de conceptos predefinidos, analíticamente separados, y luego busca medidas o aproximaciones de ellos. El problema primario es el de la evaluación, la cual persigue una estimación del progreso logrado por la investigación; la solución que ofrecen es que no existen medidas directas de lo que se quiere medir, y por lo tanto hay que emplear indicadores determinados parcialmente por las magnitudes que interesan.

Hay que señalar un primer punto en relación con esta propuesta, y es que aparece como algo implícito la total separación entre los conceptos fundamentales a cuya medida se aspira (producción, progreso, calidad, importancia, impacto) y las cuantificaciones disponibles, básicamente los recuentos de citas y de publicaciones. Invita a asumir que esos conceptos son independientes, previos a la disponibilidad de los instrumentos que pueden servir como medidas directas, indirectas o como indicadores parciales de ellos. Por ello, no extraña nada que esos indicadores tengan que ser necesariamente parciales: no existe ninguna conexión necesaria entre esas medidas y los conceptos citados, por lo que su enlace es, en cierto sentido, accidental. Parece ignorarse la posibilidad de que esa colección de conceptos haya sido configurada sobre la base de las medidas disponibles.

La segunda cuestión se refiere a la precisión del significado que atribuyen a esos conceptos. Aunque es bien cierto que no se extienden en demasiadas explicaciones sobre qué significa cada uno de ellos, la sensación que produce su exposición es que se refieren a dimensiones diáfamanamente distinguidas. Pero, incluso dejando a un lado los problemas de la noción de progreso, a los que ya se ha aludido, cuando se indaga sobre las relaciones que mantienen entre sí cada una de esas dimensiones surge la duda: quizá no estén delimitadas con suficiente detalle. No se indica, por ejemplo, si la producción científica supone necesariamente o no alguna contribución al progreso. Ni tampoco el modo en que la importancia depende de la calidad, si es que depende de ella, ni si la calidad garantiza algún progreso, o alguna importancia; quizá piensen que la calidad sea una condición para considerar que algo es un resultado científico, con lo que todos los resultados deberían exhibir algún grado de calidad. Aunque el significado del impacto se parece mucho a "lo que miden las citas", en realidad se le dota de un significado propio, el de la

influencia que se ejerce sobre la investigación posterior, sin especificar en qué sentido; al menos parece que el impacto presupone la importancia, aunque pueda ser también que la constatación del impacto sea el único modo de atribuir alguna importancia...

La noción que más interesa en nuestra discusión, la de calidad, queda definida como el misterioso resultado del gusto de los científicos, sin que se observe ninguna otra razón por la que un observador o evaluador externo debiera preferir a los trabajos de mayor calidad. Pero además se deja sitio para la ambigüedad de considerar que la calidad es una propiedad del trabajo en sí. La precisa distinción conceptual no ayuda finalmente a saber a qué nos referimos cuando hablamos de calidad, ni si tiene algún valor. Pero es necesario reconocer que supone un interesante avance en cuanto a la interpretación de los indicadores y a una evaluación más sofisticada de la actividad científica y sus resultados. La colección de conceptos que Martin e Irvine proponen mejora mucho la situación respecto a hablar solamente (y de modo mucho más difuso) de cantidad y de calidad, y pone de manifiesto que la confusión en torno a la noción de calidad podía estar producida por la inclusión de aspectos que eran analíticamente ajenos a ella. Con esta precisión de los conceptos y la justificación de los límites interpretativos de las medidas, el debate parecía acercarse a su fin, especialmente en lo concerniente al uso de los indicadores.

### 5.2.3. *La "salud" de la ciencia*

El libro *Quality in Science* (La Follete (ed.), 1982) expresa la preocupación que supuso este tema en los Estados Unidos. Plantea la cuestión acerca de la construcción de indicadores de calidad sobre la ciencia en unos términos un tanto diferentes a los de las contribuciones que se han venido comentando hasta aquí. En el prólogo, el historiador de la ciencia G. Holton lo presenta como la reunión de una serie de intentos por buscar los significados operacionales del concepto de calidad científica, tanto en lo que respecta al estado de la ciencia como a su impacto sobre la vida humana. Pero, por poco preciso que sea la expresión "estado de la ciencia", el conjunto de trabajos que componen el libro no parecen dedicar igual atención a los dos polos señalados por Holton. Puede decirse que el volumen, en su conjunto, está muy atado a la necesidad de

justificar socialmente la utilidad de la ciencia y a la urgencia por planificarla, por elegir opciones científicas. En realidad, el problema que preocupa es el del control racional de la ciencia por parte de la sociedad, el de cómo justificar la dedicación de recursos a unos científicos o a otros.

Una imagen que recorre el libro es la de la *salud de la ciencia*: en un país, tal disciplina científica o tal otra son de calidad si su estado es saludable. La idea más reiterada, con distinta terminología, es la de la distinción entre dos dimensiones de la calidad, y por lo tanto de juicios acerca de la calidad, los internos y los externos. La calidad interna se refiere a la salud de la ciencia; la externa, al beneficio que tienen las consecuencias de la investigación científica en la sociedad. Ambas tienen que ver con valores, pero especialmente la externa. Lo que parece subyacer en este planteamiento es que lo preferible es lo que tiene calidad, como por definición. El apoyo a unas disciplinas o a otras puede así justificarse bien mediante las razones internas a la ciencia o bien mediante razones externas, es decir, económicas, políticas, sociales en suma.

Una posición común a varios textos del libro parece ser la de pensar que la distribución de fondos debe tener en cuenta los criterios de calidad interna, pero subordinados a los de la calidad externa; los mejores deben ser apoyados, pero siempre que sus resultados sean útiles para la sociedad que les paga; por un lado está la cuestión de establecer la competencia justa entre los científicos, por otro, el derecho de la sociedad a elegir su destino.

Harvey Brooks, por ejemplo, resalta la tensión entre los fines que autónomamente se propone la ciencia y los que la sociedad puede esperar de la ciencia a la cual mantiene. Habla de dos tipos de indicadores: *intrínsecos* y *extrínsecos*. Fija su atención en los segundos: los indicadores extrínsecos deberían poder medir no solamente el avance hacia ciertos objetivos sociales, sino también juzgar el valor relativo de los objetivos mismos (p. 3). De los intrínsecos no se ocupa especialmente, aunque lo que puede deducirse es que han de reflejar los avances en los objetivos internos de la ciencia, fijados por la persecución de la verdad. Acerca de la calidad interna, lo más que hace es equipararla a la "significación". Su preocupación fundamental en ese trabajo es la de estudiar la interacción entre la ciencia y el resto de la sociedad, en concreto, mostrar las dificultades de desarrollar medidas realistas acerca del estado de la ciencia que puedan traducirse en índices de bienestar de la sociedad (p. 2)

Bruce Mazlish también distingue dos usos del término calidad: interno y externo. Tampoco dedica una especial atención al primero, pero parece deducirse que tiene que ver con la valoración que los científicos realizan del trabajo de un colega. En su uso externo, calidad científica se refiere a lo beneficioso de su impacto en la sociedad y en la técnica. Cree que son necesarias medidas que se refieran a los dos usos, aunque subordina la interna a la externa. Advierte sobre los peligros de objetivar medidas, tal y como ha sucedido con los indicadores económicos; también indica los inconvenientes del uso de analogías como la de la *salud*, que hacen olvidar diferencias significativas y ocultan el conjunto de valores que subyacen a estos juicios. La cuestión de evaluar la ciencia no se puede reducir a su aspecto técnico, no puede ser dejado en manos de "expertos", porque importan mucho los valores sociales. Llama la atención sobre la necesidad de lograr un balance entre lo interno y lo externo.

Lewis Branscomb aporta la experiencia en la gestión de una empresa con fuerte componente de investigación, (IBM) y en la National Science Foundation, principal organismo de los Estados Unidos en cuanto a la política de investigación federal. Expone una serie de criterios mediante los cuales establecer un refuerzo de la calidad; como los anteriores autores, distingue entre calidad interna, que sólo puede ser medida por los colegas o pares, y calidad externa, que tiene varios componentes: rentabilidad, utilidad... Señala que, a pesar de la similitud con la separación entre investigación *básica* e investigación *aplicada*, la anterior distinción no es equivalente. Desde el punto de vista de la gestión empresarial de la investigación y el desarrollo, enumera una serie de dimensiones que han de tenerse en cuenta para valorar la actividad realizada: conocimiento logrado, uso efectivo de los resultados de la investigación, desarrollo del personal, presiones derivadas de la necesidad de eficiencia financiera.

El resto de los trabajos contenidos en el libro se ocupan de diferentes aspectos de las relaciones entre la ciencia y la sociedad, especialmente del modo que la segunda puede aprovechar y dirigir los logros de la primera. Como se ha podido entrever, es muy escasa la aportación en cuanto a la precisión de qué sea la calidad científica, pero es muy reseñable el esfuerzo por relacionarla con los intereses sociales ajenos a la dinámica propia de la ciencia. Podemos arriesgarnos con la siguiente interpretación global del libro: la ciencia es de buena calidad si colabora adecuadamente en la persecución de los objetivos sociales en los que

puede prestar ayuda. La calidad interna también es importante, aunque es inabordable desde juicios externos. La utilidad social de una disciplina en un país, por ejemplo, no depende solamente de su calidad interna, sino también de otros factores casi accidentales. Pero la calidad interna concreta las potencialidades de una disciplina en cuanto a su beneficio para la sociedad.

#### 5.2.4. *Dos enfoques: indefinición versus concreción.*

Dentro de la línea de investigación esforzada en encontrar un espacio de validez para los indicadores bibliométricos basados en recuentos de citas, S.M. Lawani fue uno de los autores que propusieron enseguida la necesidad de distinguir entre *cantidad* y *calidad* en las medidas sobre la producción científica (Lawani, 1977). Más tarde, en un trabajo en colaboración con A.E. Bayer (Lawani, Bayer, 1982), encuentra pruebas a favor de la capacidad discriminatoria, aunque de grano grueso, de los análisis de citas en cuanto a la calidad estimada independientemente por expertos. En él asumen que el funcionamiento de la comunidad científica (ya sea por su carácter acumulativo, por su sistema de reconocimiento o por los esfuerzos de persuasión de los investigadores) ofrece una base firme para considerar que las citas recibidas por un trabajo son un reflejo de la importancia de éste, incluso admitiendo los fallos señalados por algunas objeciones clásicas. Sin embargo, en lo que se refiere a la concepción de la *calidad* científica que los autores sostienen en este trabajo, es llamativo el escaso interés que manifiestan por precisar el significado de este término: como toda ayuda a la comprensión de ese significado, presentan la siguiente lista de términos como sinónimos entre sí: *productividad, significación, calidad, utilidad, influencia, efectividad e impacto*; también emplean indistintamente los términos *calidad e importancia* (Lawani, Bayer, 1982: 60).

Esta posición, que parece implicar un tratamiento superficial de la noción de calidad científica, puede ser tomada en realidad como un caso de sensata prudencia, quizá debida a la convicción de que no es posible especificar la lista de atributos que determinan la calidad en la ciencia; de ésta sólo podría decirse que el mejor tribunal disponible es la propia comunidad científica y que no se puede saber qué criterios concretos aplica en cada caso para juzgar cada trabajo. La calidad parece ser algo

así como las consecuencias positivas en la esfera científica, según la apreciación de los propios investigadores. Aunque sus conclusiones establecen la validez de los análisis de citas en relación a agregados y destacan los problemas en aplicarlos a trabajos individuales, no se ocupan de tratar una importante cuestión: si cambia el concepto de calidad al referirse a agregados de trabajos, a una labor en su conjunto, en lugar de a trabajos concretos.

En un trabajo posterior (Lawani, 1984), Lawani asume la distinción de dos sentidos o usos del concepto de calidad en la ciencia que ya había estado presente en el libro *Quality in Science* anteriormente comentado. Por un lado estaría lo que puede denominarse *calidad interna*, que consiste en la estimación del valor por parte de los profesionales (es decir, de los científicos) y por otro lado, la *externa*, que depende de las consecuencias de la investigación científica que se consideran beneficiosas para la sociedad, en el sentido de que permiten alcanzar fines económicos, políticos, sociales o sirven a valores o intereses humanos. Considera que los indicadores bibliométricos son, desde luego, más adecuados para captar la calidad interna que la externa. En ese trabajo sostiene la opinión de que los juicios sobre la calidad interna de un escrito científico dependen de un conjunto de factores, tales como su significación científica, la validez de la metodología, análisis y conclusiones, así como la claridad y simplicidad de la presentación. Mantiene así una cierta vaguedad como modo de afrontar lo relacionado con la calidad científica, aunque aquí deja entrever que pueden existir una serie de virtudes en el trabajo que pueden pesar en el grado de calidad que se le atribuye, sin tener que esperar al influjo que el trabajo pueda efectivamente ejercer con posterioridad.

Una curiosa propuesta en el debate acerca de la determinación de la calidad de los trabajos científicos es la de I.M. Beck (Beck, 1984). Este autor parte de la necesidad de disponer de un sistema que permita evaluar la calidad de los trabajos con independencia de la subjetividad del juez que se encargue de ello. Rechaza las medidas basadas en recuentos de publicaciones o de citas como instrumentos válidos suficientes para la evaluación de la calidad. Beck se adscribe a una teorización acerca del pensamiento científico que lo considera compuesto por seis *estaciones*, completamente relacionadas entre sí (es decir, todas con todas), cada una de las cuales tiene que ver con uno de los hechos siguientes: a)

*establecimiento de axiomas*, b) *generación de teoremas*, c) *contrastación y ajuste de modelos teóricos a la realidad*, d) *obtención de hechos empíricos*, e) *práctica heurística*, y f) *práctica rutinaria*. El orden anterior expresa la jerarquía de importancia para la evaluación de la calidad. Concibe la calidad científica como la posesión de alguna de estas dimensiones, entendiéndose que cuantas más dimensiones incluya un trabajo, de más calidad será. Puesto que existe una jerarquía entre las dimensiones, propone una escala para medir la calidad, de modo que puntúan más alto las dimensiones que aparecen al principio de la lista. Su procedimiento postula que solamente pueden poseerse dimensiones contiguas entre sí. Por lo tanto, la mayor calidad se da cuando un trabajo afecta al establecimiento de axiomas y, pasando por todas las dimensiones, implica efectos en la práctica rutinaria; la segunda mayor puntuación se obtiene si afecta a los axiomas pero sólo llega hasta la práctica heurística, y así sucesivamente, para llegar desde el 21 hasta el 1. La escala tiene una segunda mitad negativa "para referirse a la destrucción o la refutación". Según el autor, la aplicación de esta escala requiere un sólo experto, porque el desacuerdo es imposible.

En principio, su procedimiento no especifica grados dentro de cada dimensión, aunque sugiere que se construyan escalas para cada una de las dimensiones. Además, contiene varias arbitrariedades, como la de la jerarquía definida (derivada de una rudimentaria filosofía de la ciencia) o la de no admitir la existencia de contribuciones que afecten a *estaciones* no sucesivas. Ciertamente, no resulta nada fácil imaginar la aplicación práctica de este método para evaluar trabajos, que seguramente generaría fuertes contradicciones con las percepciones habituales acerca de la calidad de los trabajos científicos.

Pero, independientemente de la fortuna de la propuesta, el esfuerzo por definir el contenido de la *calidad científica* es un reflejo más de la falta de una definición generalmente aceptada acerca de esta noción. En este caso, el autor cree que una medida normalizada de la calidad en la ciencia está demandada por la necesidad de que personas externas a la ciencia cuenten con una evaluación rigurosa de los productos científicos y que se atenga a su valor interno en la ciencia.

Beck aplica lo que podría llamarse un *criterio epistemológico*, que al menos en cuanto a intención, no está nada lejos de las tradiciones de filosofía de la ciencia: la calidad epistémica sería el conjunto de características que hacen que un conocimiento pueda ser considerado mejor que

otro, o científico frente a otro que no lo es. Un rasgo común a este tipo de planteamiento es el de ignorar la posibilidad de que existan resultados científicos de calidad, obtenidos mediante la aplicación rigurosa de métodos corrientes en la ciencia, y que sin embargo no aporten finalmente ningún conocimiento como los requeridos por estas definiciones *a priori*.

#### 5.2.5. Problemas de evaluación de la ciencia

La preocupación por el contenido del concepto de calidad científica surgió, también, en el marco de los trabajos llevados a cabo por un grupo de la Universidad de Leiden dedicado a explorar las posibilidades de los análisis bibliométricos como instrumentos de apoyo a la evaluación de la investigación en instituciones universitarias (Moed, Burger, Frankfurt y van Raan, 1984). Cuando en esa universidad la *calidad científica* se convirtió en un criterio básico para la financiación de la investigación, se sintió la necesidad de estudiar el alcance de los indicadores bibliométricos en relación con su capacidad para discriminar satisfactoriamente la calidad científica de los grupos que allí trabajaban. Una de las primeras cuestiones fue la de establecer los límites de la interpretación de los indicadores bibliométricos respecto a los juicios que de ellos pueden extraerse sobre la calidad científica. En el fondo aparece de nuevo la necesidad de valorar de distinto modo los trabajos que publican los investigadores, por la convicción de que existen diferencias relevantes que quedan ocultas por los recuentos de publicaciones.

Así, asumen en primer lugar la separación de dos aspectos del rendimiento de la investigación: *producción e impacto*. Uno se refiere al corpus de resultados científicos generados por la investigación, y el otro, a la influencia que éste corpus ha tenido en las subsiguientes investigaciones relacionadas<sup>47</sup>. Es obvia la correspondencia con los dos tipos de recuentos disponibles, de publicaciones y de citas. En segundo lugar, se ocupan de depurar el significado de los dos términos de una distinción que ya había sido propuesta con anterioridad (Garfield, 1963): *calidad e impacto*, cuyos fuertes nexos ocasionan confusiones nada deseables.

---

<sup>47</sup> Para la producción, repiten casi literalmente la definición de Martin e Irvine (1980), aunque lo denominan *research output* en lugar de *research production*. Mezclan dos de las clasificaciones utilizadas por Martin e Irvine en el trabajo citado.

Existen dos clases de impacto, cada una con un significado propio: el *impacto a corto plazo*, medido aproximadamente por las citas recibidas en el periodo inmediato a la publicación (durante los dos años siguientes, por ejemplo) tiene que ver con la visibilidad del trabajo de los investigadores en el frente de investigación internacional; el *impacto a largo plazo*, por su parte, reflejado por las citas recibidas durante periodos dilatados tras la publicación, indica el grado en que los resultados científicos "han triunfado" sobre las otras alternativas y han sido asumidos por el conocimiento científico básico, el aceptado generalmente en la disciplina. El impacto a corto plazo, pues, habla de la atención despertada en la investigación del momento, del papel jugado en las discusiones científicas; el impacto a largo plazo se refiere a la permanencia alcanzada por las propuestas, a las contribuciones efectivas al acervo científico.

Para las políticas de investigación, es mucho más relevante el impacto a corto plazo. Por un lado, cuando se disponen de medidas del impacto logrado a largo plazo, suele ser demasiado tarde para que éstas tengan alguna utilidad. Pero además opinan, por otra parte, que no se puede exigir a los investigadores que sus resultados consigan permanecer e integrarse en el conjunto de conocimientos considerados básicos. Lo significativo de esta apreciación es que de ella podemos deducir que los motivos del éxito final en la ciencia caen en una buena medida fuera del ámbito de la responsabilidad de los individuos, es decir, que no les basta con ejecutar *bien* sus tareas para producir contribuciones permanentes a la ciencia. Estirando un poco más y con otras palabras, puede afirmarse entonces que la última fase del proceso de producción de conocimiento científico es esencialmente social y la desempeña la comunidad científica, por lo que en cierto sentido cae fuera de la competencia de los individuos o de los grupos aislados. Éstos sí deben, en cambio, esforzarse por mantener una presencia significativa y respetada en la solución de los problemas abiertos, lo cual se consigue, en definitiva, publicando *buenos* trabajos.

¿Por qué podemos decir de un trabajo que es *bueno*? Al ocuparse de la calidad, los autores declaran como prolegómeno que ésta es virtualmente imposible de operacionalizar, es decir, que no se pueden encontrar indicadores que la midan directamente. Esto es así porque la calidad puede referirse a una variedad de valores, que además no están determinados con precisión. Lo que denominan *calidad básica* tiene varios

niveles (Moed et al., 1984: 50): *cognitivo*, relativo a la importancia del contenido específico de ideas científicas y juzgable sólo mediante consideraciones puramente científicas; *metodológico*, que se refiere a la adecuación de los métodos, según las reglas al uso en la disciplina correspondiente; y *estético*, que se basa en el atractivo de las formulaciones y modelos matemáticos, habitualmente dependiente de la relación entre la simplicidad y la potencia explicativa. El juicio de esta calidad básica se funda en criterios intrínsecos a la investigación científica y sólo puede ser emitido por colegas expertos, por pares. Hay que notar que no especifican ningún motivo que justifique la preferencia de los expertos por estos valores ni tampoco si éstos son un mero reflejo de normas sociales extendidas entre los miembros de la comunidad científica.

Pero aquí llega una interesante innovación: sostienen que también forma parte de la calidad de la investigación la capacidad de los investigadores para mantener *relaciones públicas* productivas, en el sentido de saber hacer llegar eficazmente sus hallazgos a los colegas. Las habilidades comunicativas y sociales pueden incluso ser un requisito para que los resultados sean atendidos. Conciben así un nuevo tipo de calidad, que incluye tanto rasgos particulares del trabajo como destreza social para presentarlo. La calidad *básica*, la definida por valores intrínsecos a la ciencia, es necesaria para lograr impacto, pero no suficiente. Nuestra curiosidad por conocer la lista ejemplar de actividades "públicas" no es satisfecha por los autores; podemos imaginar que se refieren a cosas como la exposición en foros relevantes, la publicación en el momento y en los medios oportunos, el procurarse un buen puesto en la red de comunicación informal, tener contactos con investigadores relevantes... pero quizá también debamos incluir la capacidad para detectar problemas interesantes o la de realzar su trascendencia articulándolos de manera explícita con otros problemas abiertos. Éstas últimas hacen un poco borrosa la delimitación entre los dos componentes de la calidad según la entienden estos autores, porque parecen afectar a ambos lados; ofrecen así un argumento a favor de entender la calidad científica en el sentido amplio que defienden.

De este modo, los autores sortean de una manera muy elegante la limitación expuesta por Martin e Irvine: en éstos, la *influencia* estaba determinada en buena parte por la *importancia* y por la *calidad*, pero también podía depender de sesgos derivados del funcionamiento

imperfecto del sistema de comunicación. Con la noción de calidad ampliada, se traspasa la responsabilidad de ese funcionamiento a los propios autores, como parte de su capacidad científica, de modo que la calidad es la mayor causa del impacto a corto plazo.

Sostienen, en fin, que los indicadores de impacto que proponen (basados en recuentos de citas y que exhiben un grado apreciable de refinamiento) son indicadores de la calidad en este sentido extendido. Sus indicadores de impacto se refieren al *impacto a corto plazo*. De modo que asumen que las citas recibidas en el periodo inmediato a la publicación están fundamentalmente determinadas por la *calidad amplia*; no obstante, algunos factores bien conocidos (diferencias en los patrones de cada disciplina o en la cobertura de los datos bibliométricos, entre otros) pueden introducir sesgos, pero la distorsión producida por éstos puede aminorarse mediante diversas normalizaciones.

Green, y no les faltan razones, que esta noción de la calidad clarifica las discusiones en torno a la evaluación del rendimiento de la investigación. Una de sus virtudes es que, al menos en cierto grado, consigue integrar la autonomía de las evaluaciones internas de la ciencia con una razonable definición externa de objetivos por parte de los gestores de la ciencia, consistente con la práctica científica corriente, la cual está habituada a que son necesarios una gran cantidad de esfuerzos y de resultados para lograr contribuciones significativas a la ciencia, sin que eso pueda servir para negar todo valor a los resultados más cotidianos.

Este planteamiento suscita algunas reflexiones en torno a los compromisos básicos que adopta, las cuales pueden servirnos para desgranar algunos de los puntos clave que juegan un papel velado, pero determinante, en el proceso de configuración de las concepciones de la calidad científica. Uno de los puntos que llama la atención es la disyuntiva que se abre acerca de si la calidad se refiere a los *productos* científicos o a la *actividad* que ha dado lugar a esos productos. Esta duda podría involucrarnos en un debate lleno de sutilezas ontológicas acerca de las relaciones entre la acción y su efecto; pero es posible señalar aquí al menos que muchas de las posiciones sobre la calidad científica han optado implícitamente por uno de los dos polos. En este caso, sin embargo, se logra un curioso e interesante eclecticismo: la calidad afecta independientemente a ambos, tanto a los productos considerados aisladamente como a las actividades que, incluyendo aquellas posteriores a la generación de resultados, forman parte del proceso de investigación.

Otra tensión que puede adivinarse a lo largo de los esfuerzos por clarificar la naturaleza de la calidad es la que sitúa ante la alternativa de considerar si la calidad es en algún sentido objetiva, independiente del juicio de la comunidad, o si por el contrario, es algo dependiente de valores arbitrarios sostenidos por la comunidad científica; o bien ante la exigencia de aclarar en qué medida está compuesta por rasgos objetivos y en cuál otra por rasgos convencionales. De la posición que tomemos depende el que se admita que la comunidad pueda equivocarse o, antagónicamente, que ésta sea el único tribunal válido. Aunque la respuesta que nos ofrece el grupo de Leiden no parece, a este respecto, decantarse claramente por ninguno de los extremos, resulta más compatible con otorgar un mayor peso a la influencia de las convenciones de la comunidad científica que a los efectos derivados de rasgos o atributos objetivos del trabajo científico.

También es llamativo el salto que se da entre la calidad referida a los trabajos concretos y la que luego se predica de los agregados de trabajos. Al hablar de la calidad *básica* miran hacia los criterios que parecen guiar los juicios que reciben por parte de los especialistas cada uno de los trabajos individuales; los indicadores de impacto, sin embargo, se obtienen referidos a los resultados de una actividad dilatada que ha producido diversas publicaciones, y se consideran reflejo de la que hemos llamado *calidad amplia*. Cabe pensar que la maña comunicativa afecta de un modo homogéneo a cada uno de los trabajos producidos por un investigador o un grupo durante un periodo concreto, por lo que si existen diferencias de impacto entre los trabajos de ese investigador o ese grupo, éstas se deberán probablemente a diferencias en la calidad básica de cada trabajo. Parece que el modelo para sentar qué es lo propio de la calidad siempre ha sido el descubrimiento o el experimento ejemplar, pero luego, al emplear indicadores bibliométricos, se pretende ofrecer una medida de la calidad "media" del trabajo continuado de individuos o grupos, sin aclarar por qué un trabajo bueno compensa a otro malo, por ejemplo, o en qué sentido decimos que dos grupos de investigación exhiben una calidad semejante cuando uno de ellos sólo es capaz de producir trabajos corrientes mientras que el otro puede producir algunos excelentes y otros que pasan inadvertidos. La noción de calidad científica aplicada a agregados quizá debiera precisarse un poco más en relación a aquella otra a la que parece estar subordinada.

Otra clave aparece al interpretar separadamente el impacto *a corto plazo* del de *a largo plazo*: quizá podría establecerse una correspondencia de esta conceptualización con la que señala la diferencia que existe entre la calidad estimada a primera vista y la calidad que se esgrime como causa de las consecuencias permanentes en una disciplina. En el primer caso, los expertos apuestan *a priori* por la importancia que puede alcanzar un trabajo, sobre la base de determinadas características; en el segundo caso, expresan la importancia que ya ha logrado, y la calidad se le atribuye, podemos decir, *por definición*, una vez que ha contribuido de modo significativo a la ciencia. Esta distinción está llena de implicaciones para la política de la ciencia, si es que la calidad de investigación se adopta como un criterio relevante para la planificación y la toma de decisiones. Si tomamos como aceptables las discrepancias entre juicios a corto y a largo plazo, entonces la calidad no es una garantía de contribución a la ciencia, lo que impide asumir sin más que aquellos trabajos que no triunfaron carecieran de calidad.<sup>48</sup>

Estrechamente relacionado con esto último está el uso práctico que se persigue para los indicadores de impacto como estimadores de la calidad. Estos indicadores se emplean para juzgar la viabilidad y la calidad de los proyectos de investigación, pero obviamente no se refieren directamente a ellos, sino a los trabajos recientes del grupo que quiere acometer el proyecto. Así, la estimación de la calidad de los trabajos anteriores es relevante para la delimitación de las expectativas futuras. Se trata, en realidad, de una apuesta en la que el indicador de calidad se toma como muestra de la capacidad de los investigadores. El significado de la calidad científica es entonces algo así como el "buen hacer" de los investigadores.

Quizá la cuestión más interesante desde el punto de vista de la conceptualización de la calidad científica sea la relativa a las dependencias que existen entre los tipos de impacto y la calidad que denominan "básica". Ésta es sólo un componente de la noción de calidad, y el problema es que no sabemos en qué medida la visibilidad, o impacto a corto plazo, depende de esta calidad básica ni tampoco si el triunfo como contribución permanente, en iguales condiciones de relaciones públicas, está determinado por esa calidad básica. No se sabe, en fin, si esta calidad restringida es un factor determinante o en qué grado lo es. La calidad

<sup>48</sup> Claro que por detrás de este problema de la relación entre calidad y contribución asoma otro acerca de la posibilidad de una gradación de las contribuciones.

amplia como causa del triunfo es una caja negra, porque no se establece el grado en que la astucia en las relaciones públicas puede compensar la diferencia en calidad básica. Claro que la noción amplia permite sortear el problema desde el punto de vista de la gestión de la ciencia, que puede conformarse con saber quiénes son aquellos cuya actividad genera más efectos, sin preocuparse de cómo lo consiguen. Pero deja un poso de perplejidad e insatisfacción si lo que se quería era escarbar un poco bajo los mecanismos superficiales del funcionamiento de la ciencia.

De todos modos, es muy interesante concebir la calidad en este modo extendido, considerando que forma parte esencial del investigador no sólo tener a punto su instrumental y aplicar métodos rigurosos, sino también el definir o detectar problemas interesantes y la capacidad para hacer llegar eficazmente sus descubrimientos a sus colegas. En cierto sentido, esto recuerda la posición que destilan distintos trabajos de Merton, Price y Ziman de que la ciencia no conocida es ciencia que no ha existido, y refuerza la idea de que la ciencia tiene un ineludible componente social también en sus aspectos más formales. Los dominios de la ciencia no son los del conocimiento cierto, sino los del conocimiento que ha superado los procesos de crítica y elaboración más severos, que son los que aplica la comunidad científica en su conjunto.

Merece la pena abundar con otro ejemplo que refuerza la mencionada conexión entre los problemas prácticos de evaluación de la ciencia y los teóricos acerca de la calidad o el valor de los resultados de la actividad científica. En la misma línea de la evaluación de la investigación realizada en universidades, refiriéndose en este caso a la experiencia británica llevada a cabo por la *University Grants Committee* (UGC) en los años 1985-86, A.J. Phillimore (1988) también trata, aunque de manera breve, algunas cuestiones conceptuales acerca del rendimiento de la investigación. La adecuación o validez de ciertos indicadores se vincula con la claridad de los conceptos a los que se supone hacen referencia. Señala la falta de una definición común acerca de qué sea el *rendimiento de la investigación* (*research performance*) (p.263), lo que significa un escaso acuerdo sobre qué se está tratando de medir. La traducción de sus palabras es el mejor modo de acercarse a su tratamiento de la noción de rendimiento:

Pueden distinguirse como mínimo cuatro aspectos del rendimiento:

- resultados [output] son la producción de hallazgos de la investigación;

- impacto [impact]: es el efecto que tales resultados tienen sobre su audiencia;

- calidad [quality]: es el "mérito" de los resultados de la investigación;

- utilidad [utility]: es la "explotabilidad" potencial de la investigación, por ejemplo por la industria.

Además, estas nociones agregadas o generales pueden ponderarse para tomar en consideración conceptos como "eficiencia" (rendimiento en relación al coste de producción) y "productividad" (rendimiento en relación al personal) (Phillimore, 1988: 263).

Esta concisión se ve enriquecida por la lista en la que se asignan los indicadores empleados por la UGC a cada uno de los cuatro aspectos del rendimiento que se acaban de señalar: para los resultados, recuentos de publicaciones; para el impacto, recuentos de citas; para la calidad, número de becas, estudiantes en investigación, premios, presencia en comités científicos o de revistas, reputación entre los pares; para la utilidad, los recursos externos conseguidos, las patentes y licencias, o el personal externo o contratado. Phillimore subraya las principales complicaciones que presenta cada indicador. (El trabajo también recoge un esquema que resume una revisión del rechazo o la aceptación de estos indicadores como reflejo del rendimiento de la investigación).

Una de las cuestiones centrales en este trabajo es la de establecer las condiciones de aplicabilidad y de adecuación de los distintos indicadores del rendimiento. Hasta ahora, se ha visto que la imprecisión conceptual acerca del rendimiento puede ocasionar serios problemas, que se suman a los que cada indicador presenta por sí solo. Pero, además, existe otra fuente de complicaciones en lo que respecta a la validez de los indicadores: el problema de las unidades de análisis. El consejo de comparar iguales con iguales, en el que ya había insistido el trabajo de Martin e Irvine, no se cumple satisfactoriamente con la unidad elegida por la evaluación de la UGC, el *cost-centre*, que podría traducirse como unidad presupuestaria, ya que existen diferencias significativas en cuanto a tamaño y tampoco se aplican normalizaciones que eviten las diferencias entre áreas. Entre otras recomendaciones más de detalle del trabajo están, por lo tanto, la de construir escalas válidas para las

comparaciones entre áreas y aplicar relativizaciones que salven las diferencias de tamaño y presupuestos.

El objetivo principal del trabajo de Phillimore no es, como habrá podido deducirse, el de precisar conceptualmente qué sea la calidad científica; sin embargo, este enfoque tan ligado a la práctica evaluadora resulta muy clarificador en el debate acerca de las nociones relacionadas con la valoración de los logros científicos mediante medidas externalistas.

De su taxonomía acerca del rendimiento de la investigación llaman la atención varias cosas. La primera, de carácter general, es el hecho de que se considere necesaria una aclaración conceptual después de la considerable cantidad de trabajos que previamente se habían ocupado del tema de la calidad o nociones afines. Seguramente deba interpretarse esto como un reflejo del escaso acuerdo logrado, como señalaba el propio autor. La segunda, la ausencia de un tratamiento profundo de las dimensiones del rendimiento que se distinguen: una simple frase sirve para separarlas entre sí, sin ulteriores consideraciones. Puede pensarse que lo que interesa a Phillimore es meramente destacar que pueden delimitarse varias dimensiones, lo cual causa un grave problema que se ignora al emplear indistintamente todas las medidas citadas como indicadores del rendimiento. En cuanto a una mayor precisión, no se sabe si considera que son prácticamente autoexplicativas o que desarrollarlas es un problema en sí mismo, aún pendiente y que excede su propósito.

En tercer lugar, y a falta de un tratamiento más detallado, parece desprenderse que todas las dimensiones son independientes entre sí, es decir, que, en principio y siempre que haya algún resultado, cualquier combinación de resultados (cantidad), impacto, calidad o utilidad es posible. Sea o no ésta la opinión de Phillimore, es interesante en sí misma, puesto que establecería marcos valorativos separados para cada una de las dimensiones. Pero, sobre todo, abriría claramente la puerta a la posibilidad de que existieran resultados que carecieran de calidad, impacto o utilidad, y que aún así fueran científicos en algún sentido.

Por último, es destacable la introducción de un ámbito analíticamente externo a la autonomía de la actividad científica, y que ya había estado presente en el planteamiento del libro *Quality in Science*. Phillimore habla de la *utilidad* de las investigaciones. Esta dimensión no parece estar siempre al alcance de los científicos, especialmente de aquellos que cultivan lo que se llama *ciencia básica*, y parece ser ésta la razón por la

que no se le había prestado atención en las anteriores conceptualizaciones. Sin embargo, no resulta descabellado que la utilidad se convierta en un criterio para quienes evalúan la ciencia externamente, porque los responsables de su financiación sí pueden estar interesados en la aplicabilidad práctica de los descubrimientos. Lo más interesante de esta inclusión es que pone sobre la mesa el hecho de que quizá era una de las dimensiones implícitas en las anteriores discusiones en torno a la valoración de los resultados de la investigación, y que fuera la posible causa de algunas confusiones.

La posición de Phillimore, expresada de manera tan sencilla, tiene pues un cierto grado de novedad, a pesar de ser evidentemente tributaria de algunas aportaciones clásicas en ese terreno y comentadas a lo largo del presente capítulo. Al centrar su atención en la valoración del *rendimiento* de la investigación, localiza algunas de las causas de los problemas que plantea la medida de esa noción. En realidad, la fuerza de su propuesta está más en la aclaración de la noción general por medio de la especificación de sus componentes que en la definición que hace de cada uno de esos elementos.

#### 5.2.6. ¿Y si preguntamos a los científicos?

Como alternativa a los esfuerzos de precisión conceptual, por mucho que éstos hayan podido tener en cuenta determinados fenómenos de la práctica científica real, existe otro enfoque para determinar cuál es el contenido de la noción de calidad científica, uno que podríamos denominar *empírico*: consiste en preguntar a los científicos, directa o indirectamente, qué entienden por calidad en la ciencia. Un ejemplo de este tipo es el estudio llevado a cabo por S. Hemlin (Hemlin, 1992; Hemlin, Montgomery, 1990). Como condición previa, considera que existe un marco general para la formación de los juicios acerca de la calidad científica. Concibe este marco como un compuesto en el que pueden intervenir, en grado variable según la ocasión, una serie de hasta siete factores, que configuran la siguiente lista: *indicadores de calidad; esfuerzo de investigación; investigador; entorno de investigación; efectos intra- y extracientíficos; organización y política de investigación; y, por último, financiación de la investigación*. Con esta base se puede seguir la pista a los diferentes significados que puede recibir el concepto de calidad en el

contexto de la ciencia. Parece razonable interpretar que cada uno de los factores señalados anteriormente representan una especie de *ámbito de criterios*, de modo cada uno de ellos incluye cualidades importantes para la atribución de calidad científica y que los criterios de evaluación que se aplican en cada juicio concreto pueden cambiar dependiendo del factor o factores que se hayan enfatizado. Así, se intenta dar cuenta de la variabilidad que puede percibirse en la práctica real, en la que los juicios sobre la calidad parecen a veces responder a distintos significados del concepto mismo de la calidad científica.

Descendiendo a un nivel más específico, Hemlin propone una distinción que afecta a los componentes de la noción de calidad y mediante la que separa los *atributos* de los *aspectos*. Como "atributos" enumera los siguientes: *corrección, novedad, rigor, relevancia científica, relevancia extra-científica, profundidad, amplitud, productividad, relaciones internacionales*. No da una definición precisa de cada uno de ellos, quizá porque los toma como suficientemente comprensibles, o al menos operativos, dentro del vocabulario de los científicos. En cuanto a los "aspectos", parecen ser dimensiones en las que pueden darse los atributos o rasgos; son *problema, método, teoría, resultados, razonamiento, y estilo literario*. No todas las combinaciones entre atributos y aspectos son posibles, es decir, no todos los atributos tienen sentido en cualquier aspecto.

Con este enfoque, afronta un estudio empírico mediante cuestionarios a científicos de diversas disciplinas acerca de sus concepciones sobre la calidad científica. Además de plantearles una pregunta abierta, se les pidió que puntuaran, por su importancia, a los distintos atributos en relación con los aspectos en los que podían tener sentido. Sin pretender un comentario crítico en profundidad de los resultados obtenidos por este estudio, sí puede decirse al menos que el autor encuentra algunas razones que apoyan la idea de que la calidad, tal como la conciben los científicos, está compuesta más bien de atributos que de aspectos. También se dibujan algunas diferencias interesantes entre las disciplinas, particularmente entre las grandes divisiones (naturales, médicas, tecnológicas, sociales, humanidades). En cuanto a las preferencias por unos atributos u otros, el rigor, la corrección y la profundidad obtienen una puntuación algo más alta, aunque sin grandes diferencias; la productividad y las relaciones extra-científicas ocupan los últimos lugares. Es interesante señalar que se producen algunas discrepancias al

comparar las respuestas a la pregunta abierta con las respuestas a la lista que se propuso; quizá la causa fuera que el marco general se configuró de diferente modo, aunque se desconozcan cuáles fueron exactamente los motivos. Concluye que tanto la distinción entre atributos y aspectos como la consideración explícita de los factores que componen el marco general pueden ser de ayuda en la sistematización de los juicios sobre la calidad de trabajos o propuestas de investigación.

Tomaremos este ejemplo como representante de un posible tipo de enfoque que puede adoptarse ante la necesidad de precisar cuál es el significado preciso del concepto de calidad científica en los contextos en los que ésta es evaluada. Por lo tanto, los comentarios que siguen deben ser tomados más bien como referidos a las posibilidades y limitaciones del enfoque general que como críticas o anotaciones al trabajo anterior, cuya finalidad es excedida por las pretensiones que aquí se han asumido.

Es importante señalar que uno de los fundamentos de este tipo de enfoque es el concebir la calidad como un complejo de rasgos que son intrínsecos al *resultado* del trabajo científico. Esto implica, por un lado, que la molesta cuestión que plantea la variedad de juicios sobre un mismo trabajo tiene entonces su raíz en las posibilidades que, por combinatoria de rasgos y pesos, se derivan de esa complejidad, y no en una vaguedad atribuible al concepto. Por otro lado, parece significar que la calidad científica tiene un cierto carácter objetivo, lo que permite que pueda precisarse su naturaleza mediante el análisis del uso común que de ese concepto realizan quienes disfrutan de un acceso privilegiado a él, que son los propios científicos. De este modo, se intenta establecer una categorización de los componentes de los juicios acerca de la calidad como paso previo a la búsqueda de aquello en lo que consiste el acuerdo general: sumando, por así decirlo, las opiniones de todos los que por su actividad y habilidades están más cercanos a ese concepto, seguramente podremos acercarnos a determinar qué es realmente la calidad científica. La explicación del problema -la variabilidad de los juicios sobre la calidad- ayuda así a resolverlo.

Sin embargo, queda pendiente el modo en el que se podría construir una escala que permitiera unificar en la práctica los juicios sobre la calidad. Una razón de esta insuficiencia es que se mezclan fácilmente lo que parecen ser condiciones necesarias para que el trabajo sea considerado científico con propiedades que parecen ser solamente deseables o recomendables y que pueden ser poseídas en distintos grados; resulta

imprescindible decir para cada propiedad si es necesaria o sólo recomendable. Para éstas últimas habría de pergeñar medidas que permitieran discriminar esos grados; además, sería necesario establecer los criterios de ponderación que asignen un peso a cada rasgo en el valor final del trabajo. Estas cuestiones podrían ser quizá resueltas con ulteriores trabajos empíricos que fueran suficientemente finos. Claro está, de todos modos, que la construcción de escalas no es un objetivo irrenunciable, y que una opción válida dentro de este enfoque sería la de descartar la posibilidad de tales escalas de aplicación generalizada para evaluar la calidad, y conformarse con delimitar cuáles son los condicionantes contextuales que intervienen en cada juicio sobre la calidad.

Algo más problemática que la cuestión de construir escalas aplicables a la evaluación práctica resulta la alternativa ante la que nos sitúa este enfoque en relación con la interpretación de la *opinión común* que construimos a partir de las opiniones individuales de los científicos. La objetivación del concepto de calidad, que antes presentábamos como una asunción requerida por la interpretabilidad de la opinión compartida, no es la única posibilidad. Por el contrario, alguien puede sostener que, al preguntar a los investigadores acerca de su concepción de la calidad científica, sólo podremos obtener un cuadro de sus ideas, prejuicios o consensos acerca de ella; que este cuadro estará determinado por factores sociales, psicológicos, de formación cultural o de pertenencia a una comunidad, etc., y no por una colección necesaria de rasgos objetivables. La alternativa es, entonces, la siguiente: ¿la calidad en la ciencia se reduce a un acuerdo entre científicos o por el contrario ellos ofrecen un medio privilegiado para desvelar algo que está ahí? En otros términos, esto significa que debe decidirse si la lista de rasgos es arbitraria, procedente sólo de un acuerdo coyuntural, o existen razones permanentes para que los científicos incluyan esos rasgos y no otros. En el primer caso, habrá que justificar por qué la moda caprichosa en los gustos de los investigadores ha de tener alguna relevancia en la valoración de un trabajo sobre otro, especialmente para observadores externos. En el segundo caso, deberán especificarse cuáles son esas razones y aclarar los motivos por los que debemos erigir a los científicos como jueces últimos en este terreno.

Por último, debemos notar también que se echa en falta alguna consideración acerca de la relación entre la calidad de un trabajo científico y su éxito, en cualquier sentido que se defina. Parece, por tanto, que la

evaluación de la calidad científica puede realizarse completamente *a priori* y sin tener en cuenta este aspecto, lo que permite conjeturar que la calidad podría no ser una condición necesaria para el éxito aunque fueran preferibles (¿por qué?) aquellos trabajos o propuestas con calidad.

En una línea similar a la de la propuesta anterior, que considera que podrían mejorarse (normalizarse) los juicios acerca de la calidad científica, se encuadra la propuesta de K. Buchholz (1993). Su planteamiento parte de la necesidad de alcanzar una caracterización del trabajo juzgado que sea tanto *objetiva*, es decir, accesible a la prueba, como *adecuada* al contenido y nivel cognitivo del trabajo. Para él, el problema general es cómo distinguir lo que contribuye al progreso científico de aquello que no contribuye a ese progreso. Cree que aquello que no contribuye al progreso carece de calidad; en realidad, iguala "carácter científico" y "calidad": si no contribuye al progreso, entonces no es científico. El problema particular que aborda es el de ordenar los resultados, trabajos, propuestas de proyectos, etc., con una precisión mayor que la de *exce-lentes y malos*, en la que ya existe cierto acuerdo mediante los análisis cuantitativos y por expertos; la cuestión es clasificar los trabajos del rango intermedio, es decir, establecer los criterios que justifiquen la preferencia, dentro del nivel más común, por unos trabajos en relación a otros. Tras comentar algunos intentos de evaluación de los resultados y señalar sus principales insuficiencias, Buchholz afronta la especificación de una escala que, basada en argumentos acerca del progreso atribuible a cada trabajo, permita obtener juicios reproducibles y con una mayor diferenciación que la habitual.

Sea cual sea el tipo de teoría (distingue dos básicos: mecanicista o probabilística) lo que a Buchholz le importa es el hecho de la **extensión del conocimiento**: ésta puede tener lugar en el plano teórico, que debe ser accesible a prueba, o en el experimental, confirmando o criticando el marco teórico vigente. Afirma que, aunque la calidad científica requiere relaciones obvias con lo teórico y lo experimental, el progreso científico puede apoyarse primariamente en el dominio teórico. Con poco más que este equipaje, agrupa ejemplos históricos de contribuciones científicas ampliamente reconocidas en varios modos principales de producir progreso científico. En primer lugar, el relativo a la formulación de *nuevas teorías* (se supone que incondicionalmente ciertas; no dice qué pasa cuando se muestran parcialmente inadecuadas, erróneas o

cuando no son totalmente aceptadas); en segundo lugar, cita casos modélicos que han supuesto la *extensión de teorías*; el *refinamiento de teorías* existentes también lo considera un progreso estimable; por último estarían, los *pasos experimentales*, confirmatorios o refutatorios de teorías. Concluye diciendo que "con la enumeración de estos tipos de trabajo científico y de descubrimientos, la jerarquía de logros e innovaciones es obvia" (p.210). Reconoce que no parece tener sentido atribuir un rango superior a los logros de originalidad comparable e importancia complementaria sólo por ser de carácter teórico o práctico, por lo que su lista de criterios no distingue en su secuencia jerárquica entre argumentos teóricos y prácticos. (Pero no dice cómo se mide esa originalidad e importancia comparables) Además, señala que existe un gran número de resultados de la investigación que pueden encuadrarse bajo las categorías anteriores y que, sin embargo, pueden carecer de calidad, ya que son consecuencias directas de logros precedentes, lo cual configura una zona gris que plantea el auténtico problema de la evaluación de la calidad científica.

Llegado a este punto, Buchholz distingue tajantemente lo que llama *trabajo de investigación*, que persigue la creación o revelación de nuevos conocimientos, incluyendo fenómenos, del *trabajo científico*, aquél que lo consigue efectivamente. El trabajo de investigación, dice, no es científico por sí mismo. Puede decirse entonces que el *éxito* en la producción de conocimiento nuevo es, así, el criterio para delimitar lo científico, y por ende, lo que tiene calidad.

Según Buchholz, aunque sea difícil identificar la calidad científica en la mayor parte de los casos, hay que pronunciarse. La solución: un conjunto de criterios bien definidos y consensuados ayudarán a resolver el problema. Propone que, en primer lugar, debe prestarse atención a las teorías y experimentos generalmente aceptados como progreso científico, para extraer de ellos las características que son significantes para su calidad, teniendo en cuenta aquellos criterios generales que constituyen las condiciones básicas, como la originalidad y la significación teórica. En segundo lugar, hay que referirse al "conocimiento de fondo" característico de las escuelas establecidas en cada disciplina (posiblemente algo parecido a las "tradiciones de investigación" del Lakatos) para identificar, en su práctica habitual o en sus declaraciones en los grupos de trabajo, los criterios básicos que adoptan. Parece que lo que Buchholz propone sería una especie de análisis empírico de tipo cuantitativo

acerca de las prácticas científicas, aunque desde luego no lo expresa en estos términos.

Para poder evaluar la investigación común, encuentra una solución concreta: la lectura de la literatura científica actual "permite identificar y formular algunas características y criterios típicos para esa clase de resultados de la investigación", estrechamente vinculados al progreso incremental en la ciencia: incluye la extensión de un área, ya sea por extrapolación o por rellenar lagunas por intrapolación, así como algunos aspectos externos, no científicos, relacionados con el interés industrial o público. También pueden identificarse los aspectos negativos, que son aquellos que conllevan el desperdicio de recursos: esfuerzo innecesario sin posibilidad obvia de éxito; repeticiones enésimas; o trabajos sin objetivo científico obvio. En una tabla, detalla un poco más los criterios anteriores, especificando 25 puntos que agrupa en siete categorías, aunque sin especificar ninguna escala en cada punto ni dar instrucciones para ponderar las valoraciones obtenidas en uno de ellos con el fin de conseguir una valoración global; no obstante, afirma haber clasificado satisfactoriamente varios cientos de trabajos con estos criterios.

Concluye que el juicio acerca de la calidad científica requiere expertos para superar las limitaciones que presentan los métodos cuantitativos referidos a la eficiencia (introduce aquí un nuevo sinónimo de la calidad) científica. Estudios de caso, basados en juicios de expertos, pueden dar las pistas para correlacionar los datos de investigaciones cuantitativas con pruebas directas del carácter científico ligadas al contenido interno y la lógica de la disciplina analizada. Para superar las limitaciones del juicio de los expertos, ha de contarse con criterios adecuadamente formulados y seleccionados, de modo parecido a como se ha mostrado en ese trabajo.

Ciertamente, la propuesta de Buchholz que se acaba de resumir exhibe una notoria radicalidad. Sin embargo, no parece ofrecer una solución lo suficientemente útil y aplicable como para justificar la toma de posiciones tan extremas. Adopta un reduccionismo que iguala conceptos como "carácter científico", "calidad", "importancia" o "eficiencia", sobre la base del progreso científico alcanzado, a costa de despreciar toda aquella investigación llevada a cabo por científicos, pero que finalmente no haya sido coronada por un éxito claro.

Resulta completamente arbitrario, excesivo y un tanto ingenuo tomar el progreso científico, definido mediante superficiales referencias a

la línea popperiana de filosofía de la ciencia, como criterio fundamental para la valoración de los resultados de las investigaciones. Ya se ha comentado que la noción de *progreso en la ciencia* es objeto de un debate profundo y duradero. Convertirla en piedra angular implica hacer problemático todo lo que dependa de ella. En todo caso, esta postura representa el máximo énfasis en cuanto a la vinculación entre *calidad* y *éxito*: la calidad sería una condición necesaria y suficiente para el éxito, lo que aquí es decir para el progreso científico; a su vez, éste queda raquíticamente definido como la *extensión del conocimiento*.

La pretensión por minimizar la disparidad de pareceres en los juicios acerca de la calidad conduce a una objetivación del concepto de calidad: los rasgos que definen la calidad están ahí y pueden ser reconocidos inequívocamente; sólo hace falta ponerse de acuerdo acerca de cuáles se han de considerar; aunque quizá no pueda conseguirse una lista exhaustiva, sí puede lograrse un acuerdo suficiente.

Finalmente, no se soluciona nada: su propuesta práctica consiste en aconsejar hacer lo que precisamente hacen los expertos normales: estar al día y conocer la propia disciplina. Queda del todo pendiente la cuestión del establecimiento de escalas concretas y de la aplicación de los criterios que menciona. Si a esto se añade la dificultad, no tratada, que presenta la obtención de una lista de criterios de valoración, estamos como al principio, quizá algo más confusos. Para llegar a la conclusión de que aquello que aparece en los manuales universitarios de ciencia como contribuciones admitidas tiene valor científico, no hace falta discutir mucho. Pero aceptar que todo lo demás carece del menor valor para la ciencia es, a todas luces, excesivo. Esto último parece ser el resultado de una pobre concepción sobre qué es la actividad y el conocimiento científicos.

El deslizamiento hacia el polo que objetiviza los rasgos constituyentes de la calidad científica es algo que, cada una con sus matices, comparten las dos propuestas precedentes, la de Hemlin y la de Buchholz. La discordancia de juicios sobre la calidad procedería de la complejidad constitutiva de ese concepto, la cual permite que cada juez, cada experto, pueda acentuar unos rasgos distintos. La solución estaría, entonces, en hacer explícitos los componentes de la calidad y acordar una importancia para cada uno de ellos. Una dificultad de este planteamiento reside en que, aún admitiendo que una variedad de razones pueden estar en la base de la calidad de un trabajo, todavía requiere una explicación de

por qué unos expertos dan más importancia a unos rasgos frente a otros; en otras palabras, el desacuerdo en los juicios sobre la calidad sigue en realidad inexplicado. Posiblemente sea esta dificultad en explicar el desacuerdo la que conduce a reforzar el carácter objetivo de la calidad, atribuyendo las discrepancias entre los jueces, en último término, a las limitaciones, carencias o imperfecciones de los individuos. La búsqueda de un acuerdo acerca de los componentes de la calidad sería entonces el modo de salvar esas insuficiencias individuales, de modo que la imagen colectiva resultante reflejaría más adecuadamente los caracteres propios de la calidad en la ciencia.

G. Sonnert (1995) aborda un aspecto del problema de la evaluación de la calidad en la ciencia que guarda una estrecha conexión con la cuestión acerca de la objetividad de los juicios de los pares. Sonnert centra su objetivo en el examen de qué factores determinan las evaluaciones de la calidad científica en la revisión por pares. Para analizar estos juicios, llevó a cabo un simulacro en el que una serie de evaluadores, reconocidos en el área de la biología estadounidense, tasaron la calidad científica de selectos investigadores sobre la base de sus *curricula vitae* y otros elementos de ese tipo, en unas condiciones muy parecidas a las que se darían en la decisión por una plaza académica. Una de las preguntas concretas que interesaron a Sonnert fue la de si la predicción de los juicios sobre la calidad científica podía ser mejorada con medidas cuantitativas más sofisticadas que los recuentos de publicaciones normalizados respecto al prestigio de la revista o el tipo de autoría.

Sonnert asume que, de modo general, existen dos marcos en el que concebir qué sea la calidad: para la tradición mertoniana, la calidad del trabajo científico está en relación directa con su contribución al fin principal de la ciencia como institución, que es el de aumentar el conocimiento certificado; para la sociología de la ciencia de tipo constructivista, los criterios sobre la calidad son relativos y en cierto modo arbitrarios, dependientes del contexto científico.

La cuestión central es el modo en que pueden emplearse los resultados de la evaluación por pares para validar las medidas cuantitativas. La situación, ampliamente sentida, acerca de la utilidad de los indicadores cuantitativos y las revisiones por expertos es que, entre ambos, mantienen una insatisfactoria complementariedad. Mientras los primeros aportan *fiabilidad* y objetividad, adolecen de una escasa *validez*, es decir, no se sabe si miden exactamente lo que se espera de ellos; por el contrario,

aunque a las evaluaciones de expertos se les supone validez en el sentido anterior, les falta objetividad y fiabilidad (parecen existir notorias discrepancias).

Lo interesante, a este respecto, es la duda que despierta Sonnert: ¿realmente los jueces evalúan la calidad, o algo relacionado, pero conceptualmente distinto, como la visibilidad o la eminencia? (p.39). El hecho es que el juicio de los pares ha sido tomado muchas veces como criterio absoluto de validación de los indicadores cuantitativos, sin tener en cuenta la posibilidad de que ese mismo juicio debiera ser evaluado de acuerdo a algún patrón de calidad normativo, independiente. Tanto el acuerdo de los juicios de pares con algunas medidas cuantitativas como su discrepancia deberían ser recibidos, entonces, de manera muy cautelosa.

Sonnert compara las clasificaciones derivadas de las evaluaciones por los pares con las que resultarían de unos veinte distintos indicadores cuantitativos, algunos simples y otros más elaborados, tanto bibliométricos como de otros tipos que tenían en cuenta el prestigio del lugar de formación o de trabajo, por ejemplo. El principal resultado para lo que aquí interesa es muy llamativo: el mejor predictor de la evaluación por los pares fue la *productividad anual*, en términos de trabajos publicados. Este sencillo indicador superó a otras medidas más sofisticadas, aunque alguna de las basadas en recuentos de citas (concretamente, el número medio de citas recibidas por los tres trabajos más citados de cada autor) estuvo cerca de la adecuación conseguida por ese indicador. Para Sonnert, una posible explicación de este hecho es que quizá los evaluadores no midieran estrictamente la calidad, sino más bien su preferencia por las publicaciones de elevada visibilidad. También considera que este resultado apoyaría la conexión entre la cantidad y la calidad en el sistema de publicaciones, a pesar de que se detecta una diferencia entre el comportamiento productivo de investigadoras e investigadores: las mujeres son evaluadas ligeramente por encima, lo que podría atribuirse a un mayor "perfeccionismo". Los meros recuentos perjudicarían entonces a las mujeres.

El trabajo de Sonnert ordena de un modo claro los términos de un aspecto central en el problema del concepto de calidad en la ciencia. Sirve para cuestionar uno de los supuestos que se habían utilizado sin apenas crítica, el de que los juicios de investigadores expertos, los pares, era el mejor tribunal disponible para evaluar la calidad. A la disyuntiva

general entre un concepto objetivo de la calidad y otro relativista, muestra que se puede pensar que ciertos juicios que se toman como reflejo de la calidad se refieran en realidad a otras cosas.

Un breve comentario general sobre el trabajo de Sonnert. Es de resaltar que lo que se juzga es la calidad científica de los investigadores sobre la base de su contribución global, en lugar de juzgar trabajos particulares. Esto pone sobre el tapete una interesante cuestión acerca de la calidad científica: ¿se refiere a la actividad, a los resultados obtenidos o a los agentes? en otras palabras, ¿de qué objeto puede predicarse el que tenga o no calidad científica? Puede evaluarse el modo en que se ha llevado la investigación, ciertos aspectos intrínsecos o la importancia de los resultados, o la capacidad de los investigadores para producir ciencia. Quizá la variedad de opiniones o juicios sobre la calidad se deba a un desacuerdo subyacente en cuanto al objeto que se juzga.

### *5.2.7. Conclusión*

El tema de la calidad en la ciencia ha dado lugar a una apreciable cantidad de publicaciones, como permite entrever la muestra de la que aquí nos hemos servido para intentar delimitar los principales puntos del debate. Aunque es fácil percatarse de la enorme atención despertada por este tema, no lo es tanto el exponer de modo completo las razones que justifican ese interés, ya que proceden de ámbitos variados: la evaluación externa (e interna) de la ciencia, el estudio sociológico, psicológico o filosófico del conocimiento científico... El concepto de calidad científica parece haberse convertido en un término genérico que se aplica a muchos aspectos interrelacionados, pero aún así cabe ver una significativa conexión entre los papeles que juega ese concepto en los enfoques señalados. Sin embargo, incluso un vistazo superficial a las páginas anteriores muestra a las claras que ha existido una cierta confusión básica acerca de la naturaleza del problema de la calidad, lo cual se refleja palmariamente en la variedad de significados y de vocabularios empleados.

## Capítulo 6

### *Los indicadores bibliométricos y la calidad*

---

#### *6.1. Síntesis de los polos básicos en el debate sobre la calidad*

El problema acerca de la calidad en la ciencia se presentó de un modo ineludible cuando se comenzó a vislumbrar un uso extensivo de los indicadores cuantitativos más básicos, los recuentos de publicaciones, en los primeros años de la década de los sesenta. Siempre ha sido obvio que unos trabajos científicos son de mayor importancia para la ciencia que otros, y éste era un hecho que, evidentemente, era ignorado por los simples recuentos. Tanto los defensores de la significatividad de los recuentos como sus críticos admitieron enseguida que existía un dimensión, la de *calidad*, que no es reducible sin más a la de *cantidad*. En principio, no puede negarse la posibilidad de que ambas sean dimensiones independientes; de hecho, son abundantes los ejemplos de autores con pocas publicaciones, pero muy importantes, y de otros con una gran cantidad de trabajos publicados, pero de escasa relevancia. Por lo tanto, asumiendo que los recuentos miden la *cantidad*, quedaba pendiente el que sirvieran también en general para medir la *calidad*, o si existían otras medidas complementarias capaces de hacerlo.

Hasta aquí, el problema se plantea en términos bastante sencillos. Sin embargo, cuando se intentó justificar esas medidas cuantitativas, comenzaron a emerger otros aspectos. La falta, en el incipiente terreno de los estudios cuantitativos sobre la ciencia, de una definición precisa del

significado de las dos dimensiones anteriores motivó que se entremezclaran los conceptos y presupuestos procedentes de tradiciones ya asentadas: la filosofía, la sociología y la historia de la ciencia, así como los presupuestos metodológicos que conformaban la propia autoimagen de los científicos, brindaron la materia que se empleó en los sucesivos intentos de clarificación conceptual en torno al significado y la validez de los indicadores bibliométricos.

Los trabajos relacionados con la medida de la calidad fueron surgiendo, en efecto, en torno a los problemas de validación de los indicadores basados en recuentos de publicaciones y de citas. Price sugirió que había una cierta correlación entre el número de trabajos publicados por un autor y su *eminencia*, como ya se ha visto; otros investigadores pusieron a prueba la adecuación de indicadores basados en las citas recibidas en relación al prestigio (o importancia, relevancia, etc.) de los autores citados o a la significación científica de los trabajos. Los esfuerzos se centraban, en estos inicios, en encontrar pruebas empíricas sólidas en que apoyar el significado de unos y otros recuentos. Las pruebas consistían en encontrar correlaciones robustas entre las medidas que se querían validar y otras que se asumían como ya validadas, tales como las resultantes de opiniones de expertos. El objetivo explícito era comprobar la utilidad de esas medidas bibliométricas como indicadores de la *calidad científica*.

Se trataba de llevar a cabo una comprobación práctica de algo que conceptualmente aparecía casi como trivial: los indicadores basados en recuentos serían válidos si pudieran sustituir satisfactoriamente las opiniones de los expertos. Lo que en aquella fase inicial se estaba asumiendo, como algo bastante natural, era que los expertos forman su opinión sobre la importancia de un autor basándose primordialmente en la calidad de sus trabajos científicos. En otras palabras, el buen nombre de un científico se asienta en la calidad de sus trabajos. Si un autor está bien valorado por los expertos correspondientes (que pueden ser sus propios colegas) será porque éstos aprecian el conjunto de sus publicaciones.

Sin embargo, cuando se atiende al trabajo de Chase (1978) acerca de los criterios normativos empleados en la selección de trabajos para su publicación se perfila otro modo de hablar de la calidad. Ya no es la calidad de la obra completa de un investigador, sino la de un trabajo concreto que aquél presenta para la publicación. No se juzga la trascendencia o importancia para el desarrollo de la ciencia, sino la

superación de unos mínimos exigibles. Esta aproximación al problema de la calidad científica representa un modo de entender esa noción que se contrapone al anterior en varios aspectos, aunque, en cierto sentido, ambas alternativas representan actividades de evaluación muy comunes en la ciencia.

A lo largo de la exposición desarrollada en el capítulo anterior, el comentario de las sucesivas aportaciones relacionadas con el tema ha servido para llamar la atención acerca de esta clase de diferencias entre las distintas conceptualizaciones de la calidad científica. Si algo tenían éstas en común, era precisamente la convicción de que era necesaria una mayor precisión al hablar de la calidad; la variabilidad de los juicios sobre la calidad científica de trabajos, propuestas o candidatos ha sido un problema incómodo y duradero. La cuestión es que, si la calidad se convierte en un criterio en la ciencia, habrá de ser identificada de manera justa, fiable y más objetiva posible; de otro modo, el criterio efectivo estaría formado en realidad por las preferencias o caprichos de cada evaluador. Sin embargo, incluso con la mejor intención por parte de estos jueces, pueden llegar a producirse apreciables diferencias tanto en el sentido como en la justificación de sus veredictos particulares.

La más socorrida explicación de estas discrepancias consiste en señalar que el concepto de calidad científica está compuesto por alguna serie de notas constitutivas y que en cada persona o situación puede darse un peso diferente a cada uno de esos elementos. Varios de los trabajos comentados asumían explícitamente una solución de este tipo. Sin embargo, sin negar esta posibilidad, merece la pena ocuparse de los puntos básicos que han configurado las disyuntivas en las que se han movido las principales posiciones en este tema, previas a la mera combinatoria de rasgos. La posibilidad que aquí se plantea es que, por el influjo de distintas tradiciones teóricas, intereses, propósitos o contextos, quizá existan modos radicalmente diferentes de concebir qué sea lo esencial al hablar de la calidad en la ciencia. La finalidad de esta exploración es encontrar un fundamento a una concepción de la calidad que sea a la vez compatible con algún tipo de indicadores bibliométricos y relevante para las prácticas de evaluación externa de la ciencia, es decir, útil en los procesos de decisión en los que esas prácticas adquieren sentido. Enseguida podrá verse el modo en que la serie de puntos que a continuación se proponen como fundamentales en el debate acerca de la calidad puede ayudar en ese sentido.

### 6.1.1. El objeto de la calidad.

¿De qué cosas puede decirse que tienen o no calidad científica? Unas veces parece importar más el modo en que se realiza la investigación, si el conjunto de actividades que la componen están *bien hechas* en algún sentido especificado; otras, el acento se pone en los resultados, en que cumplan una serie de requisitos o en que tengan unas consecuencias determinadas; y en otras ocasiones, es el individuo, su capacidad para realizar investigaciones de modo científico o para conseguir resultados, lo que ocupa un lugar central. *Actividad, resultados e investigador* no son, evidentemente, lo mismo. Sin embargo (y nada sorprendentemente), casi siempre se entrecruzan cuando se habla de la calidad científica. Por supuesto, los tres mantienen una innegable relación entre sí. Pero la subordinación entre ellos puede establecerse de distintas maneras, porque existen varias posibilidades de concebir la conexión entre la capacidad para hacer algo, el hacerlo y el conseguir con ello lo que se pretendía, relacionadas especialmente con la importancia que se le dé al éxito científico y el modo en que éste se entienda; algunos pueden tomar el prestigio científico (en tanto que reflejo de la calidad referida a las capacidades) como un resultado de la calidad de los resultados, otros pueden evaluar las contribuciones en sí para comprobar si el prestigio del autor es el merecido, etc.

El énfasis en la *actividad* podría interpretarse como una consecuencia del metodismo que suele atribuirse a los investigadores, es decir, de la confianza casi ciega en que la ejecución de una serie concreta de pasos conduce al éxito. La atención al *resultado* quizá tenga mucho que ver con las influencias de la tradición filosófica, en su búsqueda del conocimiento superior. Fijarse más en las capacidades del investigador puede tener que ver con aspectos sociológicos, particularmente con los procesos de admisión y valoración, en suma, con los varios sentidos del reconocimiento que más atrás se han distinguido. La confluencia de distintas tradiciones y enfoques como causa de la variedad de significados de la calidad se podrá entrever también más adelante en otros puntos.

La razón de que se entremezclen juicios referidos a estos objetos es bien simple: a veces no se puede elegir. Por ejemplo, cuando se evalúa el potencial científico de un candidato primerizo, no se pueden tener en cuenta los resultados (aunque siempre puede haber quienes supediten todo a haber obtenido contribuciones relevantes). La preferencia por

uno u otro aspecto como objeto de la calidad puede afectar a la definición del científico: ¿es el capaz de aplicar métodos y técnicas específicos, el que las aplica efectivamente o sólo aquél que consigue contribuciones?

Los indicadores bibliométricos que puedan construirse como reflejo de la calidad se referirán, primariamente, a los *resultados científicos*. Las inferencias que, a partir de los resultados, conduzcan a emitir valoraciones acerca de la investigación o de su autor pertenecen a otro ámbito distinto del de los indicadores bibliométricos y que depende de la concepción que se tenga acerca del proceso de producción científica.

### 6.1.2. *El nivel de agregación*

En el ámbito científico, los juicios sobre la calidad parecen estar referidos más bien a objetos concretos, individuales: esta investigación, este resultado, este científico. Se juzga si los procedimientos han sido los adecuados y han estado correctamente aplicados; si el resultado es importante, productivo, tiene coherencia interna y con lo aceptado, etc.; si el investigador conoce los procedimientos y técnicas de su disciplina o si sus contribuciones a la ciencia son consideradas valiosas. Sean cuales sean, parece que pueden darse razones que fundamenten esos juicios. En el último caso, sin embargo, puede verse cómo la valoración de la obra completa de un autor puede convertirse en el criterio principal para formar la opinión científica sobre él. Cualesquiera que sean las razones, casi siempre es posible decir si un trabajo es o no mejor que otro; pero al hablar de agregados de trabajos, (o de conjuntos de autores, o de las actividades desarrolladas por un grupo durante un año), no es fácil, sin caer en arbitrariedades, ponderar por ejemplo el modo en que los éxitos pueden compensar los fracasos o la ausencia de resultados. Si existe algún sentido en el que se pueda hablar de la calidad científica de agregados, éste habrá de ser, aunque subsidiario del relativo a elementos singulares, diferente de éste en algún sentido; no basta con "sumar" los juicios acerca de los elementos, sino que se deberán incluir rasgos específicamente aplicables al agregado completo.

Los indicadores bibliométricos, como otros cuantitativos, requieren una base estadísticamente significativa para asegurar su validez. Es posible depositar la confianza en ellos mientras sean un resultado robusto,

derivado de la estabilidad que exhiben cantidades de datos suficientemente grandes. Las características cuantificables de los documentos científicos que pueden ser relacionadas con la calidad pueden, cuando se toma un sólo documento, conectarse fácilmente también con otros factores. Si existe un sentido en que algunos indicadores bibliométricos puedan tomarse como reflejo de la calidad, será sin duda más adecuado aquél que pueda aplicarse a agregados de documentos (en tanto que representantes de agregados de resultados científicos) y no sólo a trabajos individuales.

### 6.1.3. *El carácter objetivo de la calidad científica*

Las explicaciones acerca de las discrepancias respecto a la calidad científica entre expertos considerados igualmente competentes han de enfrentarse con la siguiente disyuntiva: o bien no todos los jueces reconocen adecuadamente una serie de rasgos objetivos, o bien cada juez construye un concepto particular acerca de qué sea la calidad. Esta polaridad básica puede plantearse bajo formas diversas. Un ejemplo temprano está representado por la distinción que propusieron los hermanos Cole entre *calidad objetiva* y *calidad socialmente determinada*. Respondía a la posibilidad de que trabajos considerados de una importancia o calidad similar fueran valorados por los colegas de modo muy diferente: podían existir factores, como el prestigio del autor, que desviarán el juicio. El estudio del sistema de recompensas se ocupó expresamente de la justicia de las valoraciones. Muy relacionado con éste ámbito está también el problema del reconocimiento tardío de contribuciones importantes, el cual tiene que ver con la tensión que puede darse entre el ojo del genio individual y la aceptación general por parte de la comunidad científica.

La tradición mertoniana, como señalaba Sonnert en su trabajo, parece estar comprometida con un concepto de calidad formado por una serie de rasgos objetivos y necesarios en algún sentido; por eso recurre a factores externos para dar cuenta de las discrepancias en las valoraciones. Pero con las críticas a esta escuela por parte de los autores que pueden agruparse bajo el rótulo de "sociología del conocimiento científico" surgió también la defensa del carácter contextual de los valores científicos, incluyendo los epistémicos. Así, la valoración de un trabajo

científico sería el resultado de la acción de un conglomerado de intereses, procesos de negociación, aplicación de criterios particulares, de la concordancia con propuestas dominantes o emergentes, y cosas por el estilo. Los juicios sobre la calidad dependerían entonces de criterios cambiantes y difícilmente identificables, desde luego, nada absolutos. En este caso sería poco imaginable el establecimiento de un control de calidad sobre los juicios acerca de la calidad.

También han sido adoptadas posturas intermedias, como las propuestas que persiguen la normalización de los juicios mediante un acuerdo previo acerca de las dimensiones, rasgos, aspectos que se han tener en cuenta. Las justificaciones pueden hacer más o menos énfasis en la necesidad objetiva de cada dimensión o bien en que reflejan las opiniones generalmente sostenidas por los científicos en activo, pero lo que persiguen es llegar a un acuerdo suficiente y estable sobre el que fundar la justicia de las valoraciones acerca de la calidad, eliminando así el máximo de factores incontrolados.

Los indicadores bibliométricos pueden aspirar a reflejar las opiniones expresadas por la comunidad científica en diferentes partes del proceso de publicación más que a detectar los rasgos objetivos que presuntamente determinan el que un trabajo sea de calidad. Lo que estos indicadores podrán expresar, por lo tanto, es algún aspecto de la valoración social (dentro de la comunidad científica) alcanzada por un conjunto de trabajos. La conexión que esta valoración tenga con razones objetivas o de otro tipo es algo que, aunque útil para mejorar la comprensión de esos indicadores, no puede en principio ser abordada mediante ellos.

#### *6.1.4. La composición de la calidad*

La revisión del apartado anterior ha mostrado que, en general, se concibe la calidad como un complejo compuesto por varios atributos, dimensiones, etc. Incluso los autores que la relacionan en último término con algún fin determinado como el progreso, el aumento del conocimiento, etc., admiten que puede contribuirse a ese fin de varias maneras y que cualquiera de ellas permite decir de un trabajo que tiene calidad. El problema de determinar los componentes de la calidad depende en cierta medida de la posición que se haya asumido respecto a la

objetividad, subjetividad o convencionalidad de la calidad, como se mencionaba en el punto precedente. Lo que aquí puede decirse es que en el conjunto de rasgos con que se defina la calidad científica habrá que especificar, para cada uno de ellos, en primer lugar si su posesión es necesaria, suficiente o solamente conveniente, preferible; en segundo lugar, si ese rasgo particular puede ser poseído en diferentes grados o sólo cabe el tenerlo o no tenerlo.

Esto pone de manifiesto el cruce entre dos modos de entender la calidad. Por un lado, como un *mínimo que satisfacer* para que algo logre la consideración de científico; las cuestiones sobre los criterios de demarcación, de las que tanto se han ocupado los filósofos de la ciencia, así como los controles de calidad con que la ciencia filtra los trabajos de dudosa credibilidad no son ajenos a esta aproximación. Por otro lado, como una *ordenación* basada en la distinta valoración. Esto está más relacionado con la necesidad de elegir, de preferir, tal y como se plantea en el mundo académico, con la selección de candidatos para un puesto o en la política científica al buscar criterios internos a la ciencia para apoyar una u otra línea de investigación. Los dos enfoques no son incompatibles, desde luego; habitualmente se plantea la necesidad de ordenar cuando ya se han superado los mínimos exigibles. Pero lo que aquí importa es que los criterios propios de cada uno pueden llegar a confundirse con facilidad en las especificaciones de los componentes de la calidad científica.

Esta cuestión sólo atañe indirectamente a la posibilidad y a la validez de los indicadores bibliométricos como medidas de la calidad en la ciencia. Como éstos se construyen sobre la base de las trazas dejadas en la literatura científica por los juicios o valoraciones de los trabajos, sólo pueden llegar a captar el veredicto, pero no los procesos internos que han llevado a él. Si pudieran enunciarse de un modo más o menos correcto los componentes de la calidad, esto quizá produciría un mayor acuerdo en los juicios y daría un soporte sobre el que determinar si esos juicios son más o menos adecuados, principalmente en los procesos de evaluación contruidos expresamente al margen de la práctica científica cotidiana. Pero tampoco puede argumentarse sin más que la aplicación de distintos criterios en cada juicio invalide el conjunto de opiniones sobre un trabajo; quizá a cada investigador le ha interesado un aspecto diferente del mismo trabajo, y el trabajo sea valioso por todos ellos al mismo tiempo.

### 6.1.5. *La relación entre calidad y éxito científicos*

Esta es una de las cuestiones que más urge aclarar cuando se habla de la calidad científica: la calidad ¿asegura el éxito, o solamente contribuye a él, aumentando las posibilidades, por ejemplo? ¿Puede llegarse a logros científicos por medio de trabajos sin calidad, o lo que es lo mismo, es necesaria la calidad para alcanzar éxitos científicos? Evidentemente, las contestaciones estarán determinadas por el modo en que se conciba la conexión entre la calidad y el éxito en la ciencia. En los trabajos que se han comentado podía encontrarse o deducirse casi cualquier respuesta a estos interrogantes. A su vez, el éxito científico puede ser entendido de múltiples maneras: como la aceptación por parte de los colegas o simplemente como el haber despertado la atención de éstos; como contribución permanente al acervo de conocimientos científicos o como medio para lograrla; como verificación o como refutación; como extensión del conocimiento o como sistematización de éste... En general, parece que el éxito científico tiene mucho que ver con el progreso en la ciencia, aunque este último concepto sea bastante problemático<sup>49</sup>. De este modo, la finalidad básica que se asuma para la ciencia influirá también en el contenido del concepto de calidad. Pero la relación misma entre la calidad y el éxito es en buena medida independiente de la idea que se tenga del éxito científico, en el sentido de que es imaginable cualquiera de las conexiones que se acaban de citar entre ambos para cada uno de los anteriores criterios o hechos representativos del éxito. Quizá, por ejemplo, la importancia alcanzada sea en parte una cuestión de suerte, una vez satisfecha la calidad mínima; a veces los autores no son conscientes de todas las implicaciones de su trabajo, ni han perseguido conscientemente estas implicaciones, por lo que resulta problemático si se debe conceder el mérito solamente por las consecuencias.

Lo que está por detrás, claro, es el alcance de la visión funcional que se tenga de la calidad científica. No es nada extraño que los juicios sobre

---

<sup>49</sup> Una pequeña sugerencia sobre la valoración de la calidad como la contribución al progreso, al margen de la extensa literatura sobre la idea misma de progreso: si pensamos que las ideas, una vez publicadas, tienen una cierta vida autónoma, podemos aplicar las categorías de las relaciones ecológicas a las ideas: simbiosis, parasitismo, colonialismo, comensalismo, depredación; hay tantas relaciones imaginables que resulta muy complicado saber la medida en que una idea ha contribuido al progreso (p.e. el hecho de vencer a otra idea, refuerza la posición y la aceptación de una idea, con lo que la vencida ha contribuido al afianzamiento de la vencedora).

la calidad tengan un componente ligado con el cumplimiento satisfactorio de algún objetivo, por lo cual, cuando el fin no está bien especificado, resulta difícil hablar de calidad (aunque no imposible, ya que la calidad también puede considerarse un valor puramente estético, normativo, etc.). Puede darse, en suma, mayor o menor importancia al valor instrumental de la calidad, y en consecuencia se podrá o no tomar el éxito científico como señal necesaria (o inequívoca, fiable, parcial, probable, meramente sugestiva...) de la presencia de calidad.

Una vez puesta de manifiesto la distinción conceptual entre calidad y éxito, la fusión de ambos conceptos en uno sólo aparece como una posibilidad entre otras (a veces, la calidad se ha visto como una condición necesaria y suficiente del éxito científico). Desde luego que el interés por identificar la calidad científica trasciende casi siempre el ámbito estético o de adecuación a normas profesionales, y se dirige a los efectos valiosos que pueda tener la calidad; algún tipo de conexión entre calidad y éxito habrá de estar, por lo tanto, en la base de la utilidad de una medida fiable de la calidad.

Todas estas consideraciones conducen a una cuestión fundamental en el contexto de la evaluación de la ciencia, que es la del momento en que se pueden efectuar los juicios acerca de la calidad. Si el éxito es el criterio definitivo, sólo serán posibles juicios que podrían denominarse *a posteriori*, es decir, tras un periodo en el que hubieran podido manifestarse los efectos pertinentes; los juicios *a priori* de la calidad sólo son posibles si se presupone una vinculación más flexible entre la calidad y el éxito. Lo que parece desprenderse de la práctica científica cotidiana es que, una vez logrado el éxito científico, bien se da por supuesto, o bien importa poco, el que hubiera calidad científica. Por otro lado, son muy frecuentes los juicios *a priori* acerca de la calidad, como es el caso del control y selección de trabajos para la publicación o la elección de los proyectos de investigación que se consideran más viables, interesantes o valiosos; en estos juicios no se espera, obviamente, a que los efectos del trabajo se desarrollen para estimar la calidad que éste posee. Parece haber, por lo tanto, un sentido de la calidad en el que no resultan en principio incompatibles el que un trabajo, propuesta, investigador, etc. sean de calidad y el que su éxito no esté completamente asegurado. En este segundo caso, la calidad y la importancia son dos dimensiones claramente diferentes, aunque es razonable pensar que se espera una mayor importancia cuando algo se juzga de mayor calidad. A primera vista, los

juicios *a priori* son más cercanos a los controles sobre la superación de un mínimo exigible, mientras que los juicios *a posteriori* parecen ser más compatibles con una gradación de la importancia lograda. Sin embargo, ambos tipos de juicios pueden dar lugar a las dos clases de consideraciones.

Para la planificación y la gestión, tareas incluidas dentro de la política científica, es mucho más atractiva la posibilidad de llevar a cabo juicios *a priori* sobre la calidad científica. Pero en otras fases de la política científica puede ser necesario realizar juicios *a posteriori* mediante los que comprobar la adecuación de los resultados a los objetivos propuestos, tanto internos a la ciencia como externos a ella. Lo que importa dejar claro es que los dos tipos de juicios sobre la calidad pueden ser necesarios y que cada uno responde a una vinculación distinta entre calidad y éxito, lo que conlleva necesariamente la existencia de conceptos de calidad distintos a este respecto.

En el proceso de difusión de la literatura científica existen, como ya se ha visto, dos momentos cruciales relacionados con la evaluación de la calidad: el primero, el filtro selectivo puesto en práctica por las revistas, y el segundo, la recepción dispensada a los trabajos tras hacerse públicos. Idealmente, cada uno de ellos se ajusta mejor a un tipo de juicios: la publicación está más cerca de los juicios *a priori*, mientras que la discusión y asimilación requieren un tiempo que permita valorar el peso de la contribución efectiva. Pueden construirse indicadores bibliométricos relativos a ambos momentos: los basados en el prestigio o la calidad de la revista se referirán a la estimación *a priori* de la calidad de conjuntos de trabajos; los que empleen análisis de citas, pueden quizá aspirar a reflejar los juicios *a posteriori*. En ambos casos es inevitable un apreciable refinamiento metodológico, con la finalidad de aislar los posibles factores distorsionantes. Pero lo que debe tenerse presente es que señalan aspectos distintos de la calidad, incluso puede decirse que modos diferentes de entenderla. Cada indicador concreto habrá de precisar, dentro de esta limitación, sus supuestos y el significado que se le puede atribuir.

#### *6.1.6. Grado de adecuación al concepto interno de calidad científica*

Los indicadores bibliométricos son medidas externas a la esfera científica en el sentido de que no requieren el concurso de expertos en ciencia o practicantes activos para ser realizadas (aunque, claro está, las opiniones de éstos sí puedan ser útiles en el proceso de validación). El esfuerzo por encontrar un significado y un ámbito de validez para este tipo de indicadores cuando se refieren a la calidad ha de responder en algún momento a la cuestión sobre hasta qué punto deben reflejar las convicciones u opiniones de los propios investigadores acerca de la calidad. En un extremo estaría la posición que supone que, en la ciencia, los juicios internos sobre la calidad son aplicaciones más o menos perfectas de una misma y compartida imagen acerca de ésta, y considera que las medidas externas, como los indicadores bibliométricos de la calidad, deberían ser entonces el resultado más o menos perfecto de la aplicación de ese mismo criterio interno.

El teórico extremo opuesto podría construirse sobre la duda de que tal imagen compartida exista e incluso de que el concepto de calidad tenga algún papel relevante en la práctica científica real; ese concepto sería, entonces, el resultado de la aplicación por extensión, llevada a cabo precisamente al desarrollar medidas externas, de un concepto genérico de calidad. Así, en el proceso de construcción de cada indicador, el refinamiento requerido por la búsqueda de un significado coherente crearía un concepto propio de calidad científica, que heredaría del concepto genérico una serie de rasgos básicos, pero que incluiría otros específicos no necesariamente relacionados con ninguna imagen interna de la calidad, sino más bien con las posibilidades de la medida y con los intereses particulares de los usuarios externos de esas medidas.

La mayor parte de los trabajos que se ocupan de la medida de la calidad científica parecen inscribirse implícitamente en el primero de los extremos descritos: la validez de los indicadores de calidad depende de su ajuste a las opiniones generales de los científicos competentes. Las contradicciones o ajustes poco satisfactorios de los indicadores en relación con las opiniones de los investigadores se interpretan como pruebas de que los indicadores son limitados o inadecuados. La posición dominante, por lo tanto, parece aspirar a que los indicadores sean sustitutos de las opiniones de expertos, ventajosos respecto a éstos en

términos de rapidez, coste, objetividad, amplitud, etc., pero dependientes de ellos en cuanto a su significación.

No obstante, la visión conjunta de esos trabajos ha dado varias ocasiones para mostrar, por un lado, la sorprendente dificultad en precisar un concepto supuestamente compartido por todos los científicos, y por otro lado, para lanzar la duda de que el concepto de calidad sea intrínseco a la práctica científica y relevante en los juicios valorativos habituales en esa práctica. Moravcsik, por ejemplo, dejaba de lado este concepto y en su lugar hablaba de productividad y de progreso, como algo más cercano a las concepciones de los científicos. Sonnert señalaba la posibilidad de que los juicios de los pares no respondieran estrictamente a consideraciones sobre la calidad. Por otro lado, algunos conceptos, como el de *impacto*, han surgido directamente de las medidas con se contaba, en el esfuerzo por dotarlas de un significado. Estos indicios apoyarían una posición que reclamara para el concepto de calidad correspondiente al significado de las medidas cuantitativas la independencia respecto a las opiniones de los científicos.

Así, si la cuestión es si las medidas cuantitativas de la calidad deben o no responder al concepto de calidad sostenido por los científicos, quizá sea conveniente averiguar en primer lugar si eso es posible, es decir, si nos enfrentamos a una auténtica disyuntiva. La carga de la prueba recaería sobre quienes defienden la opción de que las medidas de calidad deben reflejar el concepto interno de calidad, puesto que son ellos quienes deberían decir cuál es ese concepto comúnmente sostenido por los protagonistas de la ciencia. Pero, a estas alturas, no es necesario insistir más en la falta de acuerdo en este punto, con diversas propuestas irreducibles entre sí. Para poder plantearse la adecuación de las medidas a un concepto previo, debería en principio disponerse de uno que fuera a la vez *preciso y estable*. Sin embargo, si algo parecen tener en común las sucesivas propuestas acerca de la constitución del concepto de calidad sostenido por la comunidad científica, es precisamente la carencia, más o menos aguda, de precisión y de estabilidad. Por *precisión* se entiende aquí la especificación diáfana, inequívoca, del significado del concepto. La *estabilidad* está relacionada con la variedad de significados posibles, es decir, si se admite o no que el significado pueda cambiar con el tiempo o que existan varias posibilidades de entre las que cada juicio interno particular toma una como criterio. Cuanta más contextualidad se admita en el concepto de calidad, más se sacrifica su estabilidad. Precisión y

estabilidad son independientes entre sí: las distintas posibilidades pueden ser descritas más o menos precisamente, y el que sólo se admita una posibilidad no garantiza en modo alguno que ésta sea delimitada con exactitud. La exigencia o la aspiración a un ajuste estricto entre indicadores y concepto interno requiere ambas, precisión y estabilidad.

¿Existe un concepto interno de calidad científica? Aquí se ha mostrado que varios trabajos han pretendido identificarlo. Aunque quede aún pendiente un análisis crítico sobre los procedimientos para determinar la opinión común acerca de la calidad, en la línea de las advertencias realizadas por Sonnert o de la posible contaminación procedente de otros ámbitos externos a la ciencia, sí puede decirse al menos que existen en la ciencia ciertos procesos internos de selección y de valoración que se pueden relacionar con la calidad, como ya se ha hecho. El funcionamiento del sistema de recompensas está muy probablemente ligado a la calidad científica, aunque no se pueden descartar sin más otros valores ajenos a ella. El mantenimiento de un elevado grado de fiabilidad en el conocimiento que la comunidad científica avala expresamente puede ser interpretado como un fin asumido por ésta, como una condición para su propia pervivencia, o quizá de otro modo; en cualquier caso, la vigilancia necesaria para lograrlo puede tomarse como un control de calidad.

Las distintas valoraciones internas que pueden tomarse como juicios acerca de la calidad científica, en principio vinculadas a consideraciones sobre el *mérito* científico, caen seguramente bajo el espacio definido por el siguiente esquema (el cual se propone como arriesgado esbozo de una complicada síntesis). La idea es que esas valoraciones internas se mueven en mayor o menor grado en una o varias de las tres dimensiones que a continuación se distinguen. La primera dimensión es la de la *adecuación* a normas, valores o rasgos, los cuales pueden ser más o menos explícitos. La siguiente lista, deliberadamente heterogénea, puede servir de ejemplo: precisión en el lenguaje, elegancia, simplicidad, capacidad explicativa, originalidad, exhaustividad, aplicación rigurosa de métodos, síntesis... Es posible que la importancia dada a unos y a otros pueda variar con el tiempo o depender del contexto. En todo caso, es razonable pensar que estos atributos no se valoran positivamente por puro capricho, sino porque se consideran acordes con los fines perseguidos por la ciencia.

La segunda dimensión se refiere a la *dificultad en la tarea científica* realizada. Tiene que ver con la admiración despertada por haber resuelto un problema que había resistido el asalto de muchos científicos anteriores, o por haberlo hecho con un uso ingenioso de los recursos (conceptuales o materiales). Se valora entonces, en cierto sentido, la pericia, la astucia científicas. La calificación del problema resuelto y la relación de éste con los medios disponibles o empleados constituirían el núcleo de esta dimensión. Éste sería el fundamento de lo que podríamos denominar la *épica de la ciencia*, la cual proporciona los héroes científicos que sirven como modelos de comportamiento. Cabría interpretar funcionalmente este proceso, debido a su claro efecto reforzador de la institución científica.

Por último, en la tercera dimensión se moverían aquellas valoraciones que se basan en los *efectos* producidos por el trabajo de los científicos, especialmente en el ámbito interno: transformación conceptual, metodológica, extensión del conocimiento, apertura de nuevas líneas de investigación, etc., en suma, aquello próximo a lo que se considera progreso científico; pero quizá habría que incluir también el impacto beneficioso en otras esferas de la sociedad a través de la tecnología o la sanidad por ejemplo.

El supuesto concepto interno de calidad científica debería estar relacionado con estos tipos de valoraciones. Si este cuadro no es descabellado, entonces se justificaría cuando menos la posibilidad de diversos conceptos internos de calidad científica. Si existe un concepto interno de calidad, parece entonces poco probable que sea preciso o estable.

Desde esta posición, el problema de la adecuación del significado de los indicadores bibliométricos con el supuesto concepto interno de calidad se plantearía en torno al grado en que esos indicadores deben incluir o reflejar algunos de esos aspectos internos. Una consecuencia inmediata es que, independientemente del grado de adecuación exigido, la validación de los indicadores por comparación con las opiniones de expertos quedaría seriamente cuestionada, ya que las discrepancias podrían reflejar una falta de acuerdo radical y legítima. Incluso podría pensarse, al contrario, que los indicadores fueran un instrumento adecuado para analizar o validar los resultados de estudios basados en los agregados de opiniones de expertos.

Todo esto pone de manifiesto una confusión básica que puede haber afectado a los intentos de medida de la calidad, la causada por no

distinguir claramente entre el *contenido* mismo del concepto de calidad y los *procedimientos* para juzgarla. Preguntar a un grupo de científicos su opinión sobre un conjunto de trabajos o de colegas puede ser un procedimiento para medir la calidad de éstos, pero no garantiza sin más un acceso a la concepción de la calidad que sostienen, ni mucho menos que esa concepción sea la misma para todos ellos. Sin embargo, los agregados de opiniones han sido tomados frecuentemente como criterio definitivo del buen o mal funcionamiento de indicadores bibliométricos referidos a la calidad.

Lo que, en relación con ese tipo de indicadores cabe extraer aquí, es que es posible concebir que el contenido conceptual y el procedimiento no sean enteramente independientes. La separación de ambos debe servir para resaltar sus mutuas relaciones. Así, en la búsqueda por indicadores bibliométricos satisfactorios, debe tenerse presente la tensión que surge entre dos polos: por un lado, lo que se quiere medir, y por otro, las posibilidades de medida. El resultado, si todo va bien, después del refinamiento de los métodos (exigido por la adecuación a lo esencial del concepto) y de la redefinición del concepto (requerida por las limitaciones o las posibilidades sugeridas por las técnicas disponibles), será un nuevo concepto, que puede incluir aspectos no buscados en un principio. El significado de los indicadores bibliométricos de calidad (como el de otros tipos de indicadores bibliométricos) puede exhibir, por lo tanto, una idiosincrasia propia, un fundamento específico no dependiente de los supuestos conceptos internos.

Los puntos tratados en este apartado como dimensiones básicas en la configuración de un concepto de calidad científica han de articularse finalmente en torno a los intereses o finalidades de quien busca un juicio acerca de la calidad. La necesidad de saber algo sobre la calidad no procede meramente de la curiosidad. En el ámbito interno, los científicos están interesados en mantener un nivel de rigor y fiabilidad, en detectar líneas prometedoras, en seleccionar al mejor candidato para su equipo, en que los méritos sean justamente valorados, en velar porque la formación de los nuevos investigadores sea adecuada... Por el otro lado, en la política de la ciencia en sentido extenso, son muchos los niveles y tipos de actuación que pueden beneficiarse de evaluaciones de la calidad en la ciencia. Unas veces hay que apostar por las mejores promesas; otras, detectar las realidades más significativas; equipos de investigación,

especialidades emergentes, áreas completas, centros de investigación o el sistema universitario completo de un país pueden ser los centros de atención en diferentes momentos de la política de la ciencia. Los requerimientos de cada caso concreto influirán en el tipo de indicador aconsejable y los aspectos más relevantes de los que tiene que dar cuenta.

## 6.2. *Una calidad medible por los indicadores*

El anterior capítulo y los apartados precedentes de éste han servido para trazar las dimensiones básicas que intervienen en la constitución de un concepto de la calidad científica. El espacio así definido delimita el terreno de las posibilidades a este respecto: no puede hablarse de un sólo concepto de calidad científica, sino de muchos casos en los que puede resultar más adecuado el énfasis en uno u otro polo de cada una de las dimensiones.

La evaluación de la ciencia de un modo relativamente masivo, dirigida hacia el objetivo de actuar sobre un determinado sistema científico, es uno de esos casos. ¿Qué noción de calidad es adecuada en él? Por un lado, ha de ser útil en ese contexto, lo que significa que debe poder servir de fundamento para determinadas decisiones. Por otro, si la política de la ciencia aspira a evitar la arbitrariedad, ha de ser posible definir indicadores suficientemente fiables de esa noción, ya que los indicadores se erigen en los mejores instrumentos disponibles para obtener una información equiparable acerca de grandes cantidades de datos.

En el terreno de la evaluación de la ciencia con vistas a la aplicación en la política científica hay que responder, pues, a varias exigencias. Una de ellas es la realizabilidad del análisis, tanto por la disponibilidad de los datos necesarios como por los medios o el tiempo empleados. El objeto de análisis es el conjunto de los resultados científicos producidos por un sistema completo; lo que se espera es una serie de indicaciones valiosas para identificar los principales rasgos que permitan fundar ciertas actuaciones sobre ese sistema. Deben proporcionarse descripciones que abarquen de manera fiable a ese conjunto completo y que respeten las mismas condiciones cuando se refieren a cada uno de sus posibles subconjuntos.

El interés de los responsables de la política científica por los indicadores de calidad es, esencialmente, estratégico, y por eso, la valoración

que necesitan es una valoración relativa. No interesan tanto las escalas absolutas, sino la posibilidad de, en un momento concreto, establecer la posición respecto a los otros, de comparar unos sistemas con otros o las partes de un sistema entre sí. Los rasgos que hacen interesantes a esas comparaciones son aquellos que permiten anticipar ciertos efectos: que una acción de política científica esté fundada significa que existen razones que hacen plausible pensar que puede lograr sus objetivos. Aunque las fases de *evaluación* y *prospectiva* se separen claramente en la descripción formal de las políticas científicas, las acciones concretas de esas políticas están indisolublemente teñidas de ambas. La necesidad de valorar qué *ha hecho* está ligada al deseo de saber qué *puede llegar a hacer* y al contrario.

Lo que se persigue es una estimación de la calidad basada en pistas objetivas, susceptible de ser aplicada a grandes conjuntos de trabajos. Gracias a la distinción entre *resultados científicos* y *conocimiento científico*, se ha podido ver que es posible juzgar los resultados (y por ende a sus productores) con independencia de su contribución final a la formación del conocimiento; es decir, existen factores no mecánicos, fuera del control de los agentes científicos, en el proceso que termina con la generación del conocimiento científico. Lo que puede esperarse de cualquier investigador es que produzca resultados; su responsabilidad es que sean máximamente prometedores a juicio de los otros colegas.

Dentro de la comunidad científica, se llevan a cabo juicios que anticipan la importancia que puede llegar a adquirir cada resultado científico. El desenlace del debate no es predecible por quienes toman parte en él (de otro modo, sería absurdo desarrollarlo). Por lo tanto, la valoración de los resultados tiene que ver más con su potencial que con su aportación definitiva. Del mismo modo que sucede con la valoración interna de los resultados científicos, desde el punto de vista de los gestores de la ciencia, el interés por saber la calidad de un trabajo se basa en que permite realizar apuestas sobre la importancia futura de la tarea que hace un investigador o que va a hacer en los próximos años. No se aspira a saber si un trabajo en particular ha tenido mucha importancia para la ciencia (lo cual ya se sabe sin necesidad de los indicadores bibliométricos y puede además carecer por completo de interés para la política de la ciencia), pero sí es necesario, al juzgar grandes agregados de documentos, conectar de algún modo con las percepciones internas.

Un modo de entender la calidad, por lo tanto, es como la *expectativa* que genera un resultado al ser valorado por los otros científicos, es decir, como la esperanza que despierta de que jugarán un papel importante en el debate científico. Los investigadores expertos en una especialidad son capaces de detectar ciertos rasgos o características que pueden despertar esa clase de esperanza. Esta calidad se juzga *a priori*, es independiente de la contribución efectiva que suponga finalmente.

Ya se ha visto que en el ciclo de producción existe un momento directamente relacionado con este tipo de juicio acerca de la calidad, que es el de la selección de los trabajos por parte de las revistas. Los trabajos publicados por cada una de ellas son juzgados por la comunidad a través de los trabajos publicados posteriormente. Puesto que cabe esperar que se aplique un rigor similar a todos los trabajos publicados por la misma revista, el nivel medio de los trabajos de cada revista estará directamente relacionado con el nivel medio de exigencia requerido por esa revista, lo cual significa que se puede usar como estimación del listón que han tenido que superar esos trabajos.

Las citas recibidas por las revistas se han tomado como un indicador de la importancia de esas revistas en la ciencia. La base última de esta suposición ha sido que las citas recibidas por los documentos son un reflejo satisfactorio de la calidad de éstos: las revistas más importantes son las que publican los mejores artículos. Sobre este razonamiento, se han propuesto y usado diversos indicadores bibliométricos para evaluar agregados de documentos a partir de las revistas en que han aparecido.

Estos indicadores han de superar dos tipos de dificultades: técnicas y conceptuales. Algunas de las primeras se tratarán en los siguientes apartados. De las segundas, conviene ocuparse aquí de la más radical: que el proceso de citación no ofrece un fundamento para interpretar que citar un trabajo esté causado por su calidad o por su contribución a la ciencia. En el apartado dedicado al análisis del documento científico se ha ofrecido una interpretación coherente de éstos como cristalización de los resultados. De la calidad de los resultados se puede hablar en varios sentidos; el que permite una aplicación más general es aquél que se refiere a la actividad generada en el debate, más en la construcción de nuevos resultados que en la contribución directa al conocimiento científico. Las citas han de tomarse, como se ha discutido detalladamente, como reflejo directo de ese papel. El nivel de citas de una revista, como efecto del proceso de selección o filtro que aplica. Al considerar agregados

estadísticamente significativos de documentos, la revista en que cada uno ha aparecido se puede tomar como una pista valiosa de la calidad de esos resultados en el anterior sentido.

En los siguientes apartados se tratarán los aspectos técnicos de este planteamiento y se discutirán las principales propuestas de indicadores construidos, para poner de manifiesto las ventajas que aporta el método de cálculo de los indicadores *puntuación decílica (P-10)* y *porcentaje en el decil superior (%SUP)*.

### 6.3. *El factor de impacto (FI) y los indicadores derivados*

El número de citas recibidas pronto se percibió como un criterio para valorar las distintas revistas: cuantas más veces fuera citada una revista, más influyente en la ciencia se podía suponer, más interés habría despertado entre los científicos. Las citas recibidas por una revista son, naturalmente, el agregado de las citas recibidas por los documentos que esa revista ha publicado.

La aparición de un documento en una revista científica significa que éste ha superado un determinado control de calidad, como ya se ha visto en capítulos anteriores. Al comparar las citas recibidas por las distintas revistas, por tanto, se está comparando indirectamente también la eficacia general de ese control o selección previa. Pero ese control forma parte de un proceso dinámico en el que interés de los investigadores por enviar sus trabajos puede afectar y verse afectado a su vez por el nivel general de la revista, por su difusión, por su especialización, etc., de modo que, en un momento dado, las revistas no cuentan con material comparable sobre el que realizar su selección. Varios factores, por tanto, además del rigor en la selección pertinente, pueden estar por detrás del nivel general de lo que una revista publica. El recuento global de citas que ésta recibe puede en principio tomarse como un reflejo del interés que ha despertado entre los investigadores, lo cual suele llamarse *impacto* (Garfield, 1963) en la investigación subsiguiente; pero, claro está, el número de citas no depende solamente de que la revista publique artículos interesantes, sino también de cuántos publique. Las revistas mayores, en términos de trabajos publicados, tienen obviamente más posibilidades de ser citadas que las que publican un número menor de trabajos.

Una solución a este problema consiste en relativizar el número de citas recibidas respecto al número de documentos publicados, obteniendo alguna medida que exprese el *número medio* de citas recibidas por cada artículo. De este modo se evita la influencia del tamaño de la revista. El *factor de impacto*, propuesto por Garfield (1972), fue uno de los primeros<sup>50</sup> indicadores de este tipo que se pusieron en práctica y ha dado lugar a una extensa literatura en el ámbito de los estudios cuantitativos de la ciencia, en buena medida gracias a su disponibilidad: se calcula anualmente para cada una de las aproximadamente 4.500 revistas cubiertas por los *Journal Citation Reports* en la actualidad. Para un año determinado ( $a$ ) el factor de impacto ( $FI$ ) de una revista ( $R$ ) es igual al número de citas emitidas durante el año  $a$  hacia documentos publicados por  $R$  en los dos años anteriores ( $a-1$  y  $a-2$ ), dividido por el número total de documentos publicados por la revista  $R$  durante esos dos años.

El *factor de impacto* de una revista en un año puede leerse, así, como el número medio de citas que han recibido en ese año los artículos que publicó en los dos anteriores. Si, a falta de otras consideraciones, se toma el número medio de citas recibidas por cada artículo como un reflejo del interés general de la revista, los valores del factor de impacto proporcionan una base objetiva para establecer un orden entre las revistas científicas: cuanto más elevado sea su factor de impacto, más importante o más relevante puede considerarse a una revista. Las aplicaciones sugeridas por Garfield en un principio se referían principalmente a la utilidad del factor de impacto para decidir acerca de suscripciones, para localizar las principales revistas de una disciplina o, junto con otros indicadores basados en las citas, como guía para los propios editores. Pero con el tiempo y gracias a la representatividad creciente de la colección cubierta por los *Journal Citation Reports*, el factor de impacto se ha convertido en la práctica en uno de los principales indicadores de la relevancia de una revista. El entusiasmo desmedido o un insuficiente conocimiento lo han llegado a convertir en ocasiones en una especie de número mágico que se toma directamente como base para el cálculo de indicadores que pretenden reflejar la importancia o la calidad de la producción científica de países, instituciones o autores. Al mismo tiempo, voces críticas han señalado limitaciones de este indicador o han

---

<sup>50</sup> Raisig (1960) y Martyn, Giolchrist (1968) habían propuesto indicadores similares.

propuestas alternativas (en ambos casos, algunas anticipadas por el propio Garfield).

Para situar la discusión acerca de la utilidad o la validez del factor de impacto es preciso ocuparse de algunos aspectos relacionados con su construcción. Quizá lo primero que llama la atención al analizar la formulación del factor de impacto sea la aparente arbitrariedad de su cálculo: por un lado, se limita a los trabajos publicados durante dos años por la revista; por otro lado, sólo tiene en cuenta una parte de las citas recibidas por esos documentos, es decir, las emitidas durante el año de referencia. La razón última es que el factor de impacto así definido, a pesar de requerir el proceso de cantidades ingentes de datos, es una alternativa técnica ventajosa, por la relativa simplicidad de su cómputo. Su justificación procede de que la muestra de citas empleada sea considerada adecuada y sea capaz de reflejar los cambios o las tendencias satisfactoriamente.

Sería menos arbitrario, desde luego, que, en lugar de usar esa muestra de citas (u otra), se realizara el cálculo con todas las citas recibidas por todos los artículos que una revista hubiera publicado desde su fundación. Pero esto, además de ser poco práctico por el enorme requerimiento en cuanto a recogida y tratamiento de información, sería muy poco útil: cabría esperar que las variaciones diacrónicas del valor del indicador fueran cada vez más lentas por la inercia del pasado, y difíciles de interpretar, porque no se sabría cuál habría sido el periodo de la revista causante del cambio en el nivel de citas recibidas. La diferente antigüedad de las revistas entraría en juego, de modo que la historia de cada una podría ser una ventaja o una carga injustificadas. Esta alternativa no es, por tanto, viable, y su planteamiento sirve sólo para justificar la *necesidad de la limitación temporal* en el cálculo de indicadores de la familia del factor de impacto.

Garfield (1972: 476) justifica su selección de los dos años anteriores en su estudio de los patrones de citación en la base del SCI, que pone de manifiesto que el artículo citado típico recibe la mayor intensidad de citas durante los dos años siguientes a su publicación. Para cualquier año, señala, entre el 21% y el 25% de todas las referencias se dirigen a artículos de una antigüedad igual o menor a 3 años. Considera, por lo tanto, que su muestra es suficiente para establecer diferencias significativas entre las revistas.

Pero también se han considerado otras alternativas. En el mismo trabajo (1972: 478, nota 27) Garfield indica otra posibilidad de cálculo que puede servir de complemento a la primera: en lugar de contar para el denominador todos los documentos publicados por la revista, tomar sólo aquellos que han sido efectivamente citados. En una línea paralela, Schubert y Glänzel (1983) han propuesto incluir solamente los tipos de documento que consideran "ítemes citables" (artículos, revisiones, notas, cartas al editor) y llaman al indicador así obtenido *factor de impacto corregido*. También se ha señalado la conveniencia de omitir las autocitas, es decir, las citas procedentes de trabajos publicados en la misma revista (Price, 1981; Noma, 1988).

Unos años después de comenzar la publicación en los *Journal Citation Reports* de los valores anuales de los factores de impacto, se incluyó otro indicador de la misma familia, el *índice de inmediatez*, que expresa el número medio de citas recibidas por los documentos de una revista en el mismo año de su publicación<sup>51</sup>. Es interesante notar que diferencias elevadas en el valor de este indicador para las diferentes revistas podrían tomarse como una fuente de sesgo en el cálculo del factor de impacto definido por Garfield, que no hace uso de esas citas.

Moed et al. (1984) pusieron en práctica otra variante en el modo de calcular el número medio de citas recibidas y lo aplicaron para la evaluación de grupos de investigación de la universidad de Leiden. Como complemento a su análisis *de tendencia* en cuanto a publicaciones y citas, realizaron un análisis *de nivel*, para el que definieron el indicador *Journal Citation Score (JCS)* que resulta, para cada revista, de dividir por el número de artículos publicados en un año el número de citas recibidas por esos artículos dos años después (e.d., al tercer año de su publicación, el que se considera el *pico* de citación más frecuente). Este valor lo usaron para calcular una media ponderada sobre la producción de cada grupo de investigación, como medida del *impacto esperado*, que luego emplearon como referencia para valorar el *impacto actual* o número de citas recibidas efectivamente por los artículos<sup>52</sup>.

<sup>51</sup> Con una intención parecida, Schubert y Glänzel (1986) propusieron un indicador más sofisticado denominado *tasa media de respuesta (mean response rate, MRT)* para expresar la rapidez de los trabajos de una revista en recibir su primera cita:  $MRT = -\log(f_0 + f_1e^{-1} + f_2e^{-2} + f_3e^{-3} + f_4e^{-4})$ , donde  $f_i$  es la fracción de documentos que reciben su primera cita en el año  $i$ -ésimo tras su publicación (tomado de Egghe y Rousseau, 1990: 265).

<sup>52</sup> En esta ratio observado/esperado sólo consideran las revistas en las que han publi-

Las diferentes posibilidades de cálculo derivadas de los distintos periodos citados que se pueden abarcar pueden generalizarse, como ha hecho Rousseau (1988), quien formaliza la expresión para el *factor de impacto generalizado*, (IPF) el cual puede aplicarse considerando cualquier número de años antes del año citante<sup>53</sup>:

$$IPF_n(Y) = \frac{\sum_{i=1}^n CIT_{Y-i}(Y)}{\sum_{i=1}^n PUB(Y-i)}$$

Rousseau también considera el caso en el que el recorrido de  $i$  comience en 2, es decir, se salte el año inmediatamente anterior, y llama a esa medida *factor de impacto truncado*. (IPF\*). (El factor de impacto de Garfield es el caso IPF<sub>2</sub> y el JCS de Moed et al., es IPF\*<sub>2</sub>).

De la aplicación experimental que Rousseau lleva a cabo sobre las revistas de matemática pura, comparando los valores del factor de impacto calculados para distintos periodos, pueden destacarse varias consecuencias interesantes. El conjunto de revistas analizado alcanza su máximo valor en el factor de impacto cuando éste se calcula para cuatro años (es decir, para IPF<sub>4</sub>, en el además encuentra indicios de una mayor estabilidad a lo largo del tiempo), con una diferencia significativa respecto al IPF<sub>2</sub>. Esto sugiere que pueden existir sesgos entre las disciplinas debidos a la diferente rapidez o antigüedad media de las citas. Egghe y Rousseau (1990a: 264) indican que el  $\max(IPF_n)$  sería más adecuado para comparaciones entre áreas que el IPF<sub>2</sub>. Por otro lado, sin embargo, no se producen diferencias significativas entre la ordenación del conjunto de revistas mediante el IPF<sub>2</sub> y la ordenación por IPF<sub>4</sub>, aunque de modo no homogéneo: para las revistas con IPF alto, la ordenación resultante es prácticamente la misma, pero para las de IPF muy pequeño, la ordenación sí resulta alterada. Para aplicaciones que sólo tomen en cuenta el

---

cado los grupos de investigación estudiados e ignoran, por ejemplo, el resto de las revistas de la misma área. Lo que refleja es si el número de citas logrado es superior o inferior al promedio de las revistas en las que han publicado, sin hacer referencia al área completa ni tener en cuenta si el impacto de las revistas empleadas es alto o bajo.

<sup>53</sup> Se ha preferido aquí la expresión expuesta por Egghe y Rousseau (1990: 264), más específica que en Rousseau (1988: 250).

orden de las revistas<sup>54</sup>, la diferencia entre usar uno u otro periodo de cálculo es poco apreciable.

Todo esto trae a escena algunas de las limitaciones del factor de impacto para dos años, el definido por Garfield, que radican en que la muestra de citas dirigidas a los dos años anteriores puede no ser igualmente representativa para todas las áreas debido a los distintos patrones de citación que existen en ellas. El propio Garfield (1976) ya había advertido que el valor del FI podía, en disciplinas con una antigüedad mayor de los trabajos citados, incrementarse si el periodo de cálculo se alargaba; relacionó las diferencias entre áreas en cuanto al *potencial de citación* con el hecho de que el número medio de referencias por artículo podía variar significativamente entre las áreas (Garfield, 1978); también había alertado sobre el peligro de realizar comparaciones entre áreas basadas meramente en el factor de impacto, debido a los distintos patrones de cita que podían darse en cada área (Garfield, 1979a).

Otros autores también se habían ocupado de estos problemas. Pinski y Narin (1976) señalaron otras limitaciones del factor de impacto, además de la de no normalizar respecto a las peculiaridades de citación de las áreas científicas: por un lado, el factor de impacto carece de una corrección que neutralice la influencia de la longitud media de los documentos en cada revista (las que publican trabajos más largos suelen ser más citadas); por otro, considera todas las citas de igual importancia, sin distinguir la revista que las emite. Marton (1983) muestra llamativas diferencias en cuanto al factor impacto medio de las revistas de distintas áreas. Moed, Burger, Frankfort y van Raan (1984b) reforzaron las aseveraciones de Garfield al encontrar diferencias entre áreas en cuanto a número medio de referencias por artículo, porcentaje de referencias a documentos de dos años de antigüedad y media de referencias de dos años de antigüedad por artículo. Además, detectan cambios en los patrones de una misma área a lo largo del tiempo. Todo ello cuestiona la validez de indicadores derivados directamente del factor de impacto y sin normalizar para realizar comparaciones entre áreas o incluso para datos evolutivos de la misma área.

Varios autores han encaminado sus esfuerzos a la superación de estas dificultades. Pinski y Narin (1976), como respuesta a las limitaciones que habían enumerado, definieron el indicador *peso de la influencia*

---

<sup>54</sup> Véanse en el siguiente apartado los indicadores P-10 y %SUP.

(*influence weight*) para valorar la importancia de cada revista. Los valores de este indicador se obtienen tras un proceso iterativo en el que cada cita adquiere un valor dependiente de la influencia de la revista que la ha emitido. Los artículos de una revista se valoran por el número de citas (ponderado en relación a la influencia de la fuente citante) que han recibido, y al final se consigue, para cada revista, una medida que refleja la suma de la influencia de los artículos que ha publicado. Los valores del indicador *peso de la influencia* se refieren a un marco configurado por un conjunto de revistas relacionadas entre sí, por lo que, según los autores, permite comparaciones entre áreas. Este indicador ha sido empleado principalmente en los estudios bibliométricos de la empresa *Computer Horizons, Inc.* (CHI), la cual genera una base bibliográfica normalizada para estudios bibliométricos a partir de los datos del *Science Citation Index*. Geller (1978), Todorov (1984) y Noma (1988) refinaron la medida del *peso de la influencia*. Anderson, Narin y McAllister (1978) lo emplean como base del *indicador de calidad bibliométrica*.

Braun, Schubert y Glänzel, investigadores de ISSRU (*Information Science and Scientometrics Unit*, de Hungría) han llevado a cabo varios estudios en los que analizan, por países y áreas la producción bibliográfica recogida en el SCI. Para comparar el impacto de las producciones de distintos países, ponen en práctica una aproximación del tipo de ratio observado / esperado, concretada en el indicador *tasa relativa de citación* (*relative citation rate, RCR*) (Braun, Glänzel, Schubert, 1985; Schubert, Braun, 1986). El indicador *RCR* relativiza el número de citas efectivamente recibidas en un periodo determinado por un agregado de documentos respecto al número de citas esperado para ese agregado. El número esperado de citas se calcula a partir del número medio de citas recibidas en el mismo periodo por las revistas en las que esos artículos han aparecido publicados. El indicador *RCR* puede calcularse separadamente para cada campo científico y ofrecer información sobre el nivel relativo de citas de cada país en cada campo, por ejemplo. Los autores defienden que el *RCR* es independiente del área y que por lo tanto sirve para realizar comparaciones entre las áreas.

El modo en que calculan las citas esperadas ha tenido algunas variantes. Al analizar el periodo 1978-80 (Braun, Glänzel, Schubert, 1987; Braun, Schubert, 1988) emplean una muy interesante alternativa al factor de impacto que cuenta las citas recibidas por los artículos publicados en *un año durante los dos siguientes*: por ejemplo, el impacto de los

artículos publicados por una revista en el año 1978 se calcula dividiendo por éstos el número de citas que han recibido en el año 1979 y en el 1980. A diferencia de las otras alternativas anteriormente comentadas, que sólo tienen en cuenta las citas emitidas en un año (es decir, un año citante y uno o varios citados), ésta realiza el cálculo de un modo conceptualmente más acorde a las razones ya conocidas sobre el *pico de citación*.

Sin embargo, en los estudios referidos a los periodos 1981-85 (Schubert, Glänzel, Braun, 1989) y 1986-89 cambia el modo de cálculo del impacto esperado: en estos casos, los años citantes y los años citados son los mismos. El impacto esperado de una revista en el periodo 1981-85, por ejemplo, se halla dividiendo por el total de documentos (citables) publicados en esos años el número de citas dirigidas a ellos por documentos publicados durante esos mismos años. Esto supone que los documentos publicados al principio del periodo tienen más posibilidades de ser citados: para un trabajo de 1981 se cuentan las citas que ha recibido en los años 1981, 1982, 1983, 1984 y 1985; para uno publicado en 1985, sólo se tienen en cuenta las citas dirigidas a él durante ese mismo año. La principal razón que apoya esta elección es técnica, la de posibilitar una mayor rapidez en la disponibilidad de los datos de un periodo (con la fórmula anterior había que esperar dos años tras el último año del periodo).

Esta serie de estudios, publicados en la revista *Scientometrics*, ha recibido una atención considerable gracias a su extensa cobertura y a su fácil acceso, convirtiéndose para muchos casos en la única fuente disponible. Desgraciadamente, algunos aspectos de su metodología pueden ocasionar interpretaciones erróneas de los datos que ofrecen. Merecen especial atención tres de ellos: el *significado* del indicador *RCR*, la evolución que puede producirse *dentro* de cada periodo estudiado, y los posibles sesgos producidos por considerar únicamente el *primer firmante* de cada trabajo.

El indicador *tasa relativa de citación* (*RCR*) es una ratio que compara las citas recibidas con la media de los documentos publicados en la misma revista. Aplicado a un sólo documento, diría si recibe más o menos citas que las que reciben como media los trabajos de su revista. Leído de otro modo, permite saber si ese documento es responsable de subir o de bajar el promedio de citas de su revista. Es simplemente una expresión de la proporción que guardan entre sí los términos de la ratio, sin decir

nada acerca de la magnitud de esos términos. Es obvio que la misma proporción puede darse entre dos magnitudes muy grandes o dos muy pequeñas, por lo que el valor del RCR puede ser equívoco:  $RCR = 1,5$  puede deberse a 18 citas recibidas en una revista con promedio de 12, o a 3 citas en una con promedio de 2, aunque estos dos casos no parecen estrictamente equivalentes. Por otro lado, pueden obtenerse valores inferiores a 1 con citas observadas y esperadas muy elevadas, siempre que la esperada sea ligeramente superior, o darse el caso contrario. En cuanto a la evolución de los valores del RCR, puede darse el caso de que los dos componentes crezcan y el indicador descienda, si el denominador crece más rápidamente, y el caso contrario. Quien interprete el indicador RCR como reflejo de la calidad supone que los valores promedio de citas a una revista son independientes de los trabajos que ésta ha publicado cuando en realidad están debidos directamente a ellos. En la misma disciplina, ser citado 20 veces es mejor síntoma que ser citado 1 sola vez, aunque el primer trabajo aparezca en una revista con un promedio (gracias también a ese trabajo) de 21 citas y el segundo en una de 0,5 citas. Por otra parte, debido a las diferencias entre disciplinas en cuanto a los valores máximos y mínimos de las revistas, el modo en que se compensan artículos sobrecitados e infracitados puede ser muy diferente (en algunas disciplinas, una sola cita de más puede compensar a 10 trabajos sin citar en la revista con menor impacto). Para terminar, hay que señalar que la distribución de citas dentro de una misma revista sigue el patrón de la ley de Lotka, por lo que hay que esperar que los artículos citados más que el promedio sean significativamente menos de la mitad. En resumen, el indicador RCR es poco significativo si se ignora el valor de sus componentes y casi innecesario si se conocen.

En segundo lugar, la duración del periodo que se toma como unidad puede afectar el valor total de aquellos agregados que hayan evolucionado durante los años de ese periodo, ya que en el modo de cálculo de las citas recibidas y promedio tienen un mayor peso los primeros años del periodo. Aquellos países, por ejemplo, que hayan mejorado serán infravalorados, mientras que los que hayan empeorado serán sobrevalorados. La muestra de citas a los trabajos del primer año será representativa (abarca las citas realizadas a lo largo de cuatro o cinco años); pero para los trabajos del último año sólo se contabilizan las citas emitidas en ese mismo año, lo cual significa, casi, ignorarlos.

Por último, el método de asignar las citas, en el caso de colaboración entre varios países, al país donde radica la institución del primer autor firmante puede producir también sesgos apreciables si la proporción de documentos que encabeza ese país cambia significativamente y existen diferencias entre los trabajos con y sin colaboración internacional. Por poner un ejemplo: en la producción española en el SCI el porcentaje de documentos en colaboración internacional ha pasado de 12% en 1981 a 25% en 1991. Como la proporción de esos documentos encabezados por instituciones españolas se ha mantenido prácticamente estable (del 59% al 60%), la consecuencia es que, de todos los documentos con participación española, en el 94% de la producción de 1981 la primera institución era española, con lo que sus citas serían asignadas a España; en 1985 esta proporción era del 92%, y en 1991, del 85%. La calidad media de ese 25% con colaboración internacional es claramente superior a la del resto de la producción española (Maltrás y Quintanilla, 1995) y el 15% que no cuenta para España pertenece a ese subconjunto selecto. Esto significa que una proporción creciente de los mejores documentos españoles habría sido ignorada durante el periodo 1985-91.

Estos factores no invalidan la utilidad de estos estudios, pero sí deben alertar del peligro de una interpretación precipitada y excesiva. De hecho, en los estudios citados se ofrecen también los valores observados y los esperados, además del RCR.

Vinkler (1986), tras señalar la posible influencia en los valores del RCR de las elecciones de los investigadores en cuanto a sus canales de publicación, propone otro indicador de tipo observado / esperado, la *citación relativa a la disciplina* (*relative subfield citedness*,  $R_w$ ), que compara las citas recibidas con la media esperada para la disciplina, que es la media de los factores de impacto ( $\bar{h}$ ) de todas las revistas que se considera que constituyen esa disciplina. En este caso ignora las autocitas, que estima aproximadamente en el 15% de las citas recibidas, las citas esperadas para un conjunto de  $n$  documentos es:  $I_r = 0,85 \cdot n \cdot \bar{h}$ . Las citas efectivas a esos documentos ( $I$ ) son las obtenidas en los dos años siguientes, siguiendo el patrón del factor de impacto habitual:

$$R_w = \frac{I}{I_r} = \frac{I}{0,85 \cdot n \cdot \bar{h}}$$

Por otra parte, al comparar el impacto esperado de un conjunto de  $n$  documentos en un área con el impacto medio de las revistas de esa área,  $\bar{h}$ , define el indicador *estrategia de publicación* ( $S$ ):

$$S = \frac{\frac{\sum h_i}{n}}{\bar{h}}$$

Este indicador puede ser interpretado, según Vinkler, como un signo de si los investigadores eligen los canales de publicación adecuados: un valor superior a 1 indica que los investigadores publican en revistas con impacto superior a la media de la disciplina. Claro que no sólo depende del deseo del autor el que su trabajo aparezca en las revistas del máximo impacto del área.

En el grupo EPOC, de la Universidad de Salamanca, para un informe encargado por el Consejo de Universidades como parte del Programa Piloto de Evaluación de las Universidades, (Maltrás y Vidal, 1993) empleamos un indicador análogo a éste último, con la diferencia de que el factor de impacto medio de cada área (según una clasificación derivada de la del SCI) se calculó teniendo en cuenta el número de documentos publicados por cada revista, es decir, ponderado por el tamaño de la revista. Esta corrección se consideró necesaria al observar los sesgos que las diferentes distribuciones de tamaños de revistas entre las áreas podrían ocasionar.

Este indicador puede ofrecer una estimación de trazo grueso sobre la que apoyar comparaciones entre instituciones y áreas. Sin embargo, algunas de esas comparaciones pueden ser poco fiables debido a que los extremos máximos y mínimos de las áreas son muy variados (en un área, 1,6 puede ser el máximo teórico posible, y en otra 1,8 ser superado por varias instituciones). En general, pueden realizarse comparaciones cautelosas dentro de la misma área, y entre áreas atendiendo sólo a si está por encima o por debajo del valor medio. El recorrido, la distribución y la densidad de los valores del factor de impacto pueden mostrar diferencias que oscurezcan la interpretación de estos indicadores que los toman directamente.

El siguiente esquema puede valer como resumen de lo dedicado al factor de impacto y a sus variantes. Las posibilidades de cálculo resultan de la combinación entre años citantes ( $\leftarrow$ ) y años citados ( $\circ$ ).

patrón	nombre	referencia
OO<	factor de impacto	Garfield, (1972)
⊗	índice de inmediatez	Journal Citation Reports
O_<	puntuación de citación de la revista (JCS)	Moed et al. (1984)
...O<	factor de impacto generalizado	Rousseau (1988)
...O_<	factor de impacto truncado	Rousseau (1988)
O<<	tasa media de citación (a)	Braun, Glänzel, Schubert (1987)
⊗⊗⊗⊗	tasa media de citación (b)	Schubert, Glänzel, Braun (1989)

Quizá el patrón más razonable de los propuestos hasta ahora, conocida la distribución habitual de las citas, sea la *tasa media de citación (a)*, (O<<), puesto que utiliza la muestra a corto plazo más sólida de las citas dirigidas a un sólo año, sin que existan solapamientos entre los periodos sucesivos (en el factor de impacto, OO<, otra parte de las citas dirigidas al segundo año entra en el cálculo del siguiente periodo). El factor de impacto tiene la ventaja de la simplicidad técnica de su cálculo; está por ver que otras alternativas obtengan resultados significativamente diferentes, sobre todo en lo concerniente al orden establecido. Un trabajo pendiente es el de comparar las listas de revistas obtenidas de la ordenación mediante éstos y otros patrones que podrían resultar plausibles: (⊗<<), una versión reforzada del *tasa media de citación (a)*, o (⊗⊗<), parecido a una fusión del factor de impacto y el índice de inmediatez, entre otros que se pueden imaginar. De todos modos, como recomendación general hay que decir que todos éstos *son indicadores y no datos*. Los valores numéricos de estas medidas deben ser manejados con mucha cautela, teniendo siempre presente la influencia que en ellos tienen las distintas características de la distribución de citas en cada área, y no como si fueran datos absolutos que reflejan distancias reales. Sin embargo, este tipo de indicadores sí puede ser muy útil en la ordenación significativa de las listas de revistas de cada área, como ya se ha comprobado (McAllister, Anderson, Narin, 1979). En el siguiente apartado se describirá el modo en que puede aprovecharse la información del orden establecido por estos indicadores sin emplear su métrica.

#### 6.4. Los indicadores P-10 (puntuación decílica) y %SUP (peso del decil superior) como estimador de la calidad

Este apartado se dedicará a examinar los fundamentos, a comentar algunos antecedentes y a describir la metodología de la construcción de dos indicadores referidos a la calidad de agregados de documentos: la *puntuación decílica* (P-10) y el *peso del decil superior* (%SUP), propuestos y aplicados en los estudios bibliométricos del grupo EPOC de la Universidad de Salamanca (Maltrás, Quintanilla, 1995). Estos dos indicadores hacen uso del factor de impacto de las revistas científicas publicado anualmente en el *Journal Citation Reports* por el ISI, (*Institute for Scientific Information*, de Filadelfia, EE.UU.) y de la clasificación temática que esta institución aplica a las revistas incluidas en el SCI. Sin embargo, en lugar de emplear directamente los valores del FI, se lleva a cabo un proceso de normalización de éstos, a partir de la ordenación derivada del FI en cada disciplina, que hace posible una interpretación homogénea entre áreas y periodos. La noción clave de este proceso es la de *distribución en intervalos decílicos de la producción científica*, que tiene como resultado una clasificación de los documentos dependiente de la revista en la que fueron publicados.

Este planteamiento se apoya en dos pilares. El primero es la respuesta que se da ante las dificultades de interpretación de los valores numéricos de los indicadores basados directamente en el factor de impacto; no es la negación de la utilidad del FI, sino la de usar la valiosa información que puede proporcionar en relación con la jerarquía que establece entre las revistas científicas de un área, definiendo escalas a partir de esa información. El segundo pilar es considerar que la revista en la que se publican los documentos puede ser un hecho significativo y apto para realizar estimaciones sobre conjuntos de ellos.

Ambos están relacionados con algunos precedentes en la literatura de los estudios cuantitativos de la ciencia que es oportuno recordar aquí. El uso de escalas basadas en el orden, como las habituales percentiles, fue sugerido por Garfield (1978) como una posible solución para la evaluación de científicos de distintas áreas sobre la base de la producción o número de citas recibidas en relación con su campo. Narin (1987) también indica la utilidad de medir la presencia en el decil superior de las citas recibidas, una técnica también empleada en los análisis de los documentos de las patentes. Van Raan y Hartmann (1986)

realizaron un estudio exploratorio sobre la distribución de los documentos de la revista *Science* en deciles según las citas recibidas. Los trabajos en los que el nivel de la revista de publicación se toma como indicador del nivel de sus documentos son numerosos, entre ellos varios de los citados en el apartado anterior. Pero hay que destacar el trabajo de Crouch, Irvine y Martin (1985), de SPRU (*Science Policy Research Unit*, de la Universidad de Sussex), que emplearon la puntuación otorgada por expertos a las revistas de dos especialidades para evaluar de ese modo indirecto la producción británica.

Una aportación reseñable es la debida a Cano, Julián (1991), de la Secretaría del Plan Nacional de I+D (Comisión Interministerial de Ciencia y Tecnología) quienes emplearon el orden de las revistas fijado por el FI para analizar cómo había evolucionado la distribución de la producción española en los diferentes puestos. Definieron una normalización por áreas para el factor de impacto de cada revista<sup>55</sup> que luego emplean para el cálculo de indicadores normalizados de calidad sobre la producción española.

La estimación indirecta de la calidad de los documentos a través del nivel de sus fuentes difusoras suele encontrar la objeción de que, incluso en las revistas más prestigiosas pueden aparecer trabajos mediocres (Martin, Irvine, 1980; Lawani, 1977) o que no se puede excluir que en revistas poco apreciadas haya perlas escondidas. Asimismo, es bien conocida la asimétrica distribución de citas entre los trabajos de una misma revista (un claro ejemplo lo muestran van Raan, Hartmann, 1986), lo cual puede sugerir diferencias proporcionales en cuanto a la calidad. En resumen, que no se puede juzgar por igual a todos los trabajos sólo por el hecho de haber sido publicados en la misma revista.

Sólo son necesarias unas breves apreciaciones a este respecto. En primer lugar, la estimación indirecta se refiere a conjuntos estadísticamente significativos, no a documentos aislados. De nuevo hay que recordar que la publicación de un documento en una revista no es circunstancial respecto al FI de esa revista, sino parte de la causa del correspondiente valor del FI. Al juzgar indirectamente agregados de documentos lo que se asume es que la probabilidad de que hayan superado un control de

<sup>55</sup>  $Z_j = (H_j - X_{\min}) / (X_{\max} - X_{\min})$ , donde  $H_j$  es el FI de la revista,  $X_{\max}$  es el valor del máximo FI en el área y  $X_{\min}$  el del mínimo. El valor normalizado  $Z_j$  está, para todas las áreas, entre 100 y 0. Este indicador no considera el número de trabajos que contiene cada revista y mantiene la proporcionalidad de las distancias expresadas por el FI.

calidad similar es muy alta. Es ese control el que fija el nivel medio de los documentos aparecidos en la revista. Es quien esté dispuesto a apostar por un agregado de documentos publicados en las revista de menor FI frente a otro publicado en las de mayor FI quien deberá explicar su preferencia y cómo es posible que los primeros no hayan bajado el FI de sus revistas ni los segundos lo hayan subido. Lo que se busca es aumentar las posibilidades de acertar en la detección de los documentos superiores. Por otro lado, existe un concepto de calidad legítimamente aplicable para este caso, que se refiere más a la *expectativa de éxito* y a la *confianza despertada* que al hecho de haber recibido después el interés correspondiente.

Los indicadores P-10 y %SUP se calculan a partir de la puntuación otorgada a cada documento sobre la base de la posición que ocupa en la producción total del área científica a la que pertenece. El área científica está formada por los documentos publicados en las revistas asignadas por el SCI a esa área de su clasificación. Para determinar cuál es la posición, se distribuye en deciles, es decir, en 10 partes iguales, la producción científica mundial recogida en el SCI correspondiente a un área, tomando como criterio básico el orden establecido por el FI entre las revistas que constituyen esa área. Cada intervalo decílico contiene, entonces, un 10% de los documentos del área, en orden descendente de FI: en el primer decil están los documentos publicados por las revistas de mayor impacto, en el segundo, los de las de mayor FI de las restantes, y así sucesivamente hasta el último decil. La posición de un documento en un área será más alta cuanto mayor sea el FI de la revista que lo publicó y cuantos menos documentos hayan aparecido en revistas de esa área con mayor impacto que la suya.

A los documentos del primer decil se les da la puntuación 10; a los del segundo, 9; a los del tercero, 8, y así hasta el décimo, que reciben 1 como puntuación<sup>56</sup>. Mediante este procedimiento, para cualquier área o periodo se sabe que, por ejemplo, dado un documento con puntuación

---

<sup>56</sup> Si la producción de una revista (o de varias con el mismo FI) se encuentra entre dos deciles sucesivos (es cortada por el límite entre dos deciles), entonces se puntúa a todos esos documentos según la proporción de ellos que correspondería a cada decil. Por ejemplo, si las revistas  $R_i$  y  $R_j$  completan con el 25 % de sus documentos el primer decil y su 75% restante se incluye en el segundo, la puntuación para todos sus documentos será:  $(10 \cdot 0,25) + (9 \cdot 0,75) = 9,25$ .

decílica 7, el 70% de los documentos de su área tienen una puntuación menor o igual en ese periodo, y el 30%, una puntuación superior.

El indicador *puntuación decílica* (P-10) de un agregado de  $n$  documentos es la media de las puntuaciones de cada uno de ellos,  $p_i$ :

$$P10 = \frac{\sum_{i=1}^n p_i}{n}$$

El valor máximo es 10 (e.d. todos los documentos están en el primer decil), el mínimo es 1 (todos los documentos están en el último decil) y el valor esperado es 5,5, que es la media para cualquier área completa en cualquier periodo, o para cualquier muestra aleatoria válida, ya que la probabilidad de haber publicado en cualquier decil es *a priori* siempre de 1/10.

El indicador *peso del decil superior* (%SUP) expresa, en porcentaje, la proporción que representan en un agregado (una institución, o un área de un país) los documentos que tienen puntuación 10, es decir, están clasificados en el primer decil. El valor esperado para una muestra aleatoria es 10%; un valor superior indica una presencia superior a la esperada en el subconjunto de documentos con mayor puntuación según el procedimiento descrito anteriormente. Puede interpretarse como un indicador de *excelencia* científica y sirve como complemento del P-10, ya que informa sobre uno de los extremos de la distribución cuya media expresa el P-10.

El cálculo de estos indicadores requiere la disponibilidad de una enorme cantidad de datos relacionados con las revistas cubiertas por el SCI: se necesita conocer el conjunto completo de revistas cubiertas por el SCI, la clasificación temática de cada una en cada periodo, el número de documentos que han publicado cada año, las citas recibidas por esos documentos al segundo año de su publicación. El tratamiento informático de esos datos para aplicar el procedimiento de distribución en intervalos decílicos y para calcular los valores de estos indicadores tampoco es sencillo.

Pero ambos presentan ventajas que compensan estos esfuerzos. Las principales son: *estabilidad y univocidad* de la interpretación a lo largo de periodos y áreas científicas; *linealidad* de la escala (un aumento de una décima en el P-10 significa que, por término medio, uno de cada

diez documentos ha ascendido un decil); y sobre todo, la *aditividad* de sus valores, lo que permite (con una ponderación adecuada) obtener valores para agregados geográficos o institucionales de los que forman parte documentos de varias áreas.

Estos indicadores han sido aplicados en el análisis bibliométrico de toda la producción científica española recogida en el SCI para los años 1986-91 (Maltrás y Quintanilla, 1995) con resultados satisfactorios, que han permitido detectar o confirmar interesantes aspectos del sistema español de ciencia y tecnología. Los distintos cruces geográficos, institucionales y temáticos han puesto de manifiesto algunos rasgos de su evolución reciente. Estos indicadores han mostrado diferencias claras y evoluciones consistentes en cuanto a la calidad, de entre las que cabría destacar la superioridad de la producción realizada en colaboración con instituciones extranjeras o el buen nivel de la investigación realizada en los hospitales españoles.

El mismo procedimiento de distribución en intervalos decílicos de la producción podría aplicarse, en lugar de a la calidad estimada mediante el factor de impacto de la revista, al número de citas efectivamente recibido por cada documento y otorgar una puntuación con los mismos criterios que en el caso expuesto. El mismo procedimiento puede aplicarse, claro está, sobre la ordenación obtenida a partir de alguna otra variante del factor de impacto o de otros métodos objetivos.

## Capítulo 7

### *Los indicadores bibliométricos de colaboración*

---

El análisis bibliométrico permite ir más allá del estudio de los productos de la ciencia: gracias a los datos que se incluyen habitualmente tanto en la propia literatura científica (que es la fuente primaria) como en las fuentes secundarias que se constituyen con las referencias bibliográficas de esa literatura, pueden rastrearse algunos aspectos acerca del modo de operar de los *productores* de esa literatura, lo que en la práctica equivale a decir de los productores del conocimiento científico. Llamaremos indicadores bibliométricos *de colaboración* a aquéllos que informan acerca de las relaciones que han existido entre los productores o agentes científicos en el proceso que ha concluido con la publicación conjunta de resultados científicos. La base de estos indicadores son los datos sobre la autoría de las publicaciones científicas y sobre las instituciones de trabajo de esos autores que aparecen en las referencias bibliográficas; el subconjunto relevante de estas referencias es el de las publicaciones que tienen más de un autor.

Los productores elementales de ciencia son los autores científicos, los investigadores particulares. El estudio de las relaciones de colaboración entre los individuos puede iluminar puntos capitales para la comprensión del desarrollo de problemas o de líneas de investigación, especialmente si se combina con otros tipos de análisis como el de redes de citas o de léxico compartido. Sin embargo, la actividad de estos individuos aparece normalmente guiada por su pertenencia o su inclusión en agregados como equipos de investigación o instituciones con actividad científica, sean éstas académicas o de otros tipos. La información

que puede proveer la indagación estadística de las coautorías puede ser muy valiosa en la detección de grupos de investigación particularmente productivos o en la localización y estudio de las relaciones entre instituciones.

Pero las posibilidades de análisis de la colaboración son mucho más variadas. En este capítulo se tratarán, en primer lugar y de un modo general, los principales aspectos que hacen de la colaboración un objeto de estudio interesante para la comprensión de la ciencia. En segundo lugar, se abordará de modo detallado el problema técnico de la medida de la colaboración, relacionándolo con el problema conceptual de definir qué sea una colaboración.

### *7.1. La colaboración en la ciencia*

En un sentido general, la empresa científica siempre se ha basado en la colaboración: su avance ha sido posible gracias a la disposición de los individuos que en ella tomaban parte a poner en común sus propios logros. De ese modo, su trabajo servía para ahorrárselo a otros, quienes podían así tomar las metas ganadas por sus colegas como puntos de partida. Pero más allá de esta consideración genérica, la colaboración en la ciencia ha cobrado una importancia creciente en otro sentido más concreto, aquél que describe el hecho de que la producción de conocimientos científicos requiera muchas veces el concurso directo de varios individuos, quienes pueden incluso pertenecer a distintas instituciones o países.

Para los estudios cuantitativos de la ciencia, el análisis de la colaboración representa, por una parte, una cuestión metodológica, como ha podido verse en los problemas de atribución de resultados que motivaba, y como se verá enseguida al tratar las posibilidades de medida. Pero, por otra parte, suscita al mismo tiempo un interés más teórico. El aumento de la colaboración es uno de los fenómenos más visibles de entre los que han conformado la transformación que la ciencia moderna ha experimentado a lo largo de los tres últimos siglos<sup>57</sup>. Afecta de manera directa al proceso de generación de conocimientos científicos, tanto en el nivel de los investigadores individuales como en el de las

---

<sup>57</sup> Price (1963) ya destacó el crecimiento cuantitativo de la colaboración desde el siglo XVII. Sobre el tema insisten Beaver y Rosen (1978a).

instituciones y los recursos materiales puestos en juego. Y genera redes con efectos cohesivos entre los elementos que constituyen la comunidad científica. Por ello, el estudio de sus conexiones con otros cambios en la ciencia que a éste han acompañado resulta del máximo interés. Determinar cuáles han sido sus causas y cuáles sus efectos ha de ayudar a mejorar la comprensión de los mecanismos que gobiernan el desarrollo de la ciencia.

Beaver y Rosen (1978a) sistematizan del siguiente modo los principales motivos que se han señalado como justificaciones para la colaboración. La colaboración puede ser un medio para *lograr acceso* a equipos especiales, a determinadas habilidades, a materiales raros o incluso a la visibilidad y al reconocimiento; también puede buscarse para *aumentar la eficiencia* en el uso del tiempo o del trabajo; para adquirir experiencia; para adiestrar investigadores; para patrocinar discípulos; para incrementar la productividad; con la intención de multiplicar la capacitación y así conseguir fuentes de financiación, visibilidad o reconocimiento; para superar el aislamiento intelectual; puede estar guiada por la necesidad de una confirmación adicional o la evaluación de un problema; o también por la necesidad de estímulo o de fertilización cruzada; incluso, puede surgir por mero accidente.

Beaver y Rosen muestran que, a pesar de que las explicaciones más socorridas para dar cuenta del aumento de la colaboración son la especialización o el aumento de la eficacia, lo más plausible es que la causa más directa fuera el proceso de *profesionalización* de los científicos. La colaboración fue, en un principio, una consecuencia de la estructuración que la comunidad científica fue alcanzando al tiempo que se consolidaba como sostenedora legítima y exclusiva de un tipo especial de conocimiento. Tanto las relaciones *externas* de la ciencia con el resto de la sociedad como las *internas* entre los científicos definieron una jerarquía entre éstos, y la colaboración sirvió como importante factor de movilidad gracias al reconocimiento que podía proporcionar.

Pero, al margen de cuál fuera el mecanismo concreto que originó la colaboración en la ciencia moderna, lo cierto es que este fenómeno no se percibe como algo circunstancial en la ciencia del S XX, especialmente en lo que se llama *Gran Ciencia*. En ésta, la colaboración es la norma; por el peso que ha llegado a alcanzar, se ha tomado a veces como uno de sus rasgos característicos. Implica a los individuos y a las instituciones. En las últimas décadas, se ha convertido en uno de los

preocupaciones centrales de los estados al diseñar sus políticas de la ciencia.

La búsqueda de reconocimiento en un medio profesionalizado puede ser también una buena explicación de la colaboración científica en nuestros días. A medida que aumenta el tamaño de la ciencia y se produce una mayor concentración relativa de los recursos, la competencia entre los científicos se hace más aguda. Al mismo tiempo, la creciente complejidad de la ciencia exige una formación altamente especializada (tanto temática como funcional) para poder competir en la obtención de resultados. Esto hace que la colaboración pueda verse como la respuesta a la necesidad de complementar capacidades, incluso como algo implícito en la formación de los investigadores: el objetivo no es llegar a ser un agente autónomo, sino alguien capaz de formar parte de un equipo científico. Así, la colaboración en la ciencia se habría institucionalizado, y éste parece ser el caso en muchas disciplinas científicas.

Pero aunque la colaboración no fuera algo necesario, puede valorarse su efecto multiplicador del reconocimiento, no sólo por la posibilidad que ofrece de figurar al lado de investigadores reconocidos. La colaboración científica parece mejorar la capacidad individual, es decir, el resultado obtenido suele ser superior a la mera suma de los resultados individuales o al mejor de éstos. Cuando esto tiene lugar, se dice que ha habido *sinergia*. Esto plantea la cuestión de si, *a priori*, deberían valorarse por encima a los resultados en colaboración. Por otro lado, abre la discusión acerca de las relaciones entre la colaboración y la competitividad, comportamientos aparentemente antagónicos, pero que pueden verse como mutuamente reforzadores. En un entorno competitivo, si la colaboración dota de una ventaja respecto a quienes no colaboran, aquellos que colaboren se verán reforzados a seguir colaborando y además serán seleccionados entre los que no colaboran.

Varios aspectos interesan desde una perspectiva más cercana a la política de la ciencia. La colaboración científica se percibe a veces como un medio útil para la integración en la vanguardia de investigadores o grupos que se hallen en periodo de formación o que se encuentren retrasados respecto a los métodos y problemas vigentes en el frente de investigación. El trabajar codo con codo, el contacto directo con el modo de llevar a cabo las tareas concretas y la planificación cotidiana, es muchas veces la única forma de aprendizaje en no pocos ámbitos de la ciencia, y casi siempre la más efectiva. Cuando se trata de técnicas

experimentales, esta intermediación puede ser necesaria para adquirir la destreza práctica requerida. La transferencia del saber-cómo (*know how*) encuentra un terreno favorable en la colaboración, la cual sustenta en general un tipo de transferencia no completamente asimétrica.

Entre los gestores de la ciencia no es rara la preocupación por evitar duplicaciones de esfuerzos, es decir, por impedir que se empleen recursos simultáneamente en varios proyectos o líneas de investigación que están en busca del mismo objetivo. En muchas ocasiones esa duplicación es inevitable e incluso positiva<sup>58</sup>. Pero esta preocupación está más justificada, desde luego, cuando los costos que supone una investigación son muy grandes. En estos casos, la colaboración posibilita que investigadores de distintos equipos, instituciones o países trabajen juntos en el mismo proyecto, que a veces no podría haber sido acometido separadamente por cada uno de los colaboradores debido a la magnitud de su coste.

Otras veces, la colaboración científica se establece con la finalidad de optimizar el uso de unas infraestructuras o instalaciones ya disponibles, pero a las que puede sacarse mayor partido si se emplean bajo algún tipo de colaboración. La necesidad de compartir equipos raros o muy específicos está entonces en el origen de esa colaboración. Por otra parte, cuando las instalaciones científicas alcanzan un cierto grado de complejidad, la colaboración entre investigadores es el único modo de que éstas funcionen, porque exige la intervención de varios especialistas en distintos campos para mantenerlas en funcionamiento.

Varios son, pues, los motivos que convierten la colaboración en objeto de interés para los responsables de política científica. La transferencia y diseminación de los resultados y del saber práctico están íntimamente relacionados con la corrección de desequilibrios entre las partes de un sistema y con el aumento de la cohesión entre esas partes. Ésta es una preocupación central en el proceso de construcción europea, por ejemplo. El reforzamiento de las relaciones de colaboración entre los países miembros se ha visto como uno de los principales instrumentos de cohesión y de convergencia, rasgos que favorecerán la constitución de un sistema transnacional.

---

<sup>58</sup> Los descubrimientos múltiples no son estrictamente equivalentes; pueden variar en pequeños detalles que proporcionan alternativas enriquecedoras sobre temas o soluciones semejantes (Price, 1963).

Las relaciones de colaboración son así una puerta hacia el análisis de las características de los sistemas científicos, tanto nacionales (Quintanilla, Maltrás, 1993) como supranacionales (Maltrás, Quintanilla, Vega, 1995). El análisis cuantitativo de la colaboración permite acceder a distintos aspectos de la organización de la ciencia.

## *7.2. La medida de la colaboración*

El análisis cuantitativo de la colaboración parte de los datos sobre la autoría y la afiliación institucional de los autores de las publicaciones científicas. Se ocupa de estudiar las características que pueden definir los patrones del comportamiento productivo de los agentes científicos en cuanto a las relaciones que mantienen entre ellos, sobre la base de los productos que han surgido de un esfuerzo cooperativo.

Los elementos básicos entre los que se establece la colaboración científica son los investigadores particulares; el estudio de las relaciones en este nivel básico despertó mucho interés desde el inicio de los estudios cuantitativos, especialmente porque permitía una aproximación empírica acerca de la estructura social de los científicos. Sin embargo, también existen otros niveles superiores, tales como las unidades de investigación, los departamentos, las instituciones, o como los sucesivos niveles geo-políticos, de los que se puede inferir que mantienen entre sí relaciones de colaboración científica a partir de las relaciones entre los individuos que pertenecen a ellas. De hecho, la colaboración entre los individuos puede muchas veces depender de la intención o del impulso generado en el nivel institucional o político, en forma tanto de incentivos genéricos como de proyectos específicos. La existencia de distintos niveles implica, además de una variedad en las posibilidades de interacción entre ellos, que las medidas cuantitativas puedan realizarse en cualquiera de esos niveles; por otra parte, debe decidirse si se toman o no en consideración las relaciones establecidas en los niveles inferiores.

Al lado de los distintos niveles que pueden tomarse en consideración, otras dos razones convierten la medida de la colaboración en algo complejo: por un lado, existe una multitud de aspectos medibles que pueden servir para construir una imagen de la colaboración; por otro lado, la noción de colaboración puede concebirse de distintos modos, y esto produce distintas alternativas metodológicas para la medida. La

riqueza de aspectos cuantificables emergerá casi por sí sola; en cuanto a las alternativas metodológicas y las concepciones de la colaboración, nos ocuparemos de ellas enseguida con cierto detalle.

Antes conviene señalar algunas cuestiones previas acerca de los indicadores bibliométricos de colaboración. En primer lugar, hay que recordar que *miden solamente la colaboración que ha tenido éxito*, es decir, la que ha producido resultados publicados<sup>59</sup>. Es un hecho obvio, pero que no hay que perder de vista en la interpretación de estos indicadores: la colaboración que no ha publicado queda fuera de su alcance, y por lo tanto la imagen que ofrece es la de la colaboración exitosa.

Otra cuestión de principio se refiere a la representatividad del hecho de la coautoría respecto al de la colaboración científica. Desde los primeros trabajos en colaboración, la firma conjunta de un trabajo por varios autores significaba que el resultado que se presentaba era atribuible a esos autores como grupo y no por separado, que era resultado de una cooperación entre ellos. Esa cooperación puede establecerse de diferentes maneras, aunque la coautoría no es capaz de discriminar entre ellas, reflejándolas a todas por igual; tampoco se puede descartar que se den otros tipos de cooperación que no terminen en coautorías. Estas dudas están también en relación con las costumbres de firma en cuanto a la atribución de los resultados de la investigación, que ya hemos tratado anteriormente en el capítulo correspondiente, dentro de los indicadores de producción. Lo que estaría en juego son las razones que determinan si alguien que ha trabajado en una investigación aparece finalmente como autor de las publicaciones a las que dé origen, y si estas razones pueden ser sorteadas mediante algún tipo de fraude. Asumimos que, para un conjunto suficiente de trabajos, las coautorías ofrecen muy pocas dudas sobre lo que reflejan: los trabajos firmados por varios investigadores muestran con mucha fidelidad una colaboración para esos trabajos; en otras palabras, la coautoría es un fenómeno indicativo de la colaboración. Es más, aquí reservaremos el término "colaboración" precisamente para definir el tipo de cooperación que puede terminar en una coautoría y que está detrás de las coautorías (Beaver y Rosen, 1979a).

Nos ocuparemos aquí de los aspectos metodológicos relevantes para el análisis de la colaboración en sistemas científicos con diferentes

---

<sup>59</sup> Sin embargo, cuando se analizan los datos de proyectos de investigación suele suceder lo contrario, esto es, que no se sabe cuáles de esas colaboraciones han sido más productivas.

niveles de agregación. El estudio de la colaboración entre individuos, interesante y rico en sí mismo, es no obstante de una complejidad técnica menor y cae dentro de las consideraciones que se refieren a los sistemas científicos. Esto no niega que el análisis de la colaboración entre individuos tiene algunas aplicaciones específicas, como la detección y estudio de los equipos de investigación, con sus propias peculiaridades.

Para empezar, será útil distinguir dos tipos de aproximación al estudio cuantitativo de la colaboración: llamaremos *simple* a uno y *relacional* al otro. El primero construye indicadores que nos ofrecen información sobre el nivel de colaboración que exhibe la producción científica de un agregado cualquiera de agentes<sup>60</sup>, o sobre otras características de la colaboración, sin discriminar con qué otros agregados tiene lugar esa colaboración, es decir, estudiando en conjunto la colaboración que ha llevado a cabo. Se obtendrán indicadores de este tipo al calcular, por ejemplo, el porcentaje de documentos en colaboración, el número medio de autores, instituciones, tipos institucionales que participan en la producción de ese agregado, o la distribución de su producción en rangos definidos por el número de autores, instituciones, tipos institucionales o países.

Los indicadores de colaboración *relacionales*, en cambio, permiten estudiar la colaboración entre los agregados de un determinado sistema desde la perspectiva que se centra en las redes que se establecen en ese sistema, es decir, identificando separadamente los agregados concretos que intervienen en cada colaboración y construyendo una imagen global de las relaciones entre los agregados a partir de la *suma* de las colaboraciones concretas que han tenido lugar entre ellos.

El cálculo de indicadores relacionales de la colaboración exige un proceso más complicado para el tratamiento de los datos, pero sus posibilidades son mucho mayores que las que ofrece la aproximación que hemos llamado *simple*; además, los indicadores simples pueden definirse como subconjunto de los que se pueden construir a partir del tratamiento relacional de los datos. Por contra, su análisis e interpretación son más exigentes, mientras que la información que proporcionan los indicadores *simples* puede ser muchas veces suficiente.

---

<sup>60</sup> Un agregado está constituido por uno o más agentes; por eso, en los párrafos siguientes el término "agregado" indicará indistintamente "un agente" o "un agregado de agentes".

Por lo tanto, nos centraremos en el estudio de la metodología relativa a la construcción de indicadores relacionales entre agregados de agentes con posibles niveles inferiores, ya que ésta incluye a las otras posibilidades.

La primera fase que se requiere para estudiar relacionamente la colaboración es obtener una especie de resumen cuantitativo de las colaboraciones que han tenido lugar en un conjunto de trabajos científicos. Existen varios modos de definir esa suma, y cada uno de ellos dibujará un retrato distinto de la red establecida en el sistema. Esta es precisamente la cuestión metodológica que aquí trataremos con especial atención: *las posibilidades disponibles para obtener una cuantificación de la colaboración establecida en un sistema y su dependencia respecto a los distintos modos de concebir qué sea una colaboración.*

En segundo lugar, a partir de esa suma "bruta" puede estudiarse luego la importancia relativa de cada conexión que se establece entre cada par de agregados, teniendo en cuenta el total de las relaciones establecidas por cada agregado y en el sistema en conjunto. Diversos índices estadísticos se han propuesto<sup>61</sup> para identificar las conexiones significativas mediante la detección de desviaciones respecto a lo esperado; estos índices suelen además neutralizar las diferencias de tamaño que pueden existir entre los agregados. También han tenido relación con esta fase de análisis estadístico propuestas para transformar las matrices y hacerlas adecuadas a un determinado tipo de análisis (Price, 1980a; Price, 1980b).

Por último, pueden lograrse representaciones gráficas basadas en esos índices normalizados mediante la aplicación de diversas técnicas estadísticas avanzadas<sup>62</sup> o directamente en la suma obtenida primariamente. Las primeras acentuarán los rasgos más significativos, mientras que las últimas se acercarán más a una representación "realista" de las redes de colaboración.

<sup>61</sup> Destacan los siguientes índices que han sido profusamente utilizados en las medidas relacionales: 1) *Índice de Jaccard*, que mide el grado de solapamiento entre las colaboraciones de dos elementos ( $J_{ij} = C_{ij} / (C_i + C_j - C_{ij})$ ); 2) *Índice de Salton*, que compara la frecuencia observada de la colaboración en relación a la frecuencia esperada ( $S_{ij} = C_{ij} / (C_i C_j)^{1/2}$ ); 3) *Índice de Proximidad*, que es una nueva forma de expresar la ratio observado / esperado ( $P_{ij} = C_{ij} * T / (C_i C_j)$ ).

<sup>62</sup> Técnicas que se han utilizado frecuentemente son el análisis de proximidades, análisis de racimos, escalamiento multidimensional o *minimum spanning tree*.

Hay que separar, por lo tanto, tres momentos metodológicos en el análisis cuantitativo de la colaboración:

- (a) suma o recuento, proceso por el que se obtiene una matriz bidimensional cuyas celdas cuantifican la colaboración que se ha producido entre cada par de agregados. Esta matriz es una reducción de la multidimensionalidad de la colaboración con el fin de poder aplicar técnicas estadísticas manejables en la siguiente fase;
- (b) normalización mediante índices, para aislar la influencia de las diferencias de tamaños entre agregados y resaltar las conexiones estadísticamente significativas, interpretables en términos de preferencias o tendencias de los agregados bajo estudio;
- (c) técnicas para la representación gráfica para expresar la configuración resultante de un modo que facilitara su comprensión e interpretación.

Durante las últimas décadas, el análisis de la colaboración científica, especialmente de la internacional, ha despertado un interés creciente<sup>63</sup>. Sin embargo, la atención se había centrado en la segunda y tercera fase, con la búsqueda de índices normalizadores y la aplicación de técnicas de representación recientemente desarrolladas. En línea con nuestro trabajo anterior (Maltrás, Vega, Quintanilla, 1995), nos ocuparemos aquí de la primera fase, de cuyo resultado dependen las otras dos, y que había quedado sin sistematizar, ignorándose el abanico de opciones que permite y por ende también los aspectos que quedaban ocultos al aplicar el método de recuento empleado generalmente.

### *7.2.1. Bases conceptuales y alternativas metodológicas en la cuantificación de la colaboración*

Una vez que hemos adoptado una clase de hechos como sucesos en los que se produce colaboración en la ciencia (tales como los trabajos publicados o los proyectos de investigación) y que podemos identificar a los agentes que participan en esos eventos, parece que el camino hacia la medida de la colaboración está completamente despejado: basta con sumar la colaboración que se ha producido en cada uno de esos eventos

---

<sup>63</sup> Frame y Carpenter (1979); Okubo, Miquel, Frigoletto, Doré (1992); Luukkonen, Persson, Sivertsen (1992); Luukkonen, Tijssen, Persson, Sivertsen (1993); Tijssen y Moed (1989); Schubert, Braun (1990).

de colaboración para disponer de datos brutos sobre los que realizar el análisis.

Sin embargo, cuando queremos cuantificar la colaboración que se produce entre los agentes científicos que aparecen como responsables conjuntos de un cierto trabajo, es necesario haber precisado claramente *qué entendemos por colaboración* y las circunstancias de las que depende la magnitud numérica que asignemos a cada caso; y, antes aún, cuál sea la estructura organizativa que podemos suponer que ha sostenido esa colaboración, ya que existen varias maneras en que ésta puede llevarse a cabo.

Vayamos en primer término, pues, con la exploración de la importancia que pueden tener las diferentes *estructuras de colaboración* para la cuantificación de ésta. Como acabamos de ver, distintos motivos pueden haber impulsado el trabajo conjunto de los científicos. Las relaciones sociales que subyacen a esa cooperación pueden estar determinadas por variados factores, tales como la aportación de medios de cada participante o la responsabilidad última del modo en que se desarrolle la tarea; por eso, tanto las razones que deciden la elección de "socios" para un trabajo determinado como el papel asumido por cada colaborador pueden ser distintos según el ámbito que se estudie. Estos dos aspectos atraen el máximo interés cuando se pretende comprender las relaciones de colaboración establecidas en el seno de un sistema científico. De ahí que al querer medir la colaboración deba prestarse atención en primer lugar al tipo que cabe esperar que se dé en el conjunto de casos bajo escrutinio.

Dependiendo de que exista o no una diferenciación clara en el nivel de responsabilidad de unos colaboradores y otros, podemos distinguir *dos clases básicas de estructuras de colaboración*: en una, todos los participantes en un trabajo científico realizan tareas complementarias y todos pueden participar en la coordinación de esas tareas; por eso el término *democrática* es adecuado para designarla; en la otra, existe algún participante que asume la dirección y la responsabilidad del trabajo, al tiempo que los otros realizan las tareas que éste les encomienda en un contexto de dependencia funcional; a esta clase la denominaremos *jerárquica*. Mientras que en la estructura democrática suponemos que todos los participantes en un evento concreto colaboran entre sí, es decir todos son colaboradores respecto al resto de los participantes, en la jerárquica la colaboración se establece solamente entre los niveles funcionales, o en

otras palabras, cada agente dependiente colabora con el agente principal, pero se considera que los agentes dependientes no establecen colaboración entre sí. Dentro de la estructura jerárquica, podemos separar dos subclases: *bidireccional* si asumimos que la colaboración es algo mutuo entre el supervisor y cada dependiente; *unidireccional* si pensamos que uno es colaborador del otro pero no viceversa. La utilidad de esta diferenciación es manifiesta cuando lo que se busca son las pautas de preferencia entre agentes o agregados científicos: en ciertos contextos, la decisión en cuanto a quienes serán los socios puede partir exclusivamente de uno de los dos polos descritos. Cuando no se sabe con certeza qué tipo de colaboración se produce, también puede emprenderse la comprobación de si existen o no diferencias significativas al suponer que se dé una u otra estructura.

El esquema de colaboración cuya elección se estime más conveniente para un estudio concreto determinará entre qué agentes se contabilizará la colaboración de cada evento: si se asume que existe una colaboración jerárquica, sólo se contabilizará entre el agente principal y cada uno de los otros subordinados; si es más adecuado suponer una democrática, la colaboración se contabilizará para cada una de las posibles parejas de los agentes que participan. Pero sea cual sea el tipo de colaboración que se haya establecido, tendremos todavía que enfrentarnos a distintas alternativas en el modo de concebir qué sea una colaboración; estas alternativas definen una serie de nociones de la colaboración, asociada cada una de ellas a una manera irreductible de efectuar el recuento.

Para empezar, la ambigüedad de muchos nombres abstractos de nuestro lenguaje, que pueden referirse tanto a una *acción* como al *efecto* de esa acción, nos impone una primera elección. Podemos, efectivamente, pensar, por un lado, que la colaboración es el hecho mismo de que varios agentes figuren conjuntamente como productores. La *coaparición* es, entonces, el criterio para asignar colaboración a varios agentes. Y, por otro lado, puede llamarse colaboración a la necesaria transferencia de trabajo o de resultados para que una tarea pueda ser completada por varios agentes. En este segundo caso, es el *complemento* que cada uno aporta al total de la tarea realizada a lo que llamamos "su colaboración". Tenemos así ya dos modos de concebir (y de cuantificar) la colaboración: como *coaparición* o como *complemento*.

Pero enseguida nos vemos abocados a otra nueva decisión, la que especifica *quiénes son los sujetos* de la colaboración. Esta duda puede

parecer gratuita cuando hemos partido de la suposición de que los participantes en un trabajo de colaboración eran perfectamente identificables. Pero la cuestión es insoslayable en la práctica, ya que en los estudios de colaboración habitualmente interesa estudiar las relaciones que se establecen entre agregados significativos *a través de sus elementos constitutivos*, es decir, que se asume que los investigadores, unidades de investigación, organismos, etc., actúan en representación del agregado al que pertenecen. Ahora bien, al cuantificar la colaboración hemos de suponer o bien que ésta tiene lugar directamente entre los agregados bajo estudio, o bien que son sus unidades las que mantienen la colaboración efectiva. En el primer caso, todos los agregados participantes tienen la misma relevancia, sin importar el número de unidades que cada uno ha puesto en juego; en el segundo, por el contrario, se considera que el papel de cada agregado es en algún sentido proporcional a la cantidad de unidades que aporta a cada colaboración. Si estamos analizando la colaboración entre departamentos universitarios, por ejemplo, podemos fijarnos solamente en el hecho de que aparecen juntos autores de departamentos diferentes, independientemente de cuántos pertenezcan a cada departamento; o, por el contrario, considerar que es más intensa la colaboración cuantos más profesores aporte cada departamento. Aquellos tipos de recuento que ignoran el número de unidades de niveles inferiores de cada agregado los denominamos *holistas*; a los que tienen en cuenta el número de unidades aportadas por cada uno, los llamaremos *mereológicos*, porque consideran que la colaboración ocurre entre las partes constituyentes de los agregados. Es posible efectuar recuentos de estos dos tipos basándose en la noción de coaparición o en la de complemento.

Aún debemos solventar otra cuestión: ¿cuánta colaboración se produce en cada uno de los eventos (e.d., publicaciones, ...) de la serie que analizamos? Cada uno de esos eventos define un contexto que puede ser tasado de distinto modo. El problema que se plantea es si debemos asignar el mismo valor a la colaboración que se da entre un par de agentes cuando en la misma tarea participan también otros agentes que a otra en la que sólo participan esos dos agentes. Puede pensarse que la *intensidad* de la colaboración entre dos agentes es mayor cuantos menor sea el número de terceros participantes; pero también es legítimo considerar que la intensidad es la misma en cualquier caso. Cada opción define un tipo de recuento: de la primera resultan los recuentos *fraccionados*, en los

que el valor de la colaboración disminuye en proporción al número de colaboradores que se considere, ya que se asume que cada evento representa siempre una unidad de colaboración; en el segundo caso, tendremos recuentos *completos*, en los que la colaboración producida por un evento es mayor cuantos más colaboradores haya. La opción entre recuento completo y recuento fraccionado afecta solamente a los recuentos basados en la coaparición.<sup>64</sup>

### Clasificación de métodos de recuento de la colaboración

Co-aparición	Holistica .....	Completa .....	$C(A,B) = 1 (B)$	
		Fraccional .....	$C(A,B) = 1 / (T-1)$	
	Mereológica .	Completa ..	Simple ..	$C(A,B) = b$
			Nexos ..	$C(A,B) = a * b$ $C(A,A) = (a * (a-1))/2$
Fraccional ..		Simple ..	$C(A,B) = b / (t-a)$	
		Nexos ..	$C(A,B) = (a * b) / ((t*(t-1))/2)$ $C(A,A) = (a*(a-1)) / (t*(t-1))$	
Complemento	Holistica .....	$C(A,B) = 1/T$ $C(A,A) = 1/T$		
	Mereológica .....	$C(A,B) = b/t$ $C(A,A) = a/t$		

Para cada caso de colaboración:

**C(A,B)**: colaboración entre A - B

**A**: agregado A (país, institución...)

**B**: agregado B

**T**: total de agregados participantes

**a**: unidades de A participantes

**b**: unidades de B participantes

**t**: total de unidades participantes

El anterior análisis conceptual sistematiza las posibilidades de recuento de la colaboración. Ocho métodos básicos pueden ser aplicados

<sup>64</sup> Si tomamos como criterio primario la noción de complemento, sólo es posible con-  
testar que cada evento representa exactamente *una* colaboración, puesto que las aporta-  
ciones de todos suman la tarea finalmente realizada. En otras palabras, los recuentos  
basados en el complemento son inevitablemente *fraccionados*.

en cada una de las estructuras de colaboración distinguidas anteriormente. La tabla anterior muestra las relaciones y definición de esos ocho tipos básicos de recuento.

Por último, en el caso de la coaparición mereológica, nos encontramos con dos modos posibles de contabilizar las colaboraciones: el primero suma a la colaboración entre dos agregados tantas colaboraciones como representantes del otro agregado aparezcan; el segundo, entiende la colaboración la red de conexiones que puede haberse establecido entre todos los colaboradores. Cada participante está vinculado a todos los demás, y cada lazo cuenta como una colaboración. Llamaremos al primer caso recuento *simple*; y el segundo, de *nexos*.

Veamos, a continuación, cómo funciona cada método. Para ello aplicaremos a un mismo supuesto hipotético cada uno de los procedimientos de recuento. Supongamos que queremos calcular la matriz que representa la colaboración producida entre tres agregados,  $E$ ,  $F$ , y  $G$ , cada uno constituido respectivamente por una serie de unidades de nivel inferior

$$E=\{e_1, \dots, e_u\}; F=\{f_1, \dots, f_v\}; G=\{g_1, \dots, g_w\},$$

y que a lo largo de un periodo determinado han colaborado en tres ocasiones (artículos, proyectos, etc.):

$$\begin{array}{ll} \text{caso 1:} & e_1 e_2 e_3 f_1 f_2 g_1 \\ \text{caso 2:} & f_1 g_2 \\ \text{caso 3:} & e_1 f_2 f_3 \end{array}$$

Denotaremos por  $T$  el número de *agregados diferentes* cuyas unidades han tomado parte en cada caso, y por  $t$ , el número de *unidades* que han participado. Para abreviar, llamaremos  $e$ ,  $f$  y  $g$  respectivamente al *número de unidades* de  $E$ ,  $F$  y  $G$  que participan en cada caso. Así, en el tendremos para cada uno de los casos anteriores:

$$\begin{array}{ll} \text{caso 1:} & T=3; t=6; (e=3; f=2; g=1) \\ \text{caso 2:} & T=2; t=2; (e=0; f=1; g=1) \\ \text{caso 3:} & T=2; t=3; (e=1; f=2; g=0) \end{array}$$

Para ayudar a la comprensión podemos concretar algunos ejemplos posibles que responderían a este supuesto abstracto. Los criterios para

definir agregados pueden ser institucionales o geo-políticos; las unidades que los constituyen pertenecen todas a un nivel inferior que se determine. Así, si tomamos como agregados  $E, F, G$  un conjunto de países, las unidades  $e_i, f_i, g_i$  podrían ser las instituciones de investigación de esos países; también podríamos tomar como agregados los departamentos universitarios de un país y como sus unidades los individuos que pertenecen a cada departamento; o definir como agregados el tipo institucional y como unidades los departamentos que pertenecen a instituciones de cada tipo institucional definido<sup>65</sup>.

Los casos son las ocasiones en las que se colabora, tales como artículos científicos o proyectos de investigación en los que participan unidades de varios de los agregados que configuran el sistema bajo estudio. Por ejemplo, si analizamos la colaboración en la producción científica, los casos son las publicaciones en las que están representados varios de los agregados que nos interesan. Diremos entonces que dos agregados colaboran entre sí cuando aparecen juntos en alguno de estos casos.

Cada celda de la matriz contendrá al final el resultado de sumar la colaboración producida entre el par de agregados correspondiente a lo largo de los tres casos. La función para calcular esos sumandos dependerá del método que se elija; denotaremos esa función mediante la expresión

$$\text{col}_n(A, B)$$

que se lee "colaboración del agregado  $A$  con el agregado  $B$  en el caso  $n$ ", teniendo en cuenta que en varios métodos el valor puede cambiar con el orden de los agregados. Si unidades de  $A$  y  $B$  aparecen juntas en un caso de colaboración, ésta valdrá lo expresado por la fórmula; si no hay unidades de alguno de los agregados  $A$  o  $B$ , la colaboración entre los dos para ese caso valdrá, naturalmente, cero. Por tanto, el valor que contendrá una celda será:

---

<sup>65</sup> Habitualmente se estudia la colaboración entre los componentes de subsistemas, por lo que puede resultar útil definir como agregado "extra" el conjunto de las colaboraciones de esos componentes con el exterior del sistema: por ejemplo, si se estudia un país, la colaboración con el extranjero puede definir un agregado relevante; si se estudia un conjunto de países, lo será el agregado constituido por el resto de países no incluidos en el conjunto, y así sucesivamente.

$$AB = \Sigma \text{col}_n (A,B)$$

para los  $n$  casos en que aparezcan unidades de  $A$  y de  $B$ .

A continuación se aplican los distintos métodos y se comentan sus peculiaridades. Pero antes conviene hacer notar que las matrices resultantes se caracterizan por las siguientes propiedades:

- (a) Simetría / no simetría. Decimos que una matriz es *simétrica* cuando para cualquier par de agregados  $A$  y  $B$ , el valor de la celda  $AB$  es igual al de la celda  $BA$ , es decir:  $\Sigma \text{col}_n (A,B) = \Sigma \text{col}_n (B,A)$ , porque para cualquier caso  $i$ ,  $\text{col}_i (A,B) = \text{col}_i (B,A)$ .
- (b) Valor de la diagonal. La diagonal tiene *ceros estructurales* cuando para ningún agregado  $A$ , se puede calcular  $\text{col} (A,A)$ , es decir,  $\Sigma \text{col}_n (A,A)$  no tiene valor asignado, y la celda  $AA$  estará vacía. Sin embargo, existen métodos que sí rellenan la diagonal.
- (c) Significado de totalizaciones de filas y columnas. Dependiendo del método, los totales pueden reflejar el número de casos en que se participa, la contribución de cada agregado, el número de nexos establecidos, etc.

Una vez descritos los elementos del supuesto y establecidos los criterios para la comparación de las matrices, éstos son los resultados que produce cada método aplicado al supuesto anterior:

1. *Coaparición holista completa*. Cuenta el número de casos (publicaciones, proyectos...) en el que un cierto par de agregados aparecen juntos, sin importar el número de unidades de nivel inferior que cada uno aporte ni el número de agregados que participan en cada caso. La matriz que genera este método es simétrica, con ceros estructurales en la diagonal; el total de filas y el de columnas reflejan el recuento múltiple respecto al número de casos en los que se ha intervenido, debido a los casos en que más de dos agregados tomaron parte en la colaboración.

$$\text{col}_n(A,B) = 1$$

p.ej.:  $\text{col}_1(E,F)=1$

	E	F	G	ttl
E	-	2	1	3
F	2	-	2	4
G	1	2	-	3
ttl	3	4	3	(10)

2. *Coaparición holista fraccionada*. Con este método, por cada vez en que un agregado aparezca con otros se cuenta una sola colaboración, de modo que se reparte proporcionalmente cuando son más de un colaborador. Como no tiene en cuenta la participación de unidades de nivel inferior, la fracción es el inverso del número de agregados diferentes con los que se colabora en ese caso.

$$\text{col}_n(A,B) = 1 / (T-1)$$

p.ej.:  $\text{col}_1(E,F)= 1/2$

	E	F	G	ttl
E	-	1,5	0,5	2
	1,5	-	1,5	3
F	0,5	1,5	-	2
ttl	2	3	2	

3. *Coaparición mereológica completa (simple)*. Este tipo de recuento considera la colaboración como la aparición conjunta de unidades pertenecientes a los agregados bajo estudio. Así, la colaboración de un agregado *A* con otro *B* ocurre tantas veces como unidades pertenecientes a *B* tomen parte en el caso correspondiente.

$$\text{col}_n(A,B) = b$$

p.ej.:  $\text{col}_1(E,F)= 2$

	E	F	G	ttl
E	-	4	1	5
F	4	-	2	6
G	3	3	-	6
ttl				

4. *Coaparición mereológica completa (nexos)*. Como alternativa a la mera coaparición del método anterior (3), puede considerarse que se producen tantas colaboraciones en un caso como vínculos puedan establecerse entre las unidades pertenecientes a los agregados que intervienen, es decir, todos los posibles pares que puedan definirse entre las

unidades participantes. En una publicación en la que participan varias unidades de diferentes agregados, la colaboración entre los agregados *A* y *B* será igual al producto de las unidades aportadas por cada agregado. También tiene sentido definir en este procedimiento la colaboración entre las unidades de un mismo agregado, por lo que la matriz tiene rellena la diagonal y permite estudiar el grado de autonomía mutua de las unidades de un agregado para establecer colaboraciones con unidades pertenecientes a otros agregados.

$col_n(A,B) = a \cdot b$		<b>E</b>	<b>F</b>	<b>G</b>	<b>ttl</b>
$col_n(A,A) = (a \cdot (a-1)) / 2$	<b>E</b>	3	8	3	14
	<b>F</b>	8	2	3	13
p.ej.: $col_1(E,F) = 6$	<b>G</b>	3	3	0	6
	<b>ttl</b>	14	13	6	

**5. Coaparición mereológica fraccionada (simple).** Este método considera, como el tipo 2, que cada caso u ocasión (publicación, proyecto...) representa una sola colaboración, y por tanto que la intensidad de la colaboración entre un par de colaboradores es tanto menor cuanto mayor sea el número de otros participantes. Pero ahora se asume que los protagonistas de la colaboración son las unidades de los agregados, por lo que la colaboración se reparte en función del número de unidades de otros agregados que hayan participado en esa ocasión.

$col_n(A,B) = b / (t-a)$		<b>E</b>	<b>F</b>	<b>G</b>	<b>ttl</b>
	<b>E</b>	-	5/3	1/3	2
	<b>F</b>	7/4	-	5/4	3
p.ej.: $col_1(E,F) = 2/3$	<b>G</b>	3/5	7/5	-	2
	<b>ttl</b>				

**6. Coaparición mereológica fraccionada (nexos).** Este tipo supone que la intensidad de los vínculos generados entre las unidades desciende con el aumento de unidades participantes, de modo que asigna una fracción proporcional al número de unidades, dividiendo entre todos los nexos posibles el único que se supone tiene lugar entre todas las unidades en cada caso. También se cuenta la colaboración que tiene lugar entre las unidades del mismo agregado, por lo que las diagonales de la

matriz tienen sentido.

$$\text{col}_n(A,B) = (a*b)/((t*(t-1))/2)$$

$$\text{col}_n(A,A) = (a*(a-1))/(t*(t-1))$$

p.ej.:  $\text{col}_1(E,F) = 6/15$

	E	F	G	ttl
E	0,2	1,07	0,2	1,47
F	1,07	0,4	1,13	2,6
G	0,2	1,13	0	1,33
ttl				

**7. Complemento holista.** Aquí es la aportación estimada de cada uno de los agregados lo que se toma como colaboración. Se asume que la colaboración tiene lugar entre los agregados, por lo que se no tiene en cuenta el número de unidades de cada agregado. La colaboración de un agregado con otro es entonces la fracción proporcional al número de agregados que intervienen. Se obtiene una matriz *simétrica*, con diagonales rellenas y sumas de filas y columnas en las que aparece el número de casos en los que ha participado cada agregado.

$$\text{col}_n(A,B) = 1 / T$$

$$\text{col}_n(A,A) = 1 / T$$

p.ej.:  $\text{col}_1(E,F) = 1/3$

	E	F	G	ttl
E	0,83	0,83	0,33	2
F	0,83	1,33	0,83	3
G	0,33	0,83	0,83	2
ttl	2	3	2	

**8. Complemento fraccionado.** El concepto es igual al anterior, salvo que la aportación de cada agregado se supone proporcional al número de unidades que pone en juego. Pero la matriz que resulta es *asimétrica*, ya que el número de unidades de cada agregado puede ser diferente, tiene las *diagonales rellenas* con la aportación del agregado correspondiente, y en las sumas de las filas aparecen el *número total de casos* en los que intervienen unidades del agregado correspondiente. Las columnas reflejan el exceso debido a la repetición de las aportaciones del agregado que las encabeza.

$col_n(A,B) = b / t$		<b>E</b>	<b>F</b>	<b>G</b>	<b>ttl</b>
$col_n(A,A) = a / t$	<b>E</b>	0,83	1	0,17	2
	<b>F</b>	0,83	1,5	0,67	3
p.ej.: $col_1(E,F) = 2/6$	<b>G</b>	0,5	0,83	0,67	2
	<b>ttl</b>				

### 7.2.2. Comentarios sobre los métodos de recuento de la colaboración

Lo que hemos denominado enfoque *relacional* nos ofrece la posibilidad de enriquecer de modo superlativo las posibilidades del enfoque que llamábamos *simple*, al mismo tiempo que nos brinda una sistematización de éstos de la que hasta ahora carecíamos. Así, podemos obtener los indicadores *simples* al uso y otros análogos (más finos), partiendo de un tratamiento de los datos de tipo relacional. Para ello, calcularemos por un lado cuatro matrices y, por otro, el número de casos totales en los que participa cada agregado, incluyendo aquellos en los que no se colabora. Las cuatro matrices son las siguientes:

- (a) número de **Casos de CoLaboración** en los que aparecen juntos cada pareja de agregados (matriz **CCL**);
- (b) **Suma del número de AGregados** diferentes representados en cada coaparición de cada pareja, (matriz **SAG**) y
- (c) número total de **Participaciones de UniDades**, para las coapariciones de cada pareja de agregados (matriz **PUD**);
- (d) una versión modificada de la matriz obtenida mediante el recuento nº 3 (coaparición mereológica completa simple), cuyas diagonales se han rellenado con el número de unidades que aporta cada agregado cuando colabora con otros; en la celda **AB**, por ejemplo, está el total de las apariciones de **Unidades** de **B** en los casos en que ha **CoLaborado** con **A**. La abreviatura para esta matriz será **UCL**.

Los valores de cada una de estas matrices para el mismo caso hipotético anterior se muestran a continuación. Además, se ha supuesto que los agregados **E**, **F** y **G** habían participado en otros 12, 15 y 19 casos sin colaboración respectivamente (por eso el total de casos para cada agregado es de 14, 18 y 21).

(Recordemos los tres casos:  $e_1 e_2 e_3 f_1 f_2 g_1 - f_1 g_2 - e_1 f_2 f_3$ ).

casos. CCL				agregados SAG (sumas de T)				unidades PUD (sumas de t)				TTL
	E	F	G		E	F	G		E	F	G	
E	2	2	1	E	5	5	3	E	9	9	6	14
F		3	2	F		7	5	F		11	8	18
G			2	G			5	G			8	21
ttl			3	ttl			7	ttl			11	49

UCL unidades aportadas			
	E	F	G
E	4	4	1
F	4	5	2
G	3	3	2
ttl			

Con los datos contenidos en estas matrices y en la columna del total de casos, podemos obtener fácilmente una serie de indicadores simples y sus equivalentes para cada pareja de agregados que nos interese, es decir, lo que podríamos llamar sus *equivalentes relacionales* o indicadores simples aplicados a parejas determinadas. Podemos expresar formalmente los indicadores posibles si tomamos cada matriz como una función que asigna un valor a cada una de sus celdas. Para dos elementos cualesquiera,  $A$  y  $B$ , y para el sistema en su conjunto, tendríamos entonces:

Tasa de colaboración del sistema:	.....	$CCL(ttl) / TTL(ttl)$
Tasa de colaboración de A:	.....	$CCL(AA) / TTL(AA)$
Presencia de A en la colaboración:	.....	$CCL(AA) / CCL(ttl)$
.....	.....	$SAG(AA) / CCL(AA)$
Simetría de la colaboración entre A y B	.....	$UCL(AB) / UCL(BA)$
Presencia de B en la colaboración de A	.....	$SAG(AB) / SAG(AA)$
Peso de B en la colaboración de A	.....	$UCL(AB) / PUD(AA)$

Del mismo modo pueden definirse algunos indicadores más<sup>66</sup>, cuya necesidad vendrá dada por cada estudio concreto.

Otro interesante apartado de indicadores de colaboración simples es el constituido por el estudio que podemos llamar "espectrográfico" de la colaboración llevada a cabo por un agregado cualquiera. Se trata de obtener la distribución de pesos entre cada nivel de colaboración, según los segmentos que se establezcan: por ejemplo, determinar cuántos documentos iban firmados por dos autores (o unidades, instituciones...), cuántos por tres, cuatro,... etc. Puede servir para detectar patrones particulares en la producción de instituciones, de áreas científicas o de otros agregados que se definan.

**Una nota metodológica.** La variedad de matrices que pueden obtenerse con los mismos datos está originada por dos conjuntos de razones principales: (a) en primer lugar, como hemos visto, de la diversidad de concepciones posibles de la colaboración misma, derivada de las distintas *estructuras* que pueden suponerse, de *nociones* básicas alternativas y de la existencia de *varios niveles* de agregación; y (b) en segundo lugar, de la decisión de *reducir* un fenómeno *multidimensional* a uno *bidimensional*, con el fin de aplicar técnicas estadísticas disponibles.

Respecto al primer conjunto de causas, hay que notar que cuando no es posible considerar unidades de nivel inferior, sólo son aplicables tres de los ocho métodos descritos (los denominados "holistas", marcados con los números 1,2 y 7). Respecto a la necesidad de reducción a una matriz bidimensional, diremos que existe la posibilidad de analizar por separado todas las combinaciones posibles entre los colaboradores: el número de subconjuntos de colaboración posibles, para un conjunto de  $n$  elementos es igual a:  $2^n - n - 1$  ( $2^n$  es la fórmula de Cantor de la potencia de un conjunto, a lo que hay que restar los subconjuntos de 1 elemento y el conjunto vacío). También podemos analizar  $2^n - 1$  subconjuntos, si queremos incluir los casos en los que no hay colaboración, lo que permite estudiar diferencias en el nivel general de colaboración de cada agregado. Una vez contabilizadas las ocasiones en las que se ha dado cada combinación, puede emprenderse un análisis de las frecuencias de cada combinación. Si los elementos son pocos, no es difícil estudiar, por ejemplo, las relaciones condicionantes entre los elementos. Los datos así

---

<sup>66</sup> Pero por otro lado, además de esos indicadores, cada una de las matrices generadas por los ocho métodos de recuento relacionales anteriormente expuestos puede por separado servir para definir indicadores acerca de los pesos de la colaboración de cada agregado o de cada pareja de agregados.

dispuestos requieren un análisis complejo, pero que también puede develar pautas interesantes, tales como: "A y B sólo colaboran cuando aparece también C" u otras por el estilo.

El análisis de la colaboración científica presenta, como se ha visto, un rico y complejo panorama. Su estudio metodológico ha servido como muestra de la determinación que pueden ejercer las técnicas concretas de medida sobre los conceptos básicos con los que se interpretan los valores de esas medidas.

---

## CONCLUSIONES

A lo largo de la exposición anterior se han ido destacando oportunamente los puntos principales. De todos modos, me gustaría enfatizar aquellos que considero más significativos.

En primer lugar, el examen del sistema de publicación de la ciencia desde una perspectiva integradora ha permitido situarlo como pieza capital del proceso en el que los científicos producen nuevos conocimientos. Al lado de la función comunicativa y diseminadora que se suele acentuar, se ha mostrado que el servicio que presta a la comunidad científica y a sus fines alcanza también a otros aspectos fundamentales.

Se ha extendido la noción de *reconocimiento* heredada de la tradición sociológica mertoniana (independizándola de ésta) y se ha presentado como el impulsor y modelador básico de la publicación científica oficial. El intento de robustecer esta noción ha servido para resaltar lo imprecipitante que ha sido su papel en la ciencia moderna, más allá de su inmediata referencia a la dimensión colectiva de ésta.

La caracterización de la noción de *resultado científico* articula la tesis. Un resultado científico es aquello que se obtiene mediante un ejercicio riguroso de la actividad científica y que se puede percibir como *nuevo* y *relevante* para la ciencia por parte de los investigadores insertos en la especialidad correspondiente. Por un lado, este concepto dota de un sentido completo a la idea de reconocimiento que aquí se ha delineado, en sus contenidos, en la necesidad de obtenerlo que sienten los científicos y en los mecanismos por los que se obtiene. Además, al fundar la interpretación del sistema de publicación oficial como el lugar en el que se

constituyen los resultados científicos, permite alejar a este sistema de la sospecha continua originada en el cumplimiento insatisfactorio de funciones que a esta luz se revelan inadecuadas. Las publicaciones oficiales se presentan como cristalizaciones de resultados científicos y queda justificado de modo general el usarlas como fuente de datos para el análisis bibliométrico.

La noción de resultado científico está, al mismo tiempo, ligada al *contenido* de lo que hacen los científicos y al *proceso* (entendido en su sentido más amplio) por el que aspiran a conseguir ese contenido. Además de integrar estas dos vertientes tradicionalmente separadas, establece la posibilidad de penetrar en aspectos epistémicos de la ciencia por medios externalistas que parten de las trazas materiales dejadas por ese proceso.

Representa, por otro lado, al posibilidad de encontrar un significado para los dos tipos de indicadores bibliométricos más básicos. La producción científica se entiende ahora como la suma de *unidades mínimas de novedad y relevancia científicas*. De este modo se puede reconocer el valor de lo que hacen los investigadores sobre la base de pautas internas a la comunidad científica pero sin necesidad de comprometerse con exigencias desproporcionadas y difíciles de contrastar. La práctica científica cotidiana es abarcada por este enfoque de una manera natural.

Los indicadores de calidad recogen el fruto de la separación entre *resultados y conocimiento* científicos. La valoración interna de los resultados ofrece un fundamento para una concepción de la calidad que es adecuada para dirigir acciones externas sobre el sistema científico y al mismo tiempo está estrechamente vinculada a los objetivos propios de ese sistema.

Los indicadores de colaboración han mostrado las posibilidades de adentrarse en el estudio de las relaciones entre los agentes e instituciones científicos productores de resultados, al tiempo que alertaban de la delicada relación entre las operaciones concretas de medida y los conceptos que pueden atribuirse en cada caso.

A través de los análisis de la construcción de estos tres tipos de indicadores se ha puesto de manifiesto que no es pequeño el conjunto de precauciones que requiere el cálculo y la aplicación de las medidas bibliométricas.

Todo contribuye a mejorar la fundamentación de los indicadores bibliométricos. La posibilidad de una interpretación sólida de estas

medidas fortalece el uso que se pueda hacer de ellas y abre vías en el estudio cuantitativo de la ciencia. En cuanto a lo primero, legitima un uso externo de esas medidas que no traiciona las aspiraciones básicas de los científicos, aclara el significado de esas medidas y delimita todo un campo de aplicaciones concretas. En cuanto a lo segundo, ofrece un fundamento general sobre el que edificar las aproximaciones a cada aspecto particular.

Quisiera, para terminar, dejar escrito un pequeño comentario. Este trabajo ha sido concebido como una visita desde la filosofía de la ciencia a un territorio nuevo para la cosecha. Se ha evitado, en lo posible, imponer las categorías habituales en los debates filosóficos contemporáneos. Sería positivo que se hayan alcanzado aquí resultados que pudieran enriquecer esos debates con otro punto de vista. Pero, al menos, creo que esta discusión ha perfilado un foco de atención. La preocupación tradicional de los filósofos de la ciencia ha sido, de distintos modos, la caracterización y la demarcación de su objeto de estudio. Habría que preguntarse si el *modo de acceso* a ese objeto ofrece también aspectos dignos de consideración.



---

## BIBLIOGRAFÍA

- ALLEN, T.J. (1977): *Managing the flow of Technology*, Cambridge: MIT Press, 1977.
- ALLISON, P.D.; STEWART, J.A. (1974): "Productivity differences among scientists: evidence for accumulative advantage", *American Sociological Review*, v.39 (1974) 596-606.
- ALLISON, P.D.; LONG, J.S.; KRAUZE, T.K. (1982): "Cumulative advantage and inequality in science", *American Sociological Review*, v.47 (1982 oct) 615-625.
- ALLISON, P.D.; LONG, J.S. (1987): "Interuniversity mobility of academic scientists" *American Sociological Review*, v.52 (1987) 643-652.
- ALLISON, P.D.; PRICE, D.J. DE S.; GRIFFITH, B.C.; MORAVCSIK, M.J.; STEWART, J.A. (1976): "Lotka's law: A problem in its interpretation and application", *Social Studies of Science*, v.6 (1976) 269-276.
- ALLISON, P.D.; LONG, J.S. (1990): "Departmental effects on scientific productivity" *American Sociological Review*, v.55 (1990 ago) 469-478.
- ALLISON, P.D. (1978): "Measures of inequality", *American Sociological Review*, v.43 (1978 dec) 865-880.
- ALLISON, P.D. (1980): "Inequality and scientific productivity", *Social Studies of Science*, v.10, (1980) 163-179.
- AMSTERDAMSKA, O.; LEYDESDORFF, L. (1988): "Citations: indicators of significance?", *Scientometrics*, v.15, n.5-6 (1989) 449-471.
- ANDERSON, J.; COLLINS, P.M.D; IRVINE, J.; ISARD, P.A.; MARTIN, B.R.; NARIN, F.; STEVENS, K. (1988): "On-line approaches to measuring national scientific output: a cautionary tale", *Science and Public Policy*, v.15, n.3 (1988 jun) 153-161.
- ANDERSON, R.C.; NARIN, F.; McALLISTER, P. [1978]: "Publications ratings versus peer ratings of universities", *Journal of the American Society for Information Science*, v.29 (1978) 91-103.
- ANDREWS, F.(ed.): *Scientific Productivity*, Cambridge: Cambridge Univ. Press, 1979.
- ATKINSON, M. (1993): "Regulation of science by 'peer review'", *Studies in History and Philosophy of Science*, v.25, n.2 (1994) 147-158.
- AVERCH, H. (1979): "Science indicators and policy analysis", *Scientometrics*, v.6, n.5-6 (1980) 339-345.
- BAKANIC, V.; McPHAIL, C.; SIMON, R.J. (1986) "The manuscript review and decision-making process", *American Sociological Review*, v.52, n.5 (1987 oct) 631-642.
- BARBER, B. (1952): *Science and the Social Order*, New York, The Free Press, 1952.

- BARBER, B. (1961): "Resistance by scientists to scientific discovery", *Science*, v.134 (1961 sep) 596-602.
- BARNES, B. (1974): *Scientific knowledge and sociological theory*, London: Routledge & Kegan Paul, 1974.
- BARNES, B. (1977): *Interests and the growth of knowledge*, London: Routledge & Kegan Paul, 1977.
- BARNES, B. (1985): *Sobre ciencia*, Barcelona: Labor, 1987 (trad. de Faci Lacasta, J. de la v.o. *About Science*, Oxford: Blackwell, 1985).
- BARNES, B.; EDGE, D. (1982): *Science in context*, Milton Keynes, Open Univ. Press, 1982.
- BARNES, B.; DOLBY, R.G.A. (1970): "The scientific ethos: A deviant viewpoint", *Archives of European Sociology*, n.11 (1970) 3-25.
- BARRE, R. (1987): "A strategic assessment of the scientific performance of five countries", *Science & Technology Studies*, v.5 (1987) 32-38.
- BAYER, A.E.; FOLGER, J. (1966): "Some correlates of a citation measure of productivity in science", *Sociology of Education*, v.39 (1966) 381-390.
- BEAVER, D. DE B.; ROSEN, R. (1978a): "Studies in scientific collaboration: Part I. The professional origins of scientific co-authorship", *Scientometrics*, v.1, n.1 (1978) 65-84.
- BEAVER, D. DE B.; ROSEN, R. (1978b): "Studies in scientific collaboration: Part II. Scientific co-authorship, research productivity and visibility in the French scientific elite, 1799-1830.", *Scientometrics*, v.1, n.2 (1978) 133-149.
- BEAVER, D. DE B.; ROSEN, R. (1978c): "Studies in scientific collaboration: Part III. Professionalization and the natural history of modern scientific co-authorship", *Scientometrics*, v.1, n.3 (1978) 231-245.
- BECK, I.M (1984): "A method of measurement of scientific production quality", *Science of Science*, v.4, n.2 (1984) 183-195.
- BELLAVISTA, J. et al. (1993): *Política científica y tecnológica. Evaluación del I+D en la Universitat de Barcelona*, Barcelona: Universidad de Barcelona, 1993.
- BLOOR, D. (1976): *Knowledge and social imagery*, Londres: Routledge & Kegan Paul, 1976.
- BLUME, S.S. (ed.) (1977): *Perspectives in the Sociology of Science*, New York: John Wiley & Sons, 1977.
- BONZI, S.; SNYDER, H.W. (1990): "Motivations for citation: A comparison of self citation and citation to others", *Scientometrics*, v.21, n.2 (1991) 245-254.
- BOOKSTEIN, A. (1988): "Applications of the bibliometric distributions", EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 5-13.
- BOOR, M. (1982): "The citation impact factor: another dubious index of journal quality", *American Psychologist*, v.37 (1982) 975-977.
- BORGMAN, C.L. (ed.) (1990): *Scholarly communication and bibliometrics*, Beverly Hills: Sage, 1990.
- BORING, E.G. (1927): "The problem of originality in science", *American Journal of Psychology*, 39 (1927 dic) 70-90. (citado por Merton en p.400)
- BOURDIEU, P. (1975): "The specificity of the scientific field and the social conditions of the progress of reason", *Social Science Information*, v.14, n.6 (1975) 19-47.

- BRAAM, R.R.; MOED, H.; VAN RAAN, A.F.J. (1988): "Mapping of science: critical elaboration and new approaches. A case study in agricultural biochemistry", en: EGGHE, L.; ROUSSEAU, R. (eds.): *Informetrics 87/88*, Amsterdam: Elsevier, 1988, pp.15-28.
- BRAAM, R.R.; MOED, H.F.; VAN RAAN, A.F.J. (1988): "Mapping of science: critical elaboration and new approaches, a case study in agricultural biochemistry", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 15-28.
- BRADFORD, S.C. (1948): *Documentation*, London: Crosby Lockwood.
- BRAUN, T.; GLÄNZEL, W.; SCHUBERT, A. (1985): *Scientometric indicators. A 32-country comparative evaluation of publishing performance and citation impact*, Singapore, Philadelphia: World Scientific Publ., 1985.
- BRAUN, T.; GLÄNZEL, W.; SCHUBERT, A. (1987): "One more version of the facts and figures on publication output and relative impact of 107 countries, 1978-1980" *Scientometrics*, v.11, n.1-2 (1987) 9-15.
- BRAUN, T.; SCHUBERT, A. (1988): "Scientometric versus socio-economic indicators. Scatter plots for 51 countries, 1978-1980", *Scientometrics*, v13 n.1-2 (1988) 3-9.
- BRIDGSTOCK, M. (1990): "The quality of single and multiple authored papers; an unresolved problem", *Scientometrics*, v.21, n.1 (1991) 37-48.
- BROAD, W.; WADE, N.(1982): *Betrayers of the truth: Fraud and deceit in the halls of science*, New York: Simon and Schuster, 1982.
- BRONCANO, F.; VEGA, J. (1994): "Inteligencia artificial en España. Detección de áreas de investigación en emergencia mediante análisis bibliométricos", *Política Científica*, n.40 (1994 jul) 43-48.
- BRONCANO, F.; MALTRÁS, B.; VEGA, J. (1995): "Léxico compartido en la identificación de focos multidisciplinares emergentes. El caso de los nuevos materiales en la producción científica española, *Actas de las III Jornadas RICTES*, (Valencia, 1995), 367-394.
- BRONCANO, F. (1994): "Cambio tecnológico y evolución: tres concepciones sobre las relaciones entre ciencia, técnica y sociedad", *Arbor*, CLII, n.598-599 (1995 oct-nov) 27-72.
- BROOKES, B.C. (1970): "Scientific bibliography", *Nature*, v.227 (1970 sep) 1377. (carta)
- BROOKES, B.C. [1993]: "Bibliométrie, scientométrie et informétrie", en ESTIVALS, R. (dir): *Les sciences de l'écrit. Encyclopédie Internationale de Bibliologie*, Paris: AIB, 1993.
- BROOKES, B.C. (1988): "Comments on the scope of bibliometrics", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 29-41.
- BROOKES, B.C. (1990): "Biblio-, sciento-, infor-metrics??? What are we talking about?", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 41-53.
- BROOKS, T.A. (1984): "Private acts and public objects: An investigation of citer motivations", *Journal of the American Society for Information Science*, v.36, n.4 (1985) 223-229.
- BROOKS, T.A. (1985): "Evidence of complex citer motivations", *Journal of the American Society for Information Science*, v.37, n.1 (1986) 34-36.
- BROOKS, H. (1979): "Science indicators and science policy", *Scientometrics*, v.6, n.5-6 (1980) 331-337.

- BROOKS, H. (1982): Science indicators and science priorities", en: LA FOLLETE, MARCEL (ed) *Quality in science*, Cambridge, MA: MIT Press, 1982, 1-32.
- BROOKS, T.A. (1990): "Perfect Bradford multipliers: A definition and empirical investigation", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 45-55.
- BUCHHOLZ, K. (1993): "Criteria for the analysis of scientific quality", *Scientometrics*, v.32, n.2 (1995) 195-218.
- BUNGE, M. (1969): *La investigación científica*, Barcelona: Ariel, 1980.
- CALLON, M.; COURTIAL, J.P.; TURNER, W.A.; BAUIN, S. (1983): "From translation to problematic network: an introduction to co-word analysis", *Social Science Information*, v.22 (1983) 191-235.
- CALLON, M.; LAW, J.; RIP, A. (eds.) (1986): *Mapping the dynamics of science and technology*, London: Macmillan Press, 1986.
- CALLON, M. (1993): "Is science a public good?" (Fifth Mullins Lecture, Virginia Polytechnic Institute, 23 march 1993), *Science Technology and Human Values*, v.19, n.4 (1994) 395-424.
- CALLON, M.; COURTIAL, J.-P.; TURNER, W. (1991): "La méthode leximappe: un outil pour l'analyse stratégique du développement scientifique et technique", en: D. VINCK (ed.), *Gestion de la recherche*, Brussels: De Boeck, 1991, 208-277.
- CALLON, M.; COURTIAL, J.P.; PENAN, H. (1991): *Qu'est-ce la scientiométrie?*, PUF, Paris, 1991.
- CAMÍ, J.; FERNÁNDEZ, M.T.; GÓMEZ, I. (1993): "La producción científica española en biomedicina y salud. Un estudio a través del Science Citation Index (1986-91)", *Medicina Clínica*, 101 (1993) 721-731.
- CAMPBELL, C. (1982): "A dubious distinction? An inquiry into the value and use of Merton's concepts of manifest and latent function", *American Sociological Review*, v.47 (1982) 29-44.
- CANO, F.; JULIÁN, S. (1991): "Some indicators in Spanish scientific production", *Scientometrics*, v.24, n.1 (1992) 45-61.
- CAPSHAW, H.; RADER, K.A. [1992]: "Big science: Price to the present", *Osiris*, 7 (1992) 3-25.
- CARPENTER, M.P.; NARIN, F. (1973): "Clustering of scientific journals", *Journal of the American Society for Information Science*, v.24, n.6 (1973) 425-36.
- CARPENTER, M.P.; NARIN, F. (1980a): "The subject composition of the world's scientific journals", *Scientometrics*, v.2 n.1 (1980) 53-63.
- CARPENTER, M.P.; NARIN, F. (1980b): "The adequacy of the Science Citation Index (SCI) as an indicator of international scientific activity", *Journal of the American Society for Information Science*, v.32 (1981) 430-39.
- CARPENTER, M.P.; GIBB, F.; HARRIS, M.; IRVINE, J.; MARTIN, B.R.; NARIN, F. (1988): "Bibliometric profiles for British academic institutions: an experiment to develop research output indicators", *Scientometrics*, v.14, n.3-4 (1988) 213-233.
- CAWKELL, A.E. (1978): "Understanding science by analysing its literature", *Information Scientist*, 10 (1978) 3-10.

- CHASE, J. (1978): "Normative criteria for scientific publication", *American Sociologist*, 4 (1978) 262-265.
- CHUBIN, D.E.; MOITRA, S.D. (1975): "Content analysis of references: Adjunct or alternative to citation counting?", *Social Studies of Science*, 5 (1975) 423.
- CICCHETTI, D.V. (1991): "The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation", *Behavioral and Brain Sciences*, v.14 (1991) 119-186.
- CLARK, K.E. (1957): *America's Psychologist: A Survey of a Growing Profession*, Washington: American Psychological Association, 1957.
- CLARKE, B.L. (1964): "Multiple authorship trends in scientific papers", *Science*, v.143 (1964) 822-.
- COLE, J.R. (1980): "Patterns of intellectual influence in scientific research", *Sociology of Education*, v.43 (1970) 377-403.
- COLE, J.R.; COLE, S. (1967): "Scientific output and recognition: A study of the reward system in science", *American Sociological Review*, 33 (1967 jun) 397-413. (ver tb. ASR, v.32 (1967) 377-390)
- COLE, J.R.; COLE, S. (1971): "Measuring the quality of sociological research: Problems with the use of the *Science Citation Index*", *The American Sociologist*, v.6 (1971 feb) 23-29.
- COLE, S.; COLE, J.R. (1968): "Visibility and the structural bases of awareness of scientific research", *American Sociological Review*, v.32 (1968 jun) 397-413.
- COLE, S.; COLE, J.R. (1973): *Social stratification in science*, Chicago: Univ. of Chicago Press, 1973.
- COLE, J.R.; ZUCKERMAN, H. (1975): "The emergence of a scientific specialty: A sociological study of the Sociology of Science", en: COSER, L.A. (ed.): *The idea of social structure: Papers in honor of Robert K. Merton*, New York: Harcourt Brace Jovanovich, 1975.
- COLE, J.R.; COLE, S. (1972): "The Ortega hypothesis", *Science*, 178 (1972) 368-375.
- COLE, J.R.; COLE, S. (1974): "Citation analysis", *Science*, 183 (1974) 32-33.
- COLE, S. (1983): "The hierarchy of the sciences?", *American Journal of Sociology*, v.89 (1983) 111-139.
- COLE, S. (1979): "Comments on 'Indicators of scientific manpower'", *Scientometrics*, v.2, n.5-6 (1980) 405-409.
- COLE, S.; COLE, J.R.; SIMON, G.A. (1981): "Chance and consensus in peer review", *Science*, 214 (1981 nov) 881-886.
- COLE, F.J.; EALES, N.B. (1917): "The history of comparative anatomy", *Science Progress*, 11 (1917) 587-596.
- COURTIAL, J.P.: *Artificial Intelligence networks: the contribution of the co-word analysis method to the study of a scientific field*, Paris: Centre Sociologique d'Innovation, École de Mines, 1987.
- COURTIAL, J.P. (1988): "Qualitative models, quantitative tools and network analysis", *Scientometrics*, v.15 n.5-6 (1989) 527-534.
- COURTIAL, J.P.; LAW, J. (1989): "A co-word study of AI", *Social Studies of Science*, v.19 (1989) 301-311.

- COURTIAL, J.P.; CALLON, M. (1991): "Indicators for the identification of strategic themes within a research programme", *Scientometrics*, v.21 (1991) 447-458.
- COZZENS, S. (1981): "Taking the measure of science: A review of citation theories", *International Society for the Sociology of Sciences Newsletter*, v.7 (1981) 21-60.
- COZZENS, S. (1985a): "Comparing the sciences", *Social Studies of Science*, v.15 (1985) 127.
- COZZENS, S. (1985b): "Using the archive: Derek Price's theory of differences among the sciences", *Scientometrics*, v.7 (1985) 431-41.
- COZZENS, S. (1988): "What do citations count? The rethorical-first model", *Scientometrics*, v.15, n.5-6 (1989) 437-447.
- CRANE, D. (1972): *Invisible Colleges: Diffusion of knowledge in scientific communities*, Chicago: Univ. of Chicago Press, 1972.
- CRANE, D. (1969): "Social structure in a group of scientists: a test of the 'invisible college' hypothesis", *American Sociological Review*, v.34 (1969) 335-352.
- CRANE, D. (1967): "The Gatekeepers of Science. Some Factors Affecting the Selection of Articles for Scientific Journals", *The American Sociologist*, v.2 (1967) 195-210.
- CRONIN, B. (1984): *The citation process*, London: Taylor Graham, 1984.
- CRONIN, B. (1981): "Agreement and divergence on referencing practice", *Journal of Information Science*, v.3 (1981) 27-33.
- CROUCH, D.; IRVINE, J.; MARTIN, B.R. (1985): "Bibliometric analysis for science policy: an evaluation of the United Kingdom's research performance in ocean currents and protein crystallography", *Scientometrics*, v.9, n.5-6 (1986) 239-267.
- DASGUPTA, P.; DAVID, P.A. (1994): "Toward a new economics of science", *Research Policy*, 23 (1994) 487-521.
- DE LA FUENTE, A. (1995): "A note on incentives and researcher productivity in Spanish public institutions", *Investigaciones Económicas*, v.XIX, n.2 (1995 may) 291-299.
- DE BRUIN, R.E.; MOED, H.F. (1990): "The unification of addresses in scientific publications", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 65-78.
- DIAMOND, A.M.J.R.: "An Economic Model of the Life-Cycle Research Productivity of Scientists", *Scientometrics*, v.3 (1984) 189-96.
- DIEKS, D.; CHANG, H.: "Differences in impact of scientific publications: Some indices derived from a citation analysis", *Social Studies of Science*, v.6 (1976) 247-267.
- DOREIAN, P. (1987): "A revised measure of standing of journals in stratified networks", *Scientometrics*, v.11 (1987) 71-80.
- DOREIAN, P. (1988): "Measuring the relative standing of disciplinary journals", *Information Processing and management*, v.24 (1988) 45-56.
- ECHEVERRIA, J. (1995): *Filosofía de la ciencia*, Madrid, Akal, 1995.
- EDGE, D.: "Quantitative measures of communication in science", *History of Science*, v.17 (1979) 102-134.
- EDGE, D.O.; MULKAY, M. (1976): *Astronomy transformed: the emergence of radio-astronomy in Britain*, New York: Wiley-Interscience, 1976.
- EGGHE, L. (1990): "New Bradfordian laws equivalent with old Lotka laws, evolving form a source-item duality argument", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 79-96.

- EGGHE, L. (1988): "The Fussler sampling technique for populations with a discrete or a continuous distributions of thicknesses", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 65-74.
- EGGHE, L.; ROUSSEAU, R. (1990b): "Elements of concentration theory", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 97-137.
- EGGHE, L.; ROUSSEAU, R. (1990a): *Introduction to Informetrics. Quantitative methods in library, documentation and Information Science*, Amsterdam: Elsevier, 1990.
- ELKANA, YEHUDA, et al. (eds.) (1978): *Towards a metric of science: the advent of science indicators*, New York: John Wiley & Sons, 1978.
- ESTIVALS, R. [1993]: "Bibliométrie et bibliologie (histoire, sociologie et prévision intellectuelles)", en ESTIVALS, R. (dir): *Les sciences de l'écrit. Encyclopédie Internationale de Bibliologie*, Paris: AIB, 1993.
- FERNÁNDEZ DE LUCIO, I.; GUTIÉRREZ, A.; BENLLOCH, E.: "Presente y futuro de la dimensión regional de la ciencia y la tecnología", *Arbor*, n.584 (1994 ago) 107-143.
- FLECK, L.: *Genesis and Development of a Scientific Fact*, Chicago: Univ. of Chicago Press, 1979.
- FOX, M.F.: "Publication productivity among scientists: a critical review", *Social Studies of Science*, v.13 (1983) 285-305.
- FRAME, J.D.; CARPENTER, M.P. (1979): "International research collaboration", *Social Studies of Science*, 9 (1979) 481-497.
- FRANKLIN, J.J. (1987): *Introduction to the policy application of bibliometric model data: highlights of Australian research activity in the 1983-1984 model*, Wollongong: Centre for Technology and Social Change, 1987.
- FRANKLIN, J.J.; JOHNSTON, R. [1988]: "Co-citation bibliometric modelling as a tool for S&T policy and R&D management: Issues, applications and developments", en: VAN RAAN, A.F.J. (ed.): *Handbook of quantitative studies of science and technology*, Amsterdam: Elsevier, 1988, pp. 325-389.
- FREEMAN, C. (1969): *Measurement of output of research and experimental development: A review paper*, UNESCO Document: COM/CONF 22/8, Paris, 1969.
- FREEMAN, R.B. (1979): "Indicators of the impact of R&D on the economy", *Scientometrics*, v.6, n.5-6 (1980) 375-385.
- FROST, C.O. (1979): "The use of citations in literature research: A preliminary classification of citation functions", *Library Quarterly*, 49 (1979) 399-414.
- GARFIELD, E. (1955): "Citation indexes for science", *Science*, v.122 (1955) 108-111.
- GARFIELD, E. (1963): "Citation indexes in sociological and historical research", *American Documentation*, v.14, n.3 (1963) 29-31.
- GARFIELD, E. (1970): "Citation indexing for studying science", *Nature*, v.227 (1970 aug) 669-671
- GARFIELD, E. (1972): "Citation analysis as a tool in journal evaluation", *Science*, v.178 (1972 nov) 471-479.
- GARFIELD, E. (1976): "Is the ratio between number of citations and publications cited a true constant?", *Current Contents*, (1976 feb). (9 de febrero)
- GARFIELD, E. (1978): "Is citation analysis a legitimate evaluation tool?", *Scientometrics*, v.1, n.4 (1979) 359-375.

- GARFIELD, E. (1979a): *Citation indexing: its theory and application in science, technology and humanities*, New York: Wiley, 1979. (tb ISI Press)
- GARFIELD, E. (1979b): *Essays of an Information Scientist*, Philadelphia: Institute for Scientific Information, 1979.
- GARFIELD, E. (1986): "Refereeing and peer review. Part 4" *Current Contents*, v.48 (1986) 3-9.
- GARFIELD, E. (1990): "How ISI selects journals for coverage: quantitative and qualitative considerations", *Current Contents*, n.22 (1990) 5-13.
- GARFIELD, E.; WELLJAMS-DOROF, A. (1992): "Citation data: their use as quantitative indicators for science and technology evaluation and policy-making", *Science and Public Policy*, v.19, n.5, (1992 oct) 321-327.
- GARVEY, W.D. (1979): *Communication: the essence of science*, NY: Pergamon Press, 1979.
- GEISON, G.L. (1981): "Scientific change, emerging specialties and research schools", *History of Science*, v.19 (1981) 20-40.
- GELL-MANN, M. (1994): *El quark y el jaguar*, Barcelona: Tusquets, 1995.
- GELLER, N.L. (1978): "On the citation influence methodology of Pinski and Narin", *Information Processing and Management*, v.14 (1978) 93-95.
- GIERYN, T.; MERTON, R.K. (1978): "The sociological study of scientific specialties", *Social Studies of Science*, (1978) 257-261.
- GILBERT, G.N.; WOOLGAR, S. (1974): "The quantitative science: An examination of the literature", *Science Studies*, v.4 (1974) 279-294.
- GILBERT, G.N.; MULKAY, M. (1984): *Opening Pandora box: A Sociological Analysis of Scientists Discourse*, Cambridge: Cambridge University Press, 1984.
- GILBERT, G.N. (1977): "Referencing as persuasion", *Social Studies of Science*, v.7 (1977) 113-122.
- GILBERT, G.N. (1978): "Measuring the growth of science", *Scientometrics*, v.1 (1978) 9.
- GLÄNZEL, W.; SCHUBERT, A. (1990): "The acumulative advantage function. A mathematical formulation based on conditional expectations and its application to scientometric distributions", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 139-147.
- GLÄNZEL, W.; SCHUBERT, A. (1988): "Theoretical and empirical studies of the tails of scientometric distributions", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 75-83.
- GOLDSMITH, M.; MACKAY, A.L. (1964): *The science of science*, London: Souvenir Press, 1964.
- GOLDSTONE, J.A. (1979): "A deductive explanation of the Matthew effect in science", *Social Studies of Science*, v.9 (1979) 385-391.
- GÓMEZ, I.; SANZ, E.; MÉNDEZ, A. (1989): "Utility of bibliometric analysis for research policy: A case study of Spanish research in neuroscience", *Research Policy*, 19 (1990) 457-466.
- GÓMEZ, I.; FERNÁNDEZ, M.T.; ZULUETA, M.A.; CAMÍ, J. (1994): "Analysis of biomedical research in Spain", *Research Policy*, (1994).

- GONZÁLEZ BLASCO, P.; JIMÉNEZ BLANCO, J. (1979): "La producción científica española de 1965 a 1970", en: GONZÁLEZ BLASCO, P.; JIMÉNEZ BLANCO, J.; LÓPEZ PIÑERO, J.M. *Historia y sociología de la ciencia en España*, Madrid: Alianza, 1979.
- GORDON, M.D. [1980]: "A critical reassessment of inferred relations between multiple authorship, scientific collaboration, the production of papers and their acceptance for publication", *Scientometrics*, v.2 (1980) 193-201.
- GORDON, M.D. (1981): "Citation ranking versus subjective evaluation in the determination of journal hierarchies in the social sciences", *Journal of the American Society for Information Science*, (1982 ene) 55-57.
- GORDON, M.D. (1984): "How authors select journals: a test of the reward maximization model of submission behavior", *Social Studies of Science*, v.14 (1984) 27-43.
- GRIFFITH, B.C.; SMALL, H. (1974): "The Structure of Scientific Literatures II: Toward a Macro- and Micro-Structure of Science", *Science Studies*, v.4 (1974) 339-63.
- GRIFFITH, B.C. (1988): "Exact fits in bibliometrics: some tools and results", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 85-95.
- GRIFFITH, B.C.; MULLINS, N.C. (1972): "Coherent social groups in scientific change", *Science*, 177 (1972 sep) 959-964.
- GRIGG, SHEEHAN (1991): *Evaluating research: the role of the performance indicators*, Queensland, Australia: Univ. of Queensland, Office of the Academic Director of Research, 1991.
- GUSFIELD, J. (1976): "The literary rethoric of science", *American Sociological Review*, v.41 (1976) 16-34.
- HAGSTROM, W.O. (1965): *The Scientific Community*, New York: Basic Books, 1965.
- HAGSTROM, W.O. (1971): "Inputs, outputs, and the prestige of university science departments", *Sociology of Education*, v.44 (1971) 375-397.
- HAGSTROM, W.O. (1974): "Competition in science", *American Sociological Review*, v.39 (1974) 1-18.
- HAITUN, S.D. (1990): "The problem of indicator-latent relationship in metric models. Part 1: Statement and general solution", *Scientometrics*, v.23, n.2 (1992) 335-351.
- HAITUN, S.D. (1990): "The problem of indicator-latent relationship in metric models. Part 2: Metric models with a priori latent assignment", *Scientometrics*, v.24, n.2 (1992) 221-235.
- HAITUN, S.D. (1990): "Criteria of Gaussian / non-Gaussian nature of distributions and populations", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 149-161.
- HAMERS, L.; HEMERYCK, Y.; HERWEYERS, G.; JANSSEN, M.; KETERS, H.; ROUSSEAU, R.; VANHOUTTE, A. (1989): "Similarity measures in scientometric research: the Jaccard index versus Salton's cosine formula", *Information Processing and Management*, v.25 (1989) 315-318.
- HARGENS, L.L.; MULLINS, N.C.; HECHT, P.K.: "Research areas and stratification processes in science", *Social Studies of Science*, v.10 (1980) 55-74.
- HARNARD, S. (ed.). *Peer commentary on peer review*, Cambridge: Cambridge Univ. Press, 1982.

- HARNARD, S. (1985): "Rational disagreement in peer review" *Science, Technology and Human Values*, 10 (1985) 55-62.
- HARMON, L.R. (1961): "The high school backgrounds of science doctorates", *Science*, 133 (1961) 679.
- HEALEY, P.; ROTHMAN, H.; HOCH, P.K. (1986): "An experiment in science mapping for research planning", *Research Policy*, v.15 (1986) 213-251.
- HEMLIN, S. (1992): "Scientific quality in the eyes of the scientist. A questionnaire study", *Scientometrics*, v.27, n.1 (1993) 3-18.
- HEMLIN, S.; MONTGOMERY, H. (1990): "Scientists' conceptions of scientific quality", *Science Studies*, v.1 (1990) 73-81.
- HICKS, D. (1987): "Limitations of co-citation analysis as a tool for science policy", *Social Studies of Science*, v.17 (1987) 295-316.
- HICKS, D. [1991]: "A cautionary view of co-citation analysis", *Research evaluation*, v.1, n.1 (1991) 31-36.
- HIRST, G. (1978): "Discipline impact factors-method for determining core journal lists", *Journal of the American Society for Information Science*, v.24, n.4 (1978) 171-2.
- HOLBROOK, J.A.D. (1992): "Why measure science?", *Science and Public Policy*, v.19, n.5 (1992 oct) 262-266.
- HOLBROOK, J.A.D. (1992): "Basic indicators of scientific and technological performance", *Science and Public Policy*, v.19, n.5 (1992 oct) 267-273.
- HULL, D. (1988): *Science as a process*, The University of Chicago Press, Chicago, 1988.
- HUSTOPECKY, J.; VLACHY, J. (1978): "Identifying a set of inequality measures for science studies", *Scientometrics*, v.8 (1978) 85-98.
- INHABER, H.; ALVO, M. (1978): "World science as an input-output system", *Scientometrics*, v.1, n.1 (1978) 43-64.
- IRVINE, J.; MARTIN, B.R. (1983): "Assessing basic research: The case of the Isaac Newton telescope", *Social Studies of Science*, v.13 (1983) 49-86.
- IRVINE, J.; MARTIN, B.R. (1985): "Is Britain spending enough on science?", *Nature*, v.323 (1985) 591-594.
- IRVINE, J.; MARTIN, B.R. (1984): "Evaluating big science: CERN's past performance and future prospects", *Scientometrics*, v.7 n.3-6 (1985) 281-308.
- IRVINE, J.; MARTIN, B.R.; ISARD, P. (1990): *Investing in the future. An International comparison of government fundign of academic research*, Worcester; Vermont: Billing & Sons Ltd; Edward Elgar Publishing Ltd., 1990.
- JASANOFF, S.; MARKLE, G.; PETERSEN, J.C.; PINCH, T. (eds.): *Handbook of sociology of science and technology studies*, Londres: Sage, 1995.
- KAPLAN, E. (1965): "The norms of citation behavior: Prolegomena to the footnote", *American Documentation*, v.16, n.3 (1965) 179-184.
- KARA-MURZA, S.G. (1986): "The study of journal publications in order to analyse the evolution of the cognitive structure of scientific disciplines", *Science of Science*, v.6, n.3-4, (1986) 231-245.
- KESSLER, M.M. (1963): "Bibliographic coupling between scientific papers", *American Documentation*, 14 (1963) 10-25.

- KESSLER, M.M. (1963): "An experimental study of bibliographic coupling between technical papers", *IRE Transactions PGIT*, IT-9, 49.
- KING, J. (1987): "A review of bibliometric indicators and other science indicators and their role in research evaluation", *Journal of Information Science*, v.13 (1987) 261-276.
- KITCHER, P. (1993): *The advancement of science. Science without legend, objectivity without illusions*, New York: Oxford Univ. Press, 1993
- KNORR-CETINA, D.; CICOUREL, A.V.(eds.): (1981) *Advances in social theory and methodology. Toward an integration of micro- and macro-sociologies*, London: Routledge & Kegan Paul, 1981.
- KNORR-CETINA, K. (1981): *The manufacture of knowledge: An essay of constructivism and contextual nature of science*, Oxford: Pergamon Press, 1981.
- KNORR-CETINA, K. (1982): "Scientific communities or transepistemic arenas of research? a critique of quasi economic models of science", *Social Studies of Science*, v.12 (1982) 101-130.
- KNORR-CETINA, D.; MULKAY, M. (eds.) (1983): *Science observed*, Londres: Sage, 1983.
- KRACKHARDT, D. (1987): "Cognitive Social Structures", *Social Networks*, v.9 (1987) 109-134.
- KRAUZE, T.K.; MCGINNIS, R. (1978): "A matrix analysis of scientific specialties and careers in science", *Scientometrics*, v.1, n.5-6 (1979) 419-444.
- KRETSCHMER, H. (1990): "Social stratification of authors revealed from the co-authorship network", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 193-209.
- KRONICK, D.A. (1976): *History of the scientific technical periodicals: The origins and development of the scientific and technical press 1665-1790*, Metuchen, New Jersey: Scribner, 1976.
- KUHN, T.S. (1962): *La estructura de las revoluciones científicas*, México: FCE, 1995 (v.o.: *The structure of scientific revolutions*, Chicago: Univ. of Chicago Press, 1962).
- KUHN, T.S. (1977): *La tensión esencial*, México: FCE, 1982 (v.o.: *The essential tension*, Chicago: Univ. Chicago Press, 1977).
- KYVIK, S. [1989]: "Productivity difference, fields of learning, Lotka's law" *Scientometrics*, v.15 (1989) 13-25.
- LA FOLLETE, M. (ed.) (1982): *Quality in science*, Cambridge, MA: MIT Press, 1982.
- LATOUR, B. (1987): *Science in Action*, Milton Keynes Bucks.: Open University Press, 1987.
- LATOUR, BRUNO; WOOLGAR, S. (1979): *Laboratory life: the social construction of scientific facts*, Beverly Hills, CA: Sage, 1979. (London: Sage, 1986) [trad. esp.: *La vida en el laboratorio. La construcción de los hechos científicos*, Madrid: Alianza, 1995].
- LAUDAN, L. (1989): "Normative naturalism", *Philosophy of Science*, v.57 (1990) 44-59.
- LAUDAN, L. (1977): *Progress and its problems*, Berkeley: Univ of California Press, 1977.
- LAW, J. (1986) "The heterogeneity of texts", en: CALLON, M.; LAW, J.; RIP, A. (eds.): *Mapping the dynamics of science and technology*, London: Macmillan Press, 1986, 67-83.
- LAWANI, S. M. (1977): "Citation analysis and the quality of scientific productivity", *Bioscience*, v.27, n.1 (1977) 26-31.

- LAWANI, S. M. (1984): "Some bibliometric correlates of quality in scientific research", *Scientometrics*, v.9, n.1-2 (1986) 13-25.
- LAWANI, S. M.; BAYER, A.E. (1982): "Validity of citation criteria for assessing the influence of scientific publications: new evidence with peer assessment", *Journal of the American Society for Information Science*, v.34, n.1 (1983) 59-66.
- LAWANI, S. M. (1982): "On the heterogeneity and classification of author self-citations", *Journal of the American Society for Information Science*, v.33 (1982) 281-.
- LEIMKHULER, F.F. (1988): "On bibliometric modeling", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 97-104.
- LEMAINE, G.; MACLEOD, R.; MULKAY, M.; WEINGART, P. (1976) (eds.): *Perspectives on the emergence of scientific disciplines*, Paris: Mouton; Chicago: Aldine, 1976.
- LEYDESDORFF, L.; ZAAL, R. (1988): "Co-words and citations relations between documents sets and environments", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 105-119.
- LEYDESDORFF, L. (1989), "Words and Co-Words as Indicators of Intellectual Organization", *Research Policy*, 18 (1989), 209-23.
- LEYDESDORFF, L.; VAN DER SCHAAAR, P. (1987): "The use of scientometric methods for evaluating National Research Programmes", *Science & Technology Studies*, v.5, n.1 (1987) 22-31.
- LEYDESDORFF, L. (1992): "A Validation Study of 'Leximappe'", *Scientometrics*, v.25 (1992), 295-312.
- LEYDESDORFF, L. (1993): "The generation of aggregated journal-journal citation maps on the basis of the CD-ROM version of the Science Citation Index", *Scientometrics*, v.31, n.1 (1994) 59-84.
- LEYDESDORFF, L.; AMSTERDAMSKA, O. [1990]: "Dimensions of citation analysis", *Science, Technology & Human Values*, v.15, n.3 (1990) 305-335.
- LEYDESDORFF, L. (1985): "The development of frames of reference", *Scientometrics*, v.9, n.3-4 (1986) 103-125.
- LEYDESDORFF, L. (1987): "Towards a theory of citation?", *Scientometrics*, v.12, n.5-6 (1987) 305-309.
- LEYDESDORFF, L. (1988): "Problems with the 'measurement' of national scientific performance", *Science and Public Policy*, v.15, n.3 (1988) 149-152.
- LEYDESDORFF, L. (1988): "The relations between qualitative theory and scientometric methods in science and technology studies. Introduction to the topical issue", *Scientometrics*, v.15 n.5-6 (1989) 337-347.
- LEYDESDORFF, L.; COZZENS, S. (1992): "The delineation of specialties in terms of journals using the dynamic journal set of the SCI", *Scientometrics*, v.26, n.1 (1993) 135-156.
- LEYDESDORFF, L.; COZZENS, S.; VAN DEN BESSELAAR, P. (1993): "Tracking areas of strategic importance using scientometric journal mappings", *Research Policy*, v.23 (1994) 217-229.
- LEYDESDORFF, L. (1987): "Various methods for the mapping of science", *Scientometrics*, v.11 (1987) 295-324.
- LINDSAY, D. (1978): *The scientific publication system in social science*, San Francisco, CA.: Jossey-Bass, 1978.

- LINDSEY, D.: "Production and citation measures in the sociology of science: The problem of multiple authorship", *Social Studies of Science*, v.10, n.2 (1980) 145-162.
- LITTLE, A.E.; HARRIS, R.M.; NICHOLLS, P.T. (1990): "Text to reference ratios in scientific journals", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 211-216.
- LONG, J.S.; MCGINNIS, R. (1981): "Organizational context and scientific productivity", *American Sociological Review*, v.46 (1981 ago) 422-442.
- LÓPEZ PIÑERO, J.M. (1972): *El análisis estadístico y sociométrico de la literatura científica*, Valencia: Centro de Documentación e Informática Médica, 1972
- LÓPEZ PIÑERO, J.M. (1972): "La obra de Price y el análisis estadístico y sociométrico de la literatura científica", estudio preliminar a la traducción española de PRICE, D.J. DE S. (1963): *Hacia una ciencia de la ciencia*, Barcelona: Ariel, 1973, 7-19.
- LÓPEZ PIÑERO, J.M.; TERRADA, M.L. (1991): "La producción científica española y su posición en la comunidad internacional", en *España*. (vol.: *Ciencia*), Madrid: Espasa Calpe, 1991, 73-109.
- LOTKA, A. (1926): "The frequency distribution of scientific productivity", *Journal of the Washington Academy of Sciences*, v.16, n.12 (1926 jun) 317-323.
- LUUKKONEN, T. (1990): "Citations in the rethorical, reward and communication systems of science", *Acta Universitatis Tamperensis*, ser A, v.285, Tampere: University, 1990 (Academic dissertation).
- LUUKKONEN, T.; TIJSSSEN, R.W.J.; PERSSON, O.; SIVERTSEN, G. (1993): "The measurement of international scientific collaboration", *Scientometrics*, v.28, n.1 (1993) 15-36.
- LUUKKONEN, T. [1991]: "Citation indicators and peer review: their time-scales, criteria of evaluation, and biases", *Research Evaluation*, v.1, n.1 (1991) 21-30.
- LUUKKONEN, T.; PERSSON, O.; SIVERTSEN, G. (1993): "Understanding patterns of international scientific collaboration": *Science, Technology and Human Values*, v.17, n.1 (1992) 101-126.
- MACHLUP, F. (1962): *The production and distribution of knowledge in the United States*, Princeton: Princeton Univ. Press, 1962.
- MACÍAS-CHAPULA, C.A. (1990): "Production and dissemination of the Mexican biomedical journals, with some considerations of the Latin American / Caribbean region", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 217-228.
- MACKENZIE, (1981): "Interests, positivism and history", *Social Studies of Science*, v.11 (1981) 498-504.
- MACROBERTS, M.H.; MACROBERTS, B.R. (1986b): "Quantitative measures of communication in science: a study of the formal level", *Social Studies of Science*, v.16 (1986) 151-172.
- MACROBERTS, M.H.; MACROBERTS, B.R. (1986a): "Testing the Ortega hypothesis: Facts and artifacts", *Scientometrics*, v.12, n5-6 (1987) 293-295.
- MALTRÁS, B.; QUINTANILLA, M.A.; VEGA, J. (1995): "Diez años de política comunitaria en ciencia y tecnología", *Mundo Científico*, n.163 (1995 dic) 1090-1099.
- MALTRÁS, B.; QUINTANILLA, M.A. (1992): *Producción científica española, 1981-89*, Madrid: Consejo Superior de Investigaciones Científicas, 1992.

- MALTRÁS, B.; QUINTANILLA, M.A. (1994): *Indicadores de la producción científica: España, 1986-91*, Madrid: Consejo Superior de Investigaciones Científicas, 1995.
- MALTRÁS, B.; VIDAL, J. (1993): *Informe basado en indicadores bibliométricos sobre la producción científica de las Universidades españolas, 1986-91*. Informe encargado por el Consejo de Universidades.
- MALTRÁS, B.; VEGA, J.; QUINTANILLA, M.A. (1995): "Measuring multinational cooperation in science & technology: Different methods applied to the European Framework Programs", en: Koenig, M.E.D.; Brookstein, A. (eds.) *Proceedings of the Fifth Biennial Conference of International Society for Scientometrics and Informetrics*, Chicago: Learned Information Inc (Medford, N.J), 1995, 303-312.
- MANTEN, A.A. (1979): "Publication of scientific information is not identical with communication", *Scientometrics*, v.2, n.4 (1980) 303-308.
- MARGOLIS, J. (1967) "Citation indexing and evaluation of scientific papers", *Science*, 155 (1967) 1213-1219.
- MARSHAKOVA, I. (1973): "System of document connections based in references" (en ruso), *Nauchno Tekhnicheskaya Informatsiya*, s.2, 6 (1973) 3-8.
- MARTIN, B.R.; IRVINE, J. (1980): "Assessing basic research: Some partial indicators of scientific progress in radio astronomy", *Research Policy*, 12 (1983) 61-90.
- MARTIN, B.R.; IRVINE, J.; CROUCH, D. (1985): *Science Indicators for Research Policy: A Bibliometric Analysis of Ocean Currents and Protein Crystallography*, SPRU, Occasional Paper Series N° 23, University of Sussex, 1995.
- MARTIN, B.R.; IRVINE, J. (1984): "Research evaluation: Why? How?", SPRU, (1984 sep.)
- MARTIN, B.R.; IRVINE, J. (1992): "Trends in government spending on academic and related research. and international comparison", *Science and Public Policy*, v.19, n.4 (1992 oct) 311-319.
- MARTIN, B.R.; IRVINE, J.; NARIN, F.; STERRIT, C.; STEVENS, K. (1990): "Recent trends in output and impact of British science", *Science and Public Policy*, v.15 (1988) 14-26.
- MARTON, J. (1983): "Causes for low and high citation potentials in science: citation analysis of biochemistry and plant physiology journals", *Journal of the American Society for Information Science*, v.34 (1983) 244-246.
- MARTYN, J.; GILCHRIST, A. (1968): *An evaluation of British scientific periodicals*, London: Aslib (Aslib Occasional Publ. no.1), 1968.
- MAZLISH, B. [1982]: "The quality of 'The quality of science': An evaluation", en: LA FOLLETE, M.C. (ed.): *Quality in Science*, Cambridge, MA; London: MIT Press, 1982, pp. 48-67.
- MCALLISTER, P.P.; ANDERSON, R.C.; NARIN, F. (1979): "Comparison of peer and citation assessment of the influence of scientific journals", *Journal of the American Society for Information Science*, (1980 may) 147-152.
- MCREYNOLDS, P.: "Reliability of rating of research papers", *American Psychologist*, v.26 (1971) 400-401.
- MENARD, H.W. (1971): *Science: Growth and Change*, Cambridge: Harvard Univ. Press, 1971.
- MÉNDEZ, A.; GÓMEZ, I. (1988): "A comparison of citation classics in three fields of science", *Scientometrics*, v.15, n.5-6 (1989) 621-631.

- MÉNDEZ, A.; GÓMEZ, I. (1992): "Some indicators for assessing research performance without citations", *Scientometrics*, v.26, n.1 (1993) 157-167.
- MERTON, R.K. (1967): "El efecto Mateo en la ciencia", en R.K. Merton (1973) 554-578...publicado originalmente en *Science*, 159, n.3810 (1968) 56-63.
- MERTON, R.K. (1957): "Las prioridades en los descubrimientos científicos", en Merton, R.K. (1973). [publicado originalmente en *American Sociological Review*, n.6 (1957 nov) 635-659].
- MERTON, R.K. (1960): "'Reconocimiento" y "Excelencia": Ambigüedades instructivas", 531-553. en: R.K. Merton (1973): *La sociología de la ciencia*, vol.2, Madrid: Alianza, 1977.
- MERTON, R.K. (1968): "Las pautas de conducta de los científicos", en Merton, R.K. (1973), 423-443. (publicado originalmente en *American Sociological Review*, 58 (1969), 1-23).
- MERTON, R.K. (1942): "La estructura normativa de la ciencia", en: R.K. Merton (1973): *La sociología de la ciencia*, vol.2, Madrid: Alianza, 1977, 355-368. (publicado originalmente como: "Science and Technology in a democratic order", *Journal of Legal and Political Sociology*, 1 (1942) 115-126).
- MILLER, R. (1992): "The influence of primary task on R&D laboratory evaluation: a comparative bibliometric analysis", *R&D Management*, v.22, n.1 (1992) 3-20.
- MITROFF, I. (1974): "Norms and counter-norms in a select group of the Apollo moon scientist: A case study of the ambivalence of scientists", *American Sociological Review*, n.39 (1974) 579-595.
- MITTERMEIER, R.; KNORR, K.D. (1979): "Scientific productivity and accumulative advantage: a thesis reassessed in the light of international data" *R&D Management*, v.9 (1979) 235-239.
- MOED, H.F. (1988): "The use of on-line databases for bibliometric analysis", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 133-146.
- MOED, H.F.; BURGER, W.J.M.; FRANKFORT, J.G.; VAN RAAN, A.F.J. (1984a): "The application of bibliometric indicators: important field- and time-dependent factors to be considered", *Scientometrics*, v.8, n.3-4 (1985) 177-203.
- MOED, H.F.; BURGER, W.J.M.; FRANKFORT, J.G.; VAN RAAN, A.F.J. (1985): "The use of bibliometric data as tools for university research policy", *Workshop on science and technology indicators in the Higher Education sector*, (1985 jun) Paris: OECD (Directorate for Science, Technology and Industry), 15 pgs.; DSTI / SPR / 85.24 / 15.
- MOED, H.F.; VAN RAAN, A.F.J. (1988): "Indicators of research performance: Applications in university research policy", en: VAN RAAN, A.F.J. (ed.): *Handbook of quantitative studies of science and technology*, Amsterdam: Elsevier, 1988, pp. 177-192.
- MOED, H.F.; BURGER, W.J.M.; FRANKFORT, J.G.; VAN RAAN, A.F.J. (1983): *On the measurement of research performance: the use of bibliometric indicators*, Leiden: University, Science Studies Unit, 1983. (ISBN 90-9000552-8)
- MOED, H.F. (1989): *The use of bibliometric indicators for the assessment of research performance in the natural and life sciences*, Leiden: DSWO Press, 1989.

- MOED, H.F.; VAN RAAN, A.F.J. (1985): "Critical remarks on Irvine and Martin's methodology for evaluating scientific performance", *Social Studies of Science*, v.15 (1985) 539-547.
- MOED, H.F.; BURGER, W.J.M.; FRANKFORT, J.G.; VAN RAAN, A.F.J. (1984b): "The use of bibliometric data for the measurement of university research performance", *Research Policy*, v.14 (1985) 131-149.
- MOMBERS, C.; VAN HEERINGEN, A.; VAN VENETIE, R.; LE PAIR, C. (1984): "Displaying strengths and weaknesses in national R&D performance through document co-citation", *Scientometrics*, v.7, n.3-6 (1985) 341-355.
- MORAVCSIK, M.J. [1973]: "Measures of scientific growth", *Research Policy*, v.2 (1973) 266-275.
- MORAVCSIK, M.J. (1988): "The coverage of science in the third world: The 'Philadelphia' program", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 147-155.
- MORAVCSIK, M.J.; MURUGESAN, P. (1978): "Citation patterns in scientific revolutions" *Scientometrics*, v.1, n.2 (1979) 161-169.
- MORAVCSIK, M.J. (1977): "A progress report on the quantification of science", *Journal of the Scientific and Industrial Research*, 36 (1977) 195.
- MORAVCSIK, M.J. (1989): "¿Es posible planificar la ciencia?", *Revista Española de Documentación Científica*, v.12, n.2 (1989) 181-189.
- MORAVCSIK, M.J. (1989): "¿Cómo evaluar la ciencia y a los científicos?", *Revista Española de Documentación Científica*, v.12, n.3 (1989) 313-325.
- MORAVCSIK, M.J. (1989): "Transferencia de tecnología y ciencia", *Revista Española de Documentación Científica*, v.12, n.4 (1989) 451-461.
- MORAVCSIK, M.J.; MURUGESAN, P. (1975): "Some results on the function and quality of citations", *Social Studies of Science*, v.5 (1975) 86-92.
- MORITA-LOU, HIROKO (ed.) (1985): *Science and technology indicators for development*, London: Westview Press, 1985.
- MULKAY, M.J. (1976): "Norms and Ideology in Science", *Social Science Information*, v.15 (1976) 637-656.
- MULKAY, M.J. (1977): "Sociology of the scientific research community", en: SPIEGEL-RÖSING, I., SOLLA-PRICE, D.J. (eds.) *Science, technology and society: A cross-disciplinary structure*, London: Sage, 1977, 93-148.
- MULKAY, M.J. (1969): "Some aspects of cultural growth in the natural sciences", *Social Research*, 36 (1969) 22-52.
- MULKAY, N.; GILBERT, G.N. (1982): "What is the ultimate question? Some remarks in defence of the analysis of scientific discourse", *Social Studies of Science*, n.12 (1982) 309-319.
- MULKAY, M.; POTTER, J.; YEARLY, S. (1983): "Why an analysis of scientific discourse is needed", en: KNORR-CETINA, D.; MULKAY, M. (eds.): *Science observed*, Londres: Sage, 1983.
- MULLINS, N.C.; HARGENS, L.; HECHT, P.; KICK, E. [1977]: "The group structure of co-citation clusters: a comparative study", *American Sociological Review*, v.42 (1977) 323-331.

- MULLINS, N.C.; SNIZEK, W.; OEHLER, K. (1988): "The structural analysis of a scientific paper", en: VAN RAAN, A.F.J. (ed.): *Handbook of quantitative studies of science and technology*, Amsterdam: Elsevier, 1988, pp. 81-105.
- MURUGESAN, P.; MORAVCSIK, M. (1978): "Variation of the nature of citation measures with journal and scientific specialties", *Journal of the American Society for Information Science*, v.29, n.3 (1978) 141-148.
- NAGEL, E. (1961): *The structure of science: Problems in the logic of scientific explanation*, London: Routledge Kegan & Paul, 1961.
- NARIN, F. (1985): "Measuring the research productivity of Higher Education institutions using bibliometric techniques", *Workshop on science and technology indicators in the Higher Education sector*, (1985 jun) Paris: OECD (Directorate for Science, Technology and Industry), DSTI / SPR / 85.24 / 07 B.
- NARIN, F. CARPENTER, M.P (1980): "Bibliometric indicator series in the U.S. Science Indicators Data Base", *OECD: Science & Technology Indicators Conference*, París, 15-19 sept. 1980.
- NARIN, F. (1976a): *Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity*, Contract NSF C-627. NSF marzo 1976. Monograph: 456, NTIS Accession # PB252339/AS.
- NARIN, F.; CARPENTER, M.P. (1975): "National publication and citation comparisons", *Journal of the American Society for Information Science*, v.26, n.2 (1975) 80-93.
- NARIN, F. (1976b): *Evaluative bibliometrics*, Cherry Hill: Computer Horizons Inc., 1976.
- NARIN, F.; PINSKI, G.; GEE, H.H. (1976): "Structure of the biomedical literature", *Journal of the American Society for Information Science*, v.27 (1976) 25-45.
- NARIN, F.; NOMA, E. (1984): "Is technology becoming science?", *Scientometrics*, v.7 n.3-6 (1985) 369-381.
- NARIN, F. (1987): "Bibliometric techniques in the evaluation of research programs", *Science and Public Policy*, v.14, n.2 (1987 apr) 99-106.
- NARIN, F.; CARPENTER, M.P.; BERLT, N.C. (1972): "Interrelationships of scientific journals", *Journal of the American Society for Information Science*, v.23 (1972) 323-331.
- NEDERHOF, A.J. (1985): "Evaluating research output through life work citation counts", *Scientometrics*, v.7 (1985) 23-28.
- NEDERHOF, A.J. (1991): "Assesment of the international standing of university departments' research: A comparison of bibliometric methods", *Scientometrics*, v.24, n.3 (1992) 393-404.
- NOMA, E. (1988): "Using influence weights to evaluate the scientific importance of journals", en: VAN RAAN, A.F.J. (ed.) *Handbook of quantitative studies of science and technology*, Amsterdam: Elsevier, 1988, 677-703.
- NOMA, E. (1982): "An improved method for analyzing square scientometric transaction matrices", *Scientometrics*, v.4, n.4 (1982) 297-316.
- OCDE (Organización para la Cooperación y el Desarrollo Económicos) (1980): *Manual de Frascati. La medición de las actividades científicas y técnicas*, Madrid: CDTI, 1980.

- OCDE (1993): *Frascati Manual, 1993. The Measurement of Scientific and Technological Activities*, Paris: OCDE, 1994.
- OFFICE FOR TECHNOLOGY ASSESMENT (OTA): *Research funding as an investment. Can we measure the results?* A technical memorandum (U.S. Congress, OTA- TM- SET- 36, Washington, 1986)
- OKRASA, W. (1984): "Structure of output inequality in research units: An alternative approach to the analysis of scientific productivity", *Science of Science*, v.4, n.2 (1984) 197-227.
- OKUBO, Y.; MIQUEL, J.F.; FRIGOLETTO, L.; DORÉ, J.C. (1992): "Structure of international collaboration in science: typology of countries through multivariate techniques using a link indicator", *Scientometrics*, v.25, n.2 (1992) 321-351.
- PAVITT, K. (1984): "Science and technology indicators: eight conclusions", *Science and Public Policy*, (1984 feb) 21-24.
- PERITZ, B.C. (1988): "Bibliometric literature: A quantitative analysis", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 165-173.
- PERITZ, B.C. (1990): "A citation analysis of clinical trials; are definitive studies less cited than others?", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 243-250.
- PERSSON, O. (1986): "On-line bibliometrics. A research tool for everyman", *Scientometrics*, v.8 (1986) 69-75.
- PESTAÑA, A. (1996): "El sistema español de ciencia y técnica", *Investigación y ciencia*, n. 243 (1996 dic) 6-13.
- PETERS, H.P.F.; HARTMANN, D.; VAN RAAN, A.F.J. (1988): "Monitoring advances in chemical engineering", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 175-195.
- PHILLIMORE, A.J. (1988): "University research performance indicators in practice: The University Grants Committee's evaluation of British universities, 1985-86", *Research Policy*, v.18 (1989) 255-271.
- PINSKY, G.; NARIN, F. (1976): "Citation influence for journal aggregates of scientific publications: theory, with applications to the literature of physics", *Information Processing & Management*, v.12 n.5 (1976) 197-312.
- POLANYI, M. (1964): *Personal Knowledge*, Chicago: University of Chicago Press, 1964.
- POLANYI, M. (1962): "The Republic of Science, its political and economic theory", *Minerva*, v.1 (1962) 54-73.
- POMIAN, J. (1988): "Reformulation des questions dans l'interrogation des bases des données bibliographiques: système LEXIQUEST", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 197-208.
- PORTER, A.L. (1977): "Citation analysis: Queries and caveats", *Social Studies of Science*, v.7 (1977) 257-267.
- PORTER, A.L.; ROSSINI, F.A. (1985): "Peer review of interdisciplinary research proposals", *Science, Technology & Human Values*, v.10, n.3 (1985) 34-38.
- PRESSER, S. [1980]: "Collaboration and the quality of research", *Social Studies of Science*, v.10 (1980) 95-101.

- PRICE, D.J.DE S. (1950): "Quantitative measures of the development of science", *Archives Internationales d'Histoire des Sciences*, n.14 (1951) 85-93.
- PRICE, D.J.DE S. (1976): "A general theory of bibliometric and other cumulative advantage processes", *Journal of the American Society for Information Science*, v.27 (1976) 292-306.
- PRICE, D.J.DE S. (1963): *Hacia una ciencia de la ciencia*, Barcelona: Ariel, 1973 (trad. de J.M. Piñero de la edición original: *Little Science, Big Science*, New York: Columbia Univ. Press, 1963.)
- PRICE, D.J.DE S. (1971): "Collaboration in an invisible college", *American Psychologist*, v.21 (1971) 1011-1018.
- PRICE, D.J. DE S. (1970): "Smiles at the unobstrusive", *Nature*, v.226 (1970 jun) 985 (carta al editor).
- PRICE, D.J.DE S. (1965): "Networks of scientific papers", *Science*, v.149 (1965) 510-515.
- PRICE, D.J. DE S. (1980a): "The analysis of square matrices of scientometric transactions", *Scientometrics*, v.3 (1981) 55-63.
- PRICE, D.J.DE S. (1980b): "The analysis of scientometric matrices for policy implications", *Scientometrics*, v.3 (1981) 47-54.
- PRICE, D.J.DE S. (1981): "Multiple authorship", *Science*, v.212, n.4498 (1981) 986- (carta al editor)
- PRICE, D.J.DE S. (1982): "The parallel structures of science and technology", en: BARNES, B.; EDGE, D. (eds.): *Science in context*, Milton Keynes, Open Univ. Press, 1982.
- PRPIC, K. (1994): "The socio-cognitive frameworks of scientific productivity", *Scientometrics*, v.31, n.3 (1994) 293-311.
- QUINTANILLA, M.A.; MALTRÁS, B. (1993): "La colaboración entre instituciones científicas españolas (1981-1989)", *II Jornadas de Investigación sobre las relaciones entre Ciencia, Tecnología, Economía y Sociedad, RICTES*, Salamanca, (1993 oct.) 321-332.
- QUINTANILLA, M.A.; MALTRÁS, B. (1992): "La estructura de la producción científica en España (1981-1989) y las prioridades del Plan Nacional", *Arbor*, n.554-555 (1992 feb-mar) 107-130.
- QUINTANILLA, M.A. (1992): "El sistema español de ciencia y tecnología y la política de I+D", *Arbor*, n.554-555 (1992 feb-mar) 9-29.
- QUINTANILLA, M.A. (1988): *Tecnología: Un enfoque filosófico*, Madrid: Fundesco, 1989.
- RAISIG, L.M. (1960): "Mathematical evaluation of the scientific serial", *Science*, v.131 (1960) 1417-1419.
- REICHENBACH, (1934), *Experience and prediction*, Chicago: University of Chicago Press, 1934.
- RESKIN, B.F.(1977): "Scientific productivity and the reward structure of science", *American Sociological Review*, v.42 (1977 jun) 491-504.
- RIP, A.(1988): "Mapping of science: Possibilities and limitations", en: VAN RAAN, A.F.J. (ed.): *Handbook of Quantitative Studies of Science and Technology*, Amsterdam: Elsevier, 1988, 253-273.
- RIP, A.(1981): "A cognitive approach to Science Policy", *Research Policy*, v.10 (1981) 294-311.

- RIP, A. (1986): "Mobilising resources through texts", en : CALLON, M.; LAW, J.; RIP, A. (eds.): *Mapping the dynamics of science and technology*, London: Macmillan Press, 1986, 84-99.
- RIP, J.; COURTIAL, J.P. (1983): "Co-words maps of biotechnology: an example of cognitive scientometrics", *Scientometrics*, v.6, n.6 (1984) 381-400.
- ROSENBERG, N. (1979): "Comments on 'indicators of the impact of R&D on the economy'", *Scientometrics*, v.2, n.5-6 (1980) 387-393.
- ROSENBERG, N (1992): "Progreso técnico: El análisis histórico", *Economía e Historia*, n.47 Barcelona: Oikos-tau (El mundo contemporáneo), 1992
- ROUSSEAU, R. (1988): "Citation distribution of pure mathematics journals", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 249-262.
- SALTON, G. (1988): "On the relationship between theoretical retrieval models", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 263-270.
- SALTON, G.; MCGILL, M.J. (1983): *Introduction to Modern Information Retrieval*, London: McGraw Hill, 1983.
- SALTON, G.; YU, H. (1980): "The measurement of term importance in automatic indexing", *Journal of the American Society for Information Science*, v.32 (1981) 175-186.
- SANCHO, R. (1990): "Indicadores bibliométricos utilizados en la evaluación de la ciencia y la tecnología. Revisión bibliográfica", *Revista Española de Documentación Científica*, v.13, n.3-4 (1990) 842-865.
- SCHUBERT, A.; GLÄNZEL, W.; BRAUN, T. (1989): "Scientometrics datafiles. A comprehensive set of indicators on 2649 journals and 96 countries in all major science fields and subfields. 1981-1985", *Scientometrics*, v.16, n.1-6, (1989) 3-478.
- SCHUBERT, A.; GLÄNZEL, W. (1983): "Statistical reliability of comparisons based on citation impact of scientific publications", *Scientometrics*, v.5 (1983) 59-75.
- SCHUBERT, A.; GLÄNZEL, W. (1986): "Mean response time - a new indicator of journal citation speed with application to physics journals", *Czechoslovak Journal of Physics*, b.36 (1986) 126-129.
- SCHUBERT, A.; GLÄNZEL, W.; BRAUN, T. (1986): "Relative indicators and relational charts for comparative assessment of publication output and citation impact", *Scientometrics*, v.9 (1986) 281-?.
- Science Citation Index Annual. Guide and Lists of Source Publications*, Philadelphia: Institute for Scientific Information, (anual; se han utilizado las de 1987, 1989, y 1991).
- Science Citation Index. Journal Citations Reports. A bibliometric analysis of Science Journals in the ISI Database*, Philadelphia: Institute for Scientific Information,
- Science Citation Index Annual. Guide and Lists of Source Publications*, Philadelphia: Institute for Scientific Information.
- Science Indicators*, National Science Foundation, Washington DC, 1985.
- SECRETARÍA GENERAL DEL PLAN NACIONAL DE I+D (1990): *Centros de Investigación en España*, Madrid: Centro de Publicaciones. M° de Educación y Ciencia, 1990.
- SEN, B.K.; SHAIKLENDRA, K. (1991): "Evaluation of recent scientific research output by a bibliometric method", *Scientometrics*, v.23, n.1 (1992) 31-46.

- SHADISH, W.R.; TOLLIVER, D.; GRAY, M.; GUPTA, S.K.S. (1995): "Author judgements about work they cite: Three studies from psychology journals", *Social Studies of Science*, v.25 (1995) 477-498.
- SHADISH, W. R. (1989): "The perception and evaluation of quality in science", en Golsen, B., et al., *Psychology of Science. Contributions to metascience*, Cambridge University Press, Cambridge, 1989, 383-426.
- SHILS, E. (ed.) (1968): *Criteria for scientific development: Public policy and national goals*, Cambridge, MA: MIT Press, 1968.
- SHOCKLEY, W. (1957): "On the statistics of individual variations of productivity in research laboratories", *Proceedings of the Institute of Radio Engineerings*, v.45 (1957) 279 y 1409.
- SHRUM, W. (1985): *Organized Technology: networks and Innovation in Technical Systems*, West Lafayette: Purdue Univ. Press, 1985.
- SIMON, H.A. (1955): "On a class of skew distribution functions", *Biometrika*, 42 (1955), 425-440.
- SKLAIR, L.(1972): "The political sociology of science", *The Sociological Review Monograph*, 18 (1972) 43-59.
- SMALL, H. (1982): "Citation context analysis", *Progress in communication Sciences*, v.3 (1982) 287-310.
- SMALL, H. (1978): "Cited documents as concept symbols", *Social Studies of Science*, v.8 (1978) 327-340.
- SMALL, H.; GRIFFITH, B.C. (1974): "The structure of scientific literature I: identifying and graphing specialties", *Science Studies*, v.4 (1974) 17-40.
- SMALL, H.(1976): "Structural dynamics of scientific literature", *International Classification*, v.3 n.2 (1976) 67-74.
- SMALL, H.; SWEENEY, E.; GREENLEE, E. (1985a): "Clustering the Science Citation Index using co-citations, I: a comparison of methods", *Scientometrics*, v.7 (1985) 393-409.
- SMALL, H.; SWEENEY, E.; GREENLEE, E. (1985b): "Clustering the Science Citation Index using co-citations, II: Mapping Science", *Scientometrics*, v.8 (1985) 321-340.
- SMALL, H. (1977): "A co-citation model of a scientific specialty: a longitudinal study of collagen research", *Social Studies of Science*, v.7 (1977) 139-166.
- SMALL, H.; GARFIELD, E. (1985): "The geography of science: disciplinary and national mappings", *Journal of Information Science*, v.11 (1985) 147-159.
- SMALL, H. (1973): "Co-citation in the scientific literature: a new measurement of the relationship between two documents", *Journal of the American Society for Information Science*, v.24 (1973) 265-269.
- SMALL, H.; CRANE, D. [1979]: "Specialties and disciplines in science and social science: and examination of their structure using citation indexes", *Scientometrics*, v.1 (1979) 445-461.
- SMITH, D.C.; COLLINS, P.M.D.; HICKS, D.M.; WYATT, S.(1986): "National performance in Basic Research", *Nature*, 323 (1986) 681-84.
- SMITH, L.C. (1981): "Citation analysis", *Library Trends*, v.30 n.1 (1981) 83-106.
- SONNERT, G. (1995): "What makes a good scientist?: Determinants of peer evaluation among biologists", *Social Studies of Science*, v.25 (1995) 35-55.

- SPIEGEL-RÖSING, I. (1977): "Bibliometric and content analysis", *Social Studies of Science*, 5 (1977) 97.
- SPIEGEL-RÖSING, I., SOLLA-PRICE, D.J. (eds.) (1977): *Science, technology and society*, London: Sage, 1977.
- STEPHAN, P.E.; LEVIN, S.G.(1991): "Inequality in scientific performance: Adjustment for attribution and journal impact", *Social Studies of Science*, v.21 (1991) 351-368.
- STEED, J., MULLINS, N. C. (1986): "Problems and Moves: A Methodology", Paper presented at the American Association for Advancement of Science meeting, Philadelphia, May 23, 1986.
- STORER, NORMAN (1966): *The social system of science*, N.Y.:Holt, Reinhart & Wilson, 1966.
- STUDER, K.; CHUBIN, D. (1980): *The cancer mission*, Beverly Hills: Sage, 1980.
- TABAH, A.N.; SABER, A.J. (1990): "Chaotic structures in informetrics", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 281-289.
- TAGLIACOZZO, R. (1977): "Self-citation in scientific literature", *Journal of Documentation*, v.33, n.4 (1977) 251.
- TAGLIACOZZO, R. (1967): "Citation nad citation indexes: A review", *Method. Inform. Med.*, v.6 (1967) 136-145.
- TAGUE, J. (1988): "What's the use of bibliometrics?", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 271-278.
- TAYLOR, C.W.; ELLISON, R.L. (1967): "Biographical predictors of scientific performance", *Science*, v.155 (1967 mar) 1075-1080.
- The evaluation of scientific performance*, proceedings of a Ciba Foundation Conference (Wiley, 1989).
- The Magic of numbers (1968) (editorial) *Nature*, v.217 (1968) 793-794.
- THORNE, F.C.(1977): "The citation index: another case of spurious validity", *Journal of Clinical Psychology*, v.33 (1977) 1157-1161.
- TIJSSSEN, R.J.W.; MOED, H.F. (1989): "Literature-Based Statistical Analyses of International Cooperation. An exploratory case-study of the Netherlands", en VAN RAAN, A. F. J., NEDERHOF, A. J., MOED, H. M. (eds.), *Science and Technology Indicators. Their Use in Science Policy and their Role in Science Studies*, Leiden, DSWO Press, 1989, 129-145.
- TIJSSSEN, R.J.W. (1993): "A scientometric cognitive study of neural network research: expert mental maps versus bibliometric maps", *Scientometrics*, v.28, n.1 (1993) 111-136.
- TIJSSSEN, R.J.W. (1991): "A quantitative assesment of interdisciplinary structures in science and technology: Co-classification analysis of energy research", *Research Policy*, 21 (1992) 27-44.
- TODOROV, R. (1984): "Determination of influence weigths for scientific journals", *Scientometrics*, v.6 (1984) 127-138.
- TODOROV, R. (1990): "Representing Canadian geophysics: a bibliometric approach", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 291-307.

- TODOROV, R.; GLÄNZEL, W. (1988): "Journal citation measures: a concise review", *Journal of Information Science*, v.14 (1988) 47-56.
- TOMER, C. (1986): "A statistical assessment of two measures of citation: the impact factor and the immediacy index", *Information Processing & Management*, v.22 (1986) 251-258.
- TURNER, W.A.; CHARTRON, G.; MICHELET, B. (1985): *Describing scientific and technological problem networks using manually and automatically indexed full-text databases: some co-word analysis techniques*, Paris: OECD, 1985 (Report nr. DSTI/PRR/85.24/08A).
- TYSSSEN, R.J.W.; DE LEUW, J.; VAN RAAN, A.F.J. (1988): "A method for mapping bibliometric relations based on field-classifications and citations of articles", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 279-292.
- U.S. NATIONAL SCIENCE BOARD. (1985): *Science Indicators: The 1985 Report*, Washington, DC: Superintendent of Documents, 1986 and earlier editions of this biennial series.
- VAN STEIJN, F.A.; RIP, A. (1988). "The role of trade literature in the communication system", *Scientometrics*, v.13 (1988) 81-91.
- VAN RAAN, A.F.J. (1988): "Impact of research performance as measured by citations: A new model", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 87/88*, Amsterdam: Elsevier, 1988, 293-299.
- VAN RAAN, A.F.J.; HARTMANN, D. (1986): "The comparative impact of scientific publications and journals: methods of measurement and graphical display", *Scientometrics*, v.11, n.5-6 (1987) 325-331.
- VELHO, L. (1985): "The 'meaning' of citation in the context of a scientifically peripheral country", *Scientometrics*, v.9, n.1-2 (1986) 71-89.
- VIDAL, F.J. (1995): *La caja gris. Microanálisis de la actividad investigadora, su gestión y evaluación en una institución universitaria. Estudio del caso de la Universidad de León*. Salamanca, 1995 (tesis doctoral)
- VIDAL, F.J.; VILLARROEL, R. (1995): "The dynamics of research groups: representation and interpretation problems of collaboration analyses", *Proceedings of the 5th. International Conference on Scientometrics and Informetrics*, Chicago, 1995 jun.
- VINKLER, P. (1986): "Evaluation of some methods for the relative assesment of scientific publications", *Scientometrics*, v.10, n.3-4 (1986) 157-177.
- VINKLER, P. (1990): "Magic triangle for three relative impact indicators", *Scientometrics*, v.21, n.1 (1991) 143-146.
- VINKLER, P. (1987A): "A quasi-quantitative citation model", *Scientometrics*, v.12, (1988) 251-?
- VINKLER, P. (1987B): "An attempt of surveying and classifying bibliometric indicators for scientometric purposes", *Scientometrics*, v.13, n.5-6 (1988) 239-259.
- VINKLER, P. (1990): "Bibliometric analysis of publication activity of a scientific research institute", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 309-334.
- VINKLER, P. (1992): "Research contribution, authorship and team cooperativeness", *Scientometrics*, v.26, n.1 (1993) 213-230.

- VIRGO, J.A. (1977): "A statistical procedure for evaluating the importance of scientific papers", *Library Quarterly*, 47 (1977) 415-?.
- WEIMER, W.B. (1977): "Science as a Rhetorical Transaction. Toward a Nonjustificational Concept of Rhetoric", *Philosophy and Rhetoric*, v.10 (1977) 1-29.
- WEINBERG, A.M. (1961): "Impact of large-scale science on the United States", *Science*, v.134, n.3473 (1961 jul) 161-164.
- WEINBERG, A.M. (1963): "Criteria for scientific choice", *Minerva*, v.1 (1963) 159-171.
- WEINBERG, B. (1974): "Bibliographic coupling: a review", *Information Storage and Retrieval*, v.10 (1974) 189-196.
- WEINSTOCK, M. (1971): "Citation indexes", *Encyclopedia of Library and Information Science*, v.5, New York: Marcel Dekker, 1971.
- WHITE, H.D.; GRIFITH, B.C. (1980): "Author co-citation: a literature measure of intellectual structure", *Journal of the American Society for Information Science*, v.32 (1981) 163-171.
- WHITE, A.; HERNÁNDEZ, N.R. (1990): "Increasing field complexity revealed through article title analysis", *Journal of the American Society for Information Science*, v.42, n.10 (1991) 731-734.
- WHITLEY, R.D. (1970a): "The formal communication system in science", *The Sociological Review*, m.16 (1970 sep) 163-179.
- WHITLEY, R.D. (1970b): "The operation of science journals: Two case studies in British social science", *Sociological Review*, n.18 (1970 jul) 241-258.
- WILSON, F.L.; McGRATH, W.E. (1990): "Cluster analysis of title overlap in twenty-one library collections in Western New York", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 335-353.
- WHITTAKER, J. (1989) "Creativity and Conformity in Science: Titles, Keywords and Co-word Analysis", *Social Studies of Science*, 19 (1989), 473-96.
- WOLFRAM, D.; CHU, C.M.; XIN LU (1990): "Growth of knowledge: bibliometric analysis using on line database data", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 355-372.
- YABLONSKY, A.I. (1980): "On fundamental regularities of the distribution of scientific productivity", *Scientometrics*, v.2 (1980) 3-34.
- YITZHAKI, M.; BEN-TAMAR, D. (1990): "Multiple authorship in biochemistry and other fields; a case study of the Journal of Biological Chemistry throughout 1905-1988", en: EGGHE, L.; ROUSSEAU, R. (ed.) *Informetrics 89/90*, Amsterdam: Elsevier, 1990, 373-389.
- ZIMAN, J.M. [1983]: "The collectivization of science", *Proceedings of the Royal Society*, B219, London: 1983, 1-19.
- ZIMAN, J.M. (1968): *El conocimiento público: Un ensayo sobre la dimensión social de la ciencia*, México: Fondo de Cultura Económica, 1972 (versión original: *Public Knowledge. The social dimension of science*, London: Cambridge Univ. Press, 1968).
- ZUCKERMAN, H.; MERTON, R.K. (1971): "Pautas institucionalizadas de evaluación en la ciencia", en: MERTON, R.K. (1973): *La sociología de la ciencia, 2. Investigaciones teóricas y empíricas*, Madrid: Alianza Editorial, 1977, pp. 579-621. (comp. de Norman Storer) [publicado originalmente como "Patterns of Evaluation in Science:

- Institutionalisation, Structure and Functions of the Referee system", *Minerva*, v.9, n.1 (1971) 66-110].
- ZUCKERMAN, H.; MERTON, R.K. (1972): "Edad, envejecimiento y estructura de edades en la ciencia", en: MERTON, R.K. (1973): *La sociología de la ciencia*, 2. *Investigaciones teóricas y empíricas*, Madrid: Alianza Editorial, 1977, pp. 622-697.
- ZUCKERMAN, H. (1968): "Patterns of name ordering among authors of scientific papers" *American Journal of Sociology*, v.74 (1968) 276-291.
- ZUCKERMAN, H. (1987): "Citations analysis and the complex problem of intellectual influence", *Scientometrics*, v.12, n.5-6 (1987) 329-338.
- ZUCKERMAN, H.; MILLER, R.B. (1980): "Introduction", *Scientometrics*, v.6, n.5-6 (1980) 327-330.
- ZUCKERMAN, H. (1977): "Deviant behavior and social control in science", en: SAGARIN, E. (ed.): *Deviance and social change*, London: Sage, 1977.