Indicadores Cuantitativos y Cualitativos para la Evaluación de la Actividad Investigadora: ¿Complementarios? ¿Contradictorios? ¿Excluyentes?

Alfonso Juan López Baena

Unidad para la Garantía de la Calidad de la Universidad de Córdoba

Miguel Valcárces Cases

Departamento de Química analítica y Ecología de la Universidad de Córdoba

Manuel Barbancho Medina

Unidad para la Garantía de la Calidad de la Universidad de Córdoba

Resumen

En el ámbito de la evaluación de la calidad investigadora universitaria se utilizan con frecuencia indicadores, como una herramienta fundamental de apoyo al evaluador. Distintos indicadores son clasificados en cuantitativos y cualitativos por

diversos autores. En estas clasificaciones se aprecia un uso no normalizado de la terminología. Utilizando la literatura científica existente sobre el tema, se analizan los distintos criterios de clasificación existentes de los indicadores, y se hace por último una propuesta de normalización terminológica.

1. Introducción¹

Una de las herramientas más utilizadas en la evaluación de la calidad de la universidad, ya se trate de la actividad docente, la actividad investigadora, la gestión o la prestación de servicios, son los llamados indicadores. Definidos por Ginés (1999) como *“medidas objetivas, usualmente cuantitativas, de una institución o de todo un sistema de educación”* ejemplo, el sexo, la religión, o el color de los ojos.

Las variables cuantitativas son aquellas cuyos valores, además de ser diferentes, se pueden ordenar de mayor a menor. Pueden expresarse numéricamente, como el peso, la temperatura, los ingresos anuales o el número de documentos producidos por un grupo de investigación. Las variables cuantitativas pueden ser ordinales, de intervalo lineal o de razón. En las variables ordinales los valores son diferentes y se pueden ordenar; admiten postulados del

tipo “mayor que” o “menor que” (por ejemplo, el estatus social). En las variables de tipo lineal, los valores, además de ser diferentes y ordenados, conservan una unidad común; es decir, que las distancias entre valores consecutivos son idénticas. No sólo establecen un orden entre los elementos de un conjunto, sino que también lo establecen en las diferencias entre elementos (por ejemplo, el CI o la temperatura). En las variables de razón, además de los rasgos anteriores, existe un cero real (por ejemplo, el peso, la altura, la edad).

Un tipo especial de variables estadísticas son las variables cuasi-cuantitativas (Fernández et al., 1987). Inicialmente eran variables cualitativas, pero al asignarles un número que indica mayor o menor grado de atributo adquieren la capacidad de cuantificación. Por ejemplo, el grado de internacionalización de la producción cien-

tífica de un grupo de investigación: alta, media o baja. Un determinado método de investigación social (por ejemplo, una encuesta) puede producir variables tanto cuantitativas como cualitativas, dependiendo de las características de la información con él obtenida.

En el campo de la investigación social, se puede hablar de métodos cuantitativos y cualitativos de recogida de información. Son métodos cualitativos la observación, la dinámica de grupos, la entrevista o las técnicas proyectivas. Recogen una información difícilmente mensurable, como son los juicios, las actitudes o los deseos. Por su parte, los métodos cuantitativos (como las encuestas o el análisis sistemático de bases de datos) están basados en muestras obtenidas por métodos estadísticos (Grande, 1995). Su principal característica es la capacidad de extrapolación de la información².

2. Clasificación de los Indicadores en Cuantitativos y Cualitativos

En el campo de la evaluación de la investigación científica existen múltiples clasificaciones de indicadores, a menudo confluyentes. Una clasificación a menudo utilizada (Valcárcel, 1999) es la que distingue entre indicadores cuantitativos y cualitativos.

A) Indicadores cuantitativos.

Dentro de los indicadores cuantitativos se comprenden:

a) *Indicadores de la situación.*

Describen fundamentalmente aspectos relacionados con la organización, la gestión de la investigación y la infraestructura de investigación. La existencia o no de objetivos específicos para la investigación es un indicador de organización destacado. Los indicadores referidos a eficacia y eficiencia en la gestión de recursos humanos y económicos son ejemplos de indicadores de gestión. Por su parte, la proporción de infraestructura dedicada a instrumentación constituye un ejemplo claro de un indicador de infraestructura.

b) *Indicadores de los inputs.*

Describen aspectos relacionados con los recursos disponibles para hacer investigación³. Una primera clasificación distingue entre indicadores absolutos, relativos o evolutivos. Son ejemplos de indicadores absolutos el número de proyectos de investigación conseguidos por la unidad evaluada, o el número de becas con las que cuenta para el desarrollo de su labor. Los indicadores relativos de inputs pueden referirse a la actividad desarrollada por el grupo de investigación (ej., tasa de actividad en proyectos de investigación: nº de proyectos de investigación / nº de doctores del grupo), al éxito del mismo (ej.: tasa de becas concedidas sobre las solicitadas) o a la concentración de los recursos con los que dispone. Por último, los indicadores de evolución expresan la variación experimentada en una determinada magnitud con el paso del tiempo, utilizando fundamentalmente dos tipos de medidas, **(1) incrementos puros** (sobre un período base que sirve de referencia) y **(2) utilizando medias móviles para estimar la tendencia de la variable estudiada.** Una segunda clasificación de indicadores de los inputs genera dos nuevos tipos de indicadores: los indicadores económicos y los indicadores de personal. A todos los

¹ Los autores agradecen a Javier Vidal los comentarios y sugerencias aportadas, integradas en el presente artículo.

² Esta última clasificación de los métodos de recogida de información en cuantitativos y cualitativos es discutida con frecuencia por diversos autores, que opinan que lo que se considera cualitativo o cuantitativo son los datos, no las técnicas de recogida de información. Según estos autores, las distintas técnicas de recogida de información no pueden ser clasificadas como cuantitativas o cualitativas. Una misma técnica (imaginemos una encuesta) puede contener datos cuantitativos y cualitativos, indistintamente. Sin embargo, la comunidad científica, como veremos más adelante, sí adjetiviza los distintos métodos de recogida de información, prestando escasa atención a las características de los datos primarios recogidos por los mismos métodos. En otras palabras, los métodos de investigación se suelen clasificar como cuantitativos o cualitativos por las características del método *per se*, no por las características de los datos que contienen.

³ En algunas ocasiones, ciertos indicadores aquí clasificados como de inputs pueden ser considerados igualmente como de output. Por ello, en ocasiones reciben el nombre de indicadores de output intermedio.

efectos son indicadores de los inputs, ya que en definitiva miden los recursos disponibles para el desarrollo de la actividad investigadora, por lo que pueden perfectamente situarse como una subcategoría de los mismos. Sin embargo, en numerosas ocasiones se encuentran formando conjuntos independientes al de los indicadores de inputs, ya que la importancia de los aspectos económicos y de personal de la actividad investigadora justifica el que los indicadores que traten estas materias se contemplen de forma separada. Exponemos en este trabajo tanto los indicadores económicos como los de personal separadamente a los indicadores de los inputs, si bien consideramos igualmente válida la alternativa de integrarlos en los indicadores de inputs ya que, al fin y al cabo, como hemos afirmado, no son sino un tipo especial de ellos.

c) Indicadores económicos (subconjunto de los indicadores de inputs).

Describen los recursos económicos utilizados para desarrollar la investigación. Son utilizados fundamentalmente en la evaluación de la investigación en niveles agregados. El Manual de Frascati (OECD, 1994) nos proporciona diversos tipos de indicadores económicos para medir el esfuerzo investigador de una determinada unidad objeto de evaluación: c1) Los referidos a gastos intramuros⁴. Se dividen en gastos corrientes (ej.: compra de libros para la investigación o suscripción a revistas científicas) y gastos de capital (ej.: tierras y edificios, así como adquisición de instrumentos para la investigación). C2) Los referidos a las fuentes de fondos de I+D. La medición de las fuentes de los fondos destinados a I+D por la unidad objeto de evaluación se puede realizar por dos procedimientos: a) en base a la suma de dinero que una unidad recibe de otra unidad, organización o sector en función de la investigación que realiza y b) en base a las cantidades extramuros que una unidad, organización o sector declara haber pagado a otra unidad, organización o sector por el desarrollo de actividades de I+D. En este sentido, la primera de las aproximaciones es la más

recomendada por el Manual de Frascati. C3) Los referidos a gastos extramuros. C4) Los referidos a indicadores económicos nacionales.

d) Indicadores de personal (subconjunto de los indicadores de inputs).

Describen el número de personas (tipología) y su dedicación a tareas directa o indirectamente relacionadas con I+D. Los indicadores de personal pueden ser a su vez absolutos, relativos o evolutivos. Se emplean normalmente para definir otros indicadores relativos.

e) Indicadores de los outputs (o resultados).

Describen aspectos relacionados con los resultados obtenidos en la investigación. Se clasifican a su vez en directos e indirectos. Son indicadores directos los que miden directamente resultados de investigación en sí mismos, como son los referidos al número de artículos científicos publicados, el número de patentes alcanzadas o la participación en congresos científicos, entre otros. Por su parte, son indicadores indirectos los que miden no resultados en sí mismos, sino resultados que denotan indirectamente el desarrollo de una actividad investigadora. Ejemplos de indicadores indirectos son los referidos a las tesis doctorales desarrolladas en el seno de la unidad objeto de la evaluación, o los premios de investigación alcanzados por miembros de la unidad evaluada. Los indicadores de los outputs pueden asimismo ser clasificados en indicadores de outputs absolutos, relativos y evolutivos, al igual que sucedía con los indicadores de los inputs.

B) Indicadores bibliométricos.

Indicadores cuantitativos, son un caso especial de indicadores de outputs, que tienen en cuenta la producción científica de la persona / institución objeto de la evaluación. Se distinguen:

a) Los recuentos de “productos” científicos, tales como las publicaciones en revistas científicas, los libros científicos

⁴ En este sentido, se consideran gastos intramuros los realizados dentro del a unidad objeto de evaluación.

cos, las patentes o las comunicaciones a congresos.

b) Los recuentos de citas de los trabajos científicos. Se han utilizado tradicionalmente para medir la influencia de agregados de trabajos científicos, y así poder realizar comparaciones entre las distintas unidades evaluadas. Su utilización ha de hacerse guardando unas estrictas normas en lo que a la comparación se refiere, y teniendo en cuenta los múltiples inconvenientes que los recuentos de citas presentan como estimadores del impacto científico.

c) Los recuentos de publicaciones ponderados por la calidad de los trabajos científicos. Para realizar estos recuentos es necesario abordar primariamente la calidad per se de los trabajos científicos y, de forma secundaria, la calidad de las revistas científicas donde los trabajos científicos aparecen publicados. En las áreas científicas experimentales se utiliza el Factor de Impacto⁵ como estimador de la calidad de los trabajos científicos; en las áreas no experimentales este es un tema aún no resuelto⁶.

C) Indicadores cualitativos.

Son aquellos basados en encuestas o entrevistas, así como informaciones no estructuradas, métodos de investigación a partir de los cuales suele ser problemática la extrapolación estadística. Suministran información sobre aspectos no contemplados en los indicadores cuantitativos y bibliométricos, por lo que su utilización suele ser complementaria a las metodologías cuantitativas. Son imprescindibles para lograr una evaluación integral y global de la investigación.

La clasificación de indicadores anterior-

mente expuesta es en principio coherente con la tipología de métodos de investigación social contemplada en el apartado 2 del presente artículo. Así, los indicadores cuantitativos son aquellos basados en metodologías cuantitativas de obtención de información. La información que suministran es fácilmente cuantificable, y la extrapolación estadística no presenta dificultades extraordinarias. Por el contrario, los indicadores cualitativos están basados en métodos cualitativos de obtención de información. Tienen la propiedad de matizar, enriquecer a los indicadores cuantitativos, aunque en ocasiones se utilizan como sistema exploratorio, previo a otros indicadores cuantitativos) Suelen carecer de la objetividad que se atribuye a estos últimos, por lo que la extrapolación estadística a menudo no es abordable.

Diversos autores clasifican igualmente los indicadores en indicadores cuantitativos y cualitativos, de forma similar a la expuesta anteriormente. **Bellavista et al. (1991)** expone en un interesante artículo sobre la evaluación de la investigación social una clasificación de métodos coincidente: a) metodologías cuantitativas y b) cualitativas; cuando trata el tema de los indicadores (Bellavista et al, 1997: 37) clasifica los indicadores bibliométricos como indicadores cuantitativos: “*Los indicadores bibliométricos son medidas cuantitativas elaboradas a partir de la bibliografía científica*”. Bellavista considera asimismo a las técnicas bibliométricas como las más adecuadas para medir la cantidad, el impacto y el significado de la producción científica.

Cami et al. (1993), en un estudio de la producción científica española en biomedicina, clasifica los indicadores bibliométricos que utiliza en: a) cuantitativos. Basados en el número de publicaciones y b) cualitativos. Con ellos se estudia la influencia, la

⁵ El Factor de Impacto, introducido por Garfield, presenta numerosas limitaciones como estimador de la calidad de los trabajos científicos publicados por los investigadores. Aún así, una utilización racional del mismo nos pueden llevar a una correcta estimación de la calidad cuando se trata de la evaluación de agregados de documentos suficientemente grandes y representados en el Science Citation Index (SCI).

⁶ Se consideran áreas no experimentales las áreas de Ciencias Jurídicas, Humanas y Sociales, donde la producción científica española se encuentra lejos de estar suficientemente representada en las bases de datos del ISI. Actualmente se está desarrollando en la Unidad de Garantía de Calidad de la Universidad de Córdoba, en colaboración con la Unidad para la Calidad de las Universidades Andaluzas, el modelo UCO-UCUA, un trabajo de investigación encaminado a la determinación de la calidad de las publicaciones en las áreas no experimentales, basado en una metodología mixta: realización de encuestas y posterior refrendo de los resultados mediante el análisis de expertos. La clasificación de los recuentos de publicaciones como indicadores cuantitativos es comúnmente aceptada por la comunidad científica. Los recuentos de publicaciones son considerados, siempre que se cumplan unas condiciones mínimas de homogeneidad, un estimador indirecto de la cantidad de producción científica desarrollada por la unidad evaluada.

visibilidad de la producción científica y el tipo de investigación que se realiza (no cita explícitamente la dimensión calidad de los trabajos científicos). Son, según Cami, indicadores basados en diversas clasificaciones de las revistas utilizadas, así como en las citas recibidas. Vemos en este caso una clasificación divergente con las dos estudiadas anteriormente.

Al igual que hace Bellavista ⁷, Cami clasifica los recuentos de publicaciones como indicadores bibliométricos cuantitativos. Sin embargo, el resto de los indicadores bibliométricos (basados en las citas recibidas por los trabajos científicos y en la revista donde los trabajos aparecen publicados) son considerados como cualitativos. Parece que Cami asimila cuantitativo a cantidad de producción, mientras que cualitativo lo asimila a otros aspectos que se suponen orientados a la calidad, como son la visibilidad o importancia del trabajo científico (que no hay que olvidar que se consideran habitualmente como indicadores, cuando menos parciales, de la calidad de un determinado agregado de documentos científicos). No sigue, pues, el criterio del método de obtención de la información, sino el criterio de qué es lo que se pretende medir.

Escribano y Viladiú (1996) parecen coincidir con el criterio seguido por Cami. Al clasificar los análisis realizados de la producción científica de la Universidad de Barcelona, distinguen: a) análisis cuantitativos. Se refieren a la cantidad de producción, que puede estar ponderada por el tipo de aportación que se trate (ej.: artículos, libros, patentes, ...). El objetivo final es la obtención de una medida de la producción de un determinado científico (ponderándolo, como ya hemos comentado, por la importancia relativa que en el modelo se le quiera atribuir a los distintos tipos de ítems) b) análisis cualitativos. Se trata de indicadores de la calidad investigadora desarrollada por el individuo a evaluar. Viene determinada por una clasificación de los distintos outputs en varios grupos, dependiendo de la calidad de los

mismos. El análisis cualitativo viene así a ajustar las medidas ofrecidas por los análisis cuantitativos.

Tanto los análisis cuantitativos como los cualitativos (siguiendo la denominación de Escribano y Viladiú) se realizan a partir de los datos proporcionados por las bases de datos del SCI (en las áreas experimentales). En cierto modo, los indicadores cuantitativos se pueden considerar parecidos a los recuentos de publicaciones (vistos en la primera clasificación de indicadores expuesta), ya que miden la cantidad de producción científica de un sujeto determinado; los análisis cualitativos se pueden considerar similares a los recuentos de publicaciones ponderados por la importancia de la aportación. Vemos como, por lo tanto, Escribano y Viladiú siguen un criterio similar al de Cami, ya que consideran cualitativo la medida de aspectos orientados a la calidad de trabajos científicos desarrollados por el sujeto/s sometido/s a estudio.

Sancho (1990) no realiza una clasificación de los distintos indicadores existentes en indicadores cuantitativos y cualitativos. En contraposición con otros autores, afirma que la calidad de los trabajos científicos sólo puede ser medida a través de indicadores basados en percepciones (clasificados por Bellavista como cualitativos): *“ Los indicadores a través de los cuales se obtiene información sobre los aspectos de calidad, sólo pueden ser: indicadores basados en percepciones (opiniones de expertos), que juzgan las publicaciones por su contenido científico, y se basan en las estimaciones que se forman los colegas observadores a través del estudio de los trabajos publicados, cuestionarios, cartas, entrevistas, etc. acerca de la investigación llevada a cabo y los científicos que la realizan”*. Es decir, Sancho asimila los tradicionalmente llamados indicadores cualitativos con la medición de la calidad de un determinado trabajo. Por tanto, restringe la responsabilidad de su medida al examen de expertos, rechazando a los indicadores bibliométricos como herramientas útiles para ello. Para Sancho, los indicadores bibliométricos sirven para la

⁷ La clasificación de los recuentos de publicaciones como indicadores cuantitativos es comúnmente aceptada por la comunidad científica. Los recuentos de publicaciones son considerados, siempre que se cumplan unas condiciones mínimas de homogeneidad, un estimador indirecto de la cantidad de producción científica desarrollada por la unidad evaluada.

medición de la importancia o visibilidad de un determinado trabajo científico.

Maltrás (1996) analiza en su obra los indicadores bibliométricos (que hemos visto inicialmente como una tipología particular de indicadores cuantitativos) Distingue a su vez entre: a) Indicadores bibliométricos de producción. Fundamentalmente el recuento de publicaciones. b) Indicadores bibliométricos de calidad.

Este autor rehuye la problemática de la denominación de un indicador como cuantitativo o cualitativo; así, a los indicadores bibliométricos que se derivan del análisis de citas o del impacto de la revista en la que aparecen los trabajos los denomina como "*indicadores de calidad*". Reconoce, por tanto, a estos indicadores la facultad de estimar la calidad científica de un agregado de documentos, mediante la construcción de determinados índices con el objetivo de normalizar los distintos datos existentes para permitir la comparación entre autores pertenecientes a distintas disciplinas.

Diversos autores no españoles analizan en sus obras la construcción y significado de distintos indicadores para la evaluación del desempeño científico. Así, Martín e Irvine (1983) exponen como medidas para medir el progreso científico de un grupo de investigación varios indicadores parciales, basados en : **a) Los recuentos de publicaciones b) El análisis de citas y c) La revisión por expertos.** Aunque en un principio no clasifican los mismos en cuantitativos y cualitativos, afirman: (Martín e Irvine, 1983: 74): "*Todas las medidas cuantitativas de la ciencia son en el mejor caso indicadores parciales, influenciados por una red de factores interrelacionados de los cuales el tamaño de la contribución al progreso científico es uno de ellos*". Ahora bien, no especifican claramente qué medidas son cuantitativas y qué medidas son cualitativas. En un trabajo posterior (Crouch, Irvine y Martín, 1985) sí enfrentan el recuento de citas y el análisis de citas (cuantitativos) con el peer-review (cualitativo): "*Las evaluaciones cuantitativas del output y el rendimiento de la investigación son cada vez más requeridas como soporte al peer-review*".

Así pues, estos autores parecen coincidir con Bellavista, al adoptar el criterio del método de investigación social seguido; indicadores cuantitativos y cualitativos son, según este prisma científico, complementarios en su aportación a la evaluación de la actividad investigadora. En un trabajo posterior (Martín e Irvine, 1984), a través de entrevistas intensivas y estructuradas (revisión por expertos), piden a 182 físicos el establecimiento de unos rankings de instalaciones científicas, en base a la magnitud de sus contribuciones a la física de alta energía. Calculan posteriormente un ranking medio para cada instalación. En el apartado de conclusiones del artículo, denominan a esta información como "quantitative peer-evaluation data" (datos cuantitativos derivados del peer review). Estamos ante un claro ejemplo de cómo un indicador considerado inicialmente como cualitativo, por el método original de obtención de la información (el peer-review) se puede considerar posteriormente como cuantitativo, al haber introducido en su análisis métodos estadísticos, permitiendo así la posibilidad de realizar estimaciones sobre diversos aspectos de la realidad.

Moed y van Raan (1985), clasifican los indicadores bibliométricos de output (producción) e impacto utilizados por Irvine y Martín como medidas cuantitativas, al igual que los rankings generados por el examen de expertos. Sin embargo, afirman que las discrepancias observadas en los anteriores indicadores deberían llevar a cuestiones dirigidas a los expertos, a través de metodologías cualitativas. Por tanto, de nuevo vemos una doble consideración del peer-review: a) cualitativo, cuando se trata de opiniones de unos pocos expertos, difícilmente cuantificables y extrapolables y b) cuantitativo, cuando el tamaño de la muestra y la forma de organización de la información permite la cuantificación y extrapolación de forma fiable. Indicadores cualitativos en su origen se convierten en cuantitativos, por los métodos de análisis de información seguidos.

Parece así surgir un tercer criterio de clasificación de los distintos indicadores: dependiendo del método seguido en el proceso de la información; dicho en otras

palabras: dependiendo de la posibilidad de la cuantificación de la información. Este método parece coincidir con la clasificación de las variables estadísticas estudiadas en el apartado anterior. En línea con lo comentado, Moed vuelve a clasificar en un trabajo posterior (Moed et al., 1985) los indicadores bibliométricos como indicadores cuantitativos basados en la literatura científica. Hicks (1990) coincide con Moed, al clasificar los análisis de citas como métodos cuantitativos basados en la literatura. Schubert (1991) se encuentra en línea con Moed y Hicks, al titular su recopilación bibliográfica como: *“Estudios cuantitativos de la ciencia. Una bibliografía actual”*.

Englisch (1990) realiza un análisis de las medidas bibliométricas y del peer review, unificando ambas medidas en un modelo matemático, llamado cuantitativo por Englisch. Por tanto, parece seguir la clasificación de indicadores realizada por Moed y van Raan: dependiendo del método seguido en el proceso de la información.

Velho (1986) clasifica los datos utilizados para un estudio en a) cuantitativos: fundamentalmente las citas aparecidas en una gran muestra de artículos, y b) cualitativos: entrevistas con un gran número de científicos. Sigue, por tanto, el criterio de la metodología seguida de obtención primaria de información. Igual criterio sigue Alcaín (1993), al afirmar: *“los trabajos de investigación en las Ciencias Sociales y Humanas, en general, se caracterizan por la escasa utilización de métodos cuantitativos”*. Se refiere Alcaín a la utilización de métodos estadísticos, que cuantifican la información recogida en las investigaciones. Sin embargo, áreas científicas como la economía o la Psicología, encuadradas en las Ciencias Sociales, utilizan frecuentemente métodos estadísticos para el análisis de la información recogida en las investigaciones.

Un claro ejemplo de cómo una información en principio cualitativa (como es la encuesta a expertos), tratada con técnicas propias de estudios cuantitativos, puede conducir a resultados sesgados si los autores no adoptan unas mínimas cautelas metodológicas se muestra en Fernández (1999). El objetivo del anterior trabajo es la evaluación de revistas en las áreas de Ciencias de la

Educación. Para ello, se realiza una encuesta a 323 encuestados miembros de la Asociación Interuniversitaria de Investigación Pedagógica (AIDIPE); se le pide a los encuestados que valoren de 1 a 12 una serie de revistas (doce revistas de un total de 22 revistas del campo considerado). El resultado del estudio es que la revista más valorada es la Revista de Investigación Educativa, órgano portavoz de la asociación cuya población se encuesta. El principal resultado de la encuesta muestra cómo la técnica elegida de muestreo determina los resultados que de ella se deducen, resultados que son sólo representativos para la muestra encuestada. Ahora bien, cabría plantearse otra pregunta: ¿Y si la Revista de Investigación Educativa fuese de verdad la mejor revista? ¿Qué criterio habría de utilizarse para contrarrestar la opinión de 323 personas?

En un modelo de evaluación individual de la investigación desarrollado en Cataluña (Generalitat de Catalunya. CIRIT, 1999) no se hace una distinción explícita de indicadores cuantitativos y cualitativos, pero se afirma: *“El sistema de evaluación ha de incidir en aspectos cualitativos y no meramente cuantitativos”*. Posteriormente, se afirma: *“Los evaluadores... han de ajustar las puntuaciones teniendo en cuenta los aspectos cualitativos de los items seleccionados por cada investigador evaluado”*. Parece que en el anterior sistema lo cualitativo se refiere a aspectos asociados con la calidad (a juicio del evaluador externo), mientras que lo cuantitativo se refiere a la cantidad de producción científica, no ponderada por la calidad de la aportación.

Vinkler (1988), en las conclusiones de un trabajo sobre la clasificación de distintos indicadores bibliométricos para propósitos cientométricos adjetiviza los indicadores bibliométricos como cuantitativos. Sin embargo, en la introducción del mismo afirma: *“En orden a evaluar la actividad de publicación, tanto cuantitativamente como cualitativamente, se recomiendan los indicadores de citas relativos”*. Vemos de nuevo como cualitativo se refiere en ocasiones no a los métodos de investigación, sino a los aspectos de calidad de los trabajos científicos.

3. La Integración de Indicadores Cuantitativos y Cualitativos.

Una cuestión frecuentemente abordada en cualquier evaluación de la actividad investigadora es la concerniente a qué tipos de indicadores han de ser utilizados en la misma. Los diversos indicadores, según se clasifiquen en cuantitativos y cualitativos (siguiendo el criterio del método de investigación social empleado), tienen asociadas unas características derivadas directamente de su propia naturaleza. Los cuantitativos son frecuentemente considerados por la comunidad científica como medidas objetivas de evaluación, por la posibilidad de abordar su cálculo a través de procesos informáticos de tratamiento de la información. Ahora bien, están afectados (especialmente los bibliométricos) por limitaciones diversas que no conviene olvidar, algunas de las cuales ya han sido comentadas anteriormente. Algunas de ellas son derivadas de las características de la principal fuente de datos de la cual se extraen los indicadores, por lo que afectan a todos los indicadores bibliométricos basados en el SCI. Otras tienen un trasfondo más complejo, asociado con la naturaleza de la información extraída, y afectan a los diversos indicadores con distinta intensidad. Vemos a continuación las principales.

a) Limitaciones técnicas del SCI. (Hicks y Crouch, 1990). a1) Cobertura del SCI. El SCI sólo se considera válido para evaluar aquellos campos con orientación internacional, debido al perjuicio que la literatura no inglesa experimenta con su utilización, así como a su limitada cobertura, estimada frecuentemente en un 90% de la literatura científica (Reedijk, 1998). Del mismo modo, se produce una no inclusión masiva de libros, con sus respectivas referencias bibliográficas. Igualmente, se produce un importante sesgo contra la investigación aplicada, derivada de la exclusión

de productos científicos como pueden ser las patentes o la "literatura gris". a2) Consistencia del SCI. Se producen frecuentemente errores en la grabación de los datos por parte bien de los autores de los trabajos, de las revistas que publican los trabajos o por el propio ISI. Así, por ejemplo, se pueden encontrar casos en que autores distintos aparecen con el mismo nombre, o casos en que un mismo autor aparece con dos denominaciones distintas. a3) Indexado por el primer autor. Este método de indexado produce que autores que no aparecen en primer lugar no sean adecuadamente valorados utilizando técnicas bibliométricas basadas en el SCI, ya que la versión del SCI más asequible sólo contempla el primer autor de cada uno de los trabajos. Una consulta directa a la base de datos central del SCI obvia esta dificultad.

b) Limitaciones del número de citas recibidas como estimadores del impacto de un determinado trabajo científico.

b1) La cita estima el impacto a corto plazo, que puede no ser indicativo de la calidad del mismo. b2) La cita no diferencia entre impacto positivo e impacto negativo (críticas a un trabajo). b3) Distintas áreas científicas pueden tener diferentes costumbres en cuanto a citas. La comparación interáreas no es posible sin una adecuada normalización. b4) Existe un sesgo entre cita e impacto, como consecuencia de la autocita. b5) El efecto notoriedad puede influir en que un autor sea más citado que otro, aunque el impacto real de la investigación no sea significativamente superior. b6) El anteriormente comentado sesgo geográfico y lingüístico del SCI produce un sesgo en el recuento de citas como estimador del impacto científico.

c) Limitaciones del Factor de Impacto como estimador de la calidad de un documento científico determinado.

c1) El Factor de Impacto indica el impacto de una determinada revista, pero no de los artículos en ella publicados. Como indica Reedijk (1998): "Las tasa de cita de los artículos determinan el Factor de impacto, pero no al revés". Así, su utilización es sólo válida para grandes agregados de artículos. c2) El Factor de Impacto a corto plazo

presenta una gran variabilidad. Esta variabilidad condiciona su validez como indicador de la calidad de una determinada revista científica. c3) En el cálculo del factor de impacto, las citas recibidas por los artículos tienen el mismo valor independientemente tanto del año en que se recibieron como de la importancia de la revista citante. c4) La aparición de la cita tiene distinta estructura temporal dependiendo de la disciplina científica. c5) Las disciplinas consolidadas resultan beneficiadas en el cálculo del Factor de Impacto.

El conjunto de limitaciones expuestas (que no son sino una muestra) trata no de invalidar los indicadores bibliométricos, sino de alertar sobre un uso indiscriminado e irracional de los mismos. Objetividad no es, por lo tanto, garantía de validez.

Los indicadores cualitativos (examen de expertos según la clasificación adoptada), por la naturaleza de la información que de ellos emana, no se ven afectados por los errores que atañen a los cuantitativos. Sin embargo, sí presentan determinadas carencias, relacionadas con la naturaleza del ser humano, dictaminador último de la calidad de la investigación desarrollada por la unidad objeto de evaluación. Consideramos las más importantes: a) El posible empleo por los evaluadores de criterios "propios", que deriven en resultados no ajustados a los objetivos de la evaluación. b) Las posibles influencias sobre los expertos, que pueden condicionar sus opiniones en el sentido de las presiones recibidas. Esta limitación es especialmente detectable en pequeños sistemas de I+D.

En numerosas ocasiones se ha efectuado un uso alternativo y mutuamente excluyente de los indicadores bibliométricos y del examen de expertos en la evaluación de la actividad investigadora. ¿Es adecuado este procedimiento? En nuestra opinión, no es lo más indicado, fundamentalmente por dos motivos. En primer lugar, los resultados de distintas investigaciones (la principal se encuentra resumida en Moed et al, 1985b), muestran que se puede dar una falta de

acuerdo en los resultados alcanzados utilizando análisis bibliométrico y análisis de expertos. En el trabajo citado se duda del criterio empleado por los evaluadores externos, que parecen haber realizado un ranking de los distintos grupos evaluados, más que una comparación de su calidad en relación con el nivel internacional existente en la disciplina. En segundo lugar, las características de ambos métodos de evaluación son completamente distintas, por lo que es de esperar de nuevo distintos resultados en las evaluaciones.

Así pues, ¿cuál debe ser el criterio seguido en la utilización de los indicadores cuantitativos y cualitativos? A la vista de lo expuesto, lo recomendable parece la utilización conjunta e integrada de los distintos tipos de indicadores. Los indicadores bibliométricos, esencialmente objetivos, han de servir como fuente de información a los evaluadores, que han de recibir asimismo claras instrucciones sobre los criterios a seguir en la evaluación, así como sobre los objetivos de la misma. Se propone la siguiente metodología mixta: 1) Elaboración de un perfil bibliométrico de la unidad evaluada (información cuantitativa). 2) Emisión por parte de los expertos externos de un juicio sobre la unidad evaluada (información cualitativa), contando para ello con el asesoramiento de los indicadores bibliométricos elaborados (información cuantitativa).

3) En el caso de no coincidencia de los resultados alcanzados, la posibilidad de intervención de un tercer experto independiente parece aconsejable (información cualitativa) Este experto se valdría de los dos métodos anteriormente citados para dictaminar su juicio, en este caso definitivo⁸.

La utilización conjunta e integrada de los indicadores empleados se basa tanto en los distintos aspectos de la investigación en que los mismos se centran (es decir, por su carácter complementario) como en la mayor aceptación por parte de la comunidad científica del citado método.

⁸ Los críticos a este procedimiento podrían alegar que el mismo no garantiza una respuesta infalible, ya que el tercer experto no tiene por qué ser el criterio de validez. Mostramos aquí nuestro total acuerdo con esta posible objeción. En el caso de que se determinase que el tercer experto no fuese el criterio de validez, se contemplan otras alternativas, como la integración estadística de los resultados obtenidos por distintos métodos, como los clásicos análisis cluster o análisis discriminante. La utilización de procedimientos estadísticos más avanzados, como utilización de redes neuronales artificiales, resulta otra opción interesante de explorar.

Así, Martín sugiere que los científicos prefieren el uso combinado de indicadores cuantitativos y de los métodos de peer-review. (Martín, 1996, citado en Vincler, 1998).

4. Discusión

Hemos visto como distintos autores clasifican los indicadores de evaluación de la investigación científica en cuantitativos y cualitativos según diversos criterios: 1) Criterio del método de investigación social seguido para la obtención de los datos. 2) Criterio del objeto de medida. 3) Criterio del proceso estadístico aplicado a los datos primarios.

Los indicadores que generan mayor controversia en su clasificación son los indicadores bibliométricos. Vemos la situación de los indicadores bibliométricos, así como del examen por expertos:

A) Recuentos de publicaciones (ej.: número de artículos en revistas científicas) Son clasificados habitualmente por los distintos autores como indicadores cuantitativos. En principio es una clasificación coherente con los tres criterios analizados anteriormente. Si se sigue el criterio del método de investigación, se puede alegar que estos recuentos no están basados en encuestas de ningún tipo, ni en ningún método de investigación social que pudiese denominarse cualitativo. Si se sigue el criterio del objeto de medida, existe un amplio consenso científico sobre el objeto de medida de los recuentos: la cantidad de producción científica. Por último, si se sigue el criterio del proceso estadístico aplicado a los datos primarios, existe también un amplio consenso sobre la capacidad de cuantificación de los mismos (siempre que se sigan unas normas mínimas contenidas en Maltrás, 1996).

B) Recuentos de citas (ej.: número de citas recibidas por cinco artículos científicos) Sobre estos indicadores no hay, como se ha venido comentado durante

el desarrollo del presente artículo, una opinión unánime en la literatura científica. Los recuentos de citas son clasificados habitualmente como cuantitativos, siguiendo el criterio del método de investigación social empleado (no se trata de la realización de una entrevista ni de la observación, sino de un método mucho más objetivo). Los autores que, por el contrario, se rigen por el criterio del objeto de medida, clasifican estos indicadores como cuantitativos (si consideran que no pueden medir la calidad de un trabajo científico determinado) o cualitativos (si consideran que con las debidas precauciones se puede medir la calidad de un trabajo científico a través de los recuentos de citas). Los autores que se rigen por el criterio del método estadístico seguido clasifican los indicadores derivados del recuento de citas como cuantitativos, al existir la posibilidad de desarrollar modelos matemáticos cuantitativos a partir de ellos. Cualquiera de los criterios parece aceptable; ahora bien, si se sigue el criterio del método de investigación empleado, podría alegarse que estos indicadores tienen un contenido cualitativo en su origen, ya que cuando un autor está citando a otro está en definitiva expresando una opinión sobre su trabajo, o refiriéndose a él por algún otro motivo. Aunque posteriormente el conjunto de citas se cuantifique a través de las bases de datos internacionales (SCI fundamentalmente para las áreas científicas), la esencia de los recuentos podría caracterizarse por su naturaleza cualitativa. Por ello, estos indicadores podrían denominarse como cuasi-cuantitativos.

C) Los recuentos de publicaciones ponderados por la calidad de los trabajos científicos (ej.: número de artículos científicos ponderados por el Factor de Impacto de la revista en la que aparecen) Si los autores siguen el criterio del método de investigación social seguido, estamos en un caso parecido al que ocurría con los recuentos de citas. Aunque la mayoría de los autores juzgan estos indicadores como cuantitativos, por basarse en la recopilación cuantitativa y sistemática de información a partir de las bases de datos existentes, es evidente de nuevo la componente cualitativa que en su origen tienen. Cuando se trata de la evaluación de publicaciones en

ciencias experimentales, la introducción de medidas como el Factor de Impacto hace que se tenga en cuenta la opinión de la comunidad científica, cuyo origen (la cita) es, como ya vimos, un posible elemento cualitativo que puede afectar a su consideración como puramente cuantitativos. Cuando se evalúa la investigación desarrollada en las Ciencias Sociales en países no sajones, la introducción de opiniones de expertos en la valoración de las revistas científicas, que un vez cuantificadas pasan a estimar en cierto modo la calidad de las revistas científicas, introduce igualmente un elemento cualitativo en el origen de los indicadores. Por ello, podría resultar conveniente su clasificación como indicadores cuasi-cuantitativos. Cuando se sigue el criterio del objeto de medida, surge igualmente la controversia de qué es lo que mide este tipo de indicadores. Autores que consideran que el impacto de una determinada revista científica estimará adecuadamente la calidad de una determinada publicación considerarán estos indicadores como cualitativos. Si ocurre lo contrario, los considerarán como cuantitativos. Por último, si se sigue el criterio del método estadístico seguido, los recuentos de publicaciones ponderados por la importancia de la aportación se consideran fundamentalmente como indicadores cuantitativos, por motivos coincidentes con los argumentados en el caso de los recuentos de citas.

D) Examen de expertos. Sea cual sea el criterio empleado, estos indicadores están mayoritariamente considerados como cualitativos. Ahora bien, se han estudiado en el artículo varios casos de cómo los métodos cualitativos de obtención de información pueden llegar a ser cuantificados. A nuestro juicio, se han de cumplir varias condiciones para ello: 1) Tamaño de la muestra suficiente. 2) Existencia de unas escalas de medidas adecuadas. 3) Elección de la muestra adecuada, evitando sesgos en las estimaciones. Bajo las tres condiciones anteriores, los indicadores en principio considerados como cualitativos pueden pasar a denominarse indicadores cuasi-cualitativos, si se sigue el criterio del método estadístico utilizado. Son indicado-

res cualitativos en su origen, pero a los cuales es posible aplicar métodos estadísticos, llegando a alimentar en ocasiones modelos cuantitativos de evaluación de la investigación.

5. Conclusiones

Dependiendo del criterio utilizado, los indicadores son clasificados como cuantitativos o cualitativos por diversos autores. El siguiente gráfico muestra los tres criterios utilizados⁹ por los distintos autores, así como la situación de los distintos indicadores:

Las distintas clasificaciones analizadas denotan la ausencia de una terminología homogénea y generalmente aceptada sobre los indicadores de evaluación de la producción científica, especialmente los bibliométricos los cuales, como se puede observar en el anterior gráfico, cambian habitualmente su adjetivación dependiendo del criterio seguido por el autor. En ocasiones, distintos términos reflejan diferencias lingüísticas sutiles entre diferentes autores. Sin embargo, en ocasiones podrían tener como trasfondo una diferencia conceptual amplia sobre su significado y alcance (como se ha podido constatar en el presente artículo). Por ello, es recomendable ser especialmente cautelosos en la denominación que a un indicador se le aplique. En todo caso, consideramos importante el que en cada caso se especifique claramente por el autor el criterio seguido al adoptar una u otra denominación, ya que de lo contrario podría conducir a una mala interpretación por parte del resto de la comunidad científica de su visión sobre el papel que los distintos indicadores han de jugar en una materia tan importante como es la evaluación de la actividad científica. Incluso resultaría deseable el que en cada momento se definiese por el autor la escala (nominal, ordinal, lineal o de razón) de la variable aplicada.

Se proponen los siguientes criterios

⁹ Un último criterio, ya comentado, consiste en clasificar los indicadores como cualitativos o cuantitativos dependiendo de la naturaleza de la variable estadística subyacente. Así, datos cuantitativos darían lugar a indicadores cuantitativos. Datos cualitativos darían lugar a indicadores cualitativos. Sin embargo, esta clasificación, defendida por quienes argumentan que no existen métodos cuantitativos o cualitativos, sino datos cuantitativos o cualitativos, daría lugar a una clasificación de los indicadores cuando menos atípica

para la denominación de un indicador como cuantitativo o cualitativo: 1) Seguir el criterio del método de obtención de la información. Este criterio parece el más acertado, por cuanto indica las características de los mismos, así como por ser el criterio más habitualmente utilizado por la comunidad científica cuantitativa. Sin embargo, si el autor es de la opinión de que los métodos de obtención de la información no pueden ser clasificados como cuantitativos o cualitativos, este criterio ha de ser obviado, por lo que surge el siguiente criterio. 2) Seguir el criterio del proceso estadístico utilizado para analizar la información.

En todo caso, no parecen aconsejables los dos criterios siguientes: 3) Criterio del objeto de medida. La utilización de este criterio produce una confusión importante sobre cuál es la naturaleza del indicador. Si se pretende enfatizar qué es lo que mide el indicador, puede ser conseguido fácilmente con una aclaración. 4) Criterio de la naturaleza de las variables estadísticas utilizadas. Este criterio no es en nuestra opinión acertado, por su escasa utilización por la comunidad científica. El hecho de que la variable origen del indicador sea cuantitativa o cualitativa no aporta excesiva información sobre el mismo, por lo que no parece aceptable su utilización. Además, este criterio produciría una completa invalidación de las agrupaciones hasta ahora existentes.

Respecto a la utilización aislada o conjunta de los distintos tipos de indicadores, se recomienda una metodología mixta, cuantitativa y cualitativa (en este orden temporal), como mejor aproximación al estado de la actividad investigadora de la unidad objeto de evaluación.

6. BIBLIOGRAFÍA

- Alcaín, M.D.; San Millán, M.J. (1992). Uso y tendencias de las técnicas bibliométricas en ciencias sociales y humanas a nivel internacional. *Revista Española de Documentación Científica*, v. 16, n. 4. Pág. 30-41.
- Bellavista, J.; Viladiú, C.; Guardiola, E.; Escribano, L.; Grabulós, M.; Iglesias, C. (1991). Evaluación de la investigación social. *REIS*, n. 56. Pág. 219-237.
- Bellavista, J.; Guardiola, E.; Méndez, A.; Bordons, M. (1997). Evaluación de la Investigación. Colección Cuadernos Metodológicos, nº 23. Centro de Investigaciones Sociológicas. Madrid. 117 Pág.
- Camí, J.; Fernández, M.T.; Gómez, I. (1993). La producción científica española en biomedicina y salud. Un estudio a través del Science Citation Index (1986-1989). *Medicina Clínica*, v. 101, n. 19. Pág. 721-731.
- Crouch, D.; Irvine, J.; Martin, B. (1985). Bibliometric Analysis for Science Policy: an evaluation of the United Kingdom's research performance in Ocean Currents and protein crystallography. *Scientometrics*, n. 9, v. 5-6. Pág. 239-267.
- Englisch, H.; Czerwon, H.J. (1990). Quantification of the performance of research units: A simple mathematical model. *Research Policy*, n. 19. Pág. 1-26.
- Escribano, L.; Viladiú, C. (1996). Autoevaluación de las instituciones investigadoras: una perspectiva metodológica en la Universitat de Barcelona. *Política Científica*, n. 46. Pág. 27-40.
- Fernández, A (1999). Evaluación de revistas científicas del campo de la educación. *Revista de la Asociación Interuniversitaria de Investigación Pedagógica*. Edita: Facultad de Ciencias de la Educación de la Universidad de Málaga.
- Fernández, L.; Olea, J.; Collantes, J. (1987). Estadística aplicada al trabajo social. Publicaciones de la Universidad Pontificia Comillas. Madrid. 157 Pág.
- Generalitat de Catalunya. CIRIT (1999). *Avaluació de la Recerca en Humanitats i Ciències Socials*. Barcelona. 81 Pág.
- Ginés, J. (1999). Indicadores y decisiones en las universidades. Ponencia presentada en el Seminario Indicadores en la universidad: información y decisiones. León. 1999.
- Grande, I.; Abascal, E. (1995). Fundamentos y técnicas de investigación comercial. Ed. ESIC. Madrid. 562 Pág.
- Hicks, D.; Crouch, D. (1990). Can bibliometrics measure up? *Physics World*, v. 3, n. 9. Pág. 27-28.

- Leydesdorff, L.; Amsterdamska, O. (1990). Dimension of Citation Analysis. *Science, Technology and Human Values*, v. 15, n. 3. Pág. 305-335.
- Maltrás, B. (1996). Los indicadores bibliométricos en el estudio de la ciencia. *Fundamentos conceptuales y aplicación en política científica*. Universidad de Salamanca. Facultad de Filosofía. Departamento de Filosofía, Lógica y Filosofía de la Ciencia. 293 Pág.
- Martin, B.; Irvine, J. (1983). Assessing basic research. Some partial indicators of scientific progress in radio astronomy. *Research Policy*, n. 12. Pág. 61-90.
- Martin, B.; Irvine, J. (1984). CERN: Past performance and future prospects. *Research Policy*, n. 13. Pág. 183-210.
- Moed, H.F.; Burger, W.J.M.; Frankfort, J.G.; Raan, A.F.J. Van (1985). The application of bibliometric indicators: important field and time-dependent factors to be considered. *Scientometrics*, vol. 8, n. 3-4. Pág. 177-203.
- Moed, H.F.; Burger, W.J.M.; Frankfort, J.G.; Raan, A.F.J. Van (1985b). A comparative study of bibliometric past performance analysis and peer judgement. *Scientometrics*, vol. 8, n. 3-4. Pág. 149-159.
- Moed, H.F.; Raan, A.F.J. van (1985). Critical Remarks on Irvine and Martin's Methodology for Evaluating Scientific Performance. *Social Studies of Science*, n. 15. Pág. 539-547.
- OECD (1994). *The Measurement of Scientific and Technological Activities*. Frascati Manual. París. 261 Pág.
- Reedijk, J. (1998). Sense and nonsense of science citation analyses: comments on the monopoly position of ISI and citation inaccuracies. Risk of possible misuse and biased citation and impact data. *New Journal of Chemistry*, vol. 22, n. 8. Pág. 767-770.
- Sancho, R. (1990). Indicadores bibliométricos utilizados en la evaluación de la ciencia y la tecnología. *Revisión bibliográfica*. *Revista Española de Documentación Científica*, vol. 13, n. 3-4. Pág. 842-865.
- Schubert, A. (1991). Quantitative studies of science. A current bibliography. *Scientometrics*, v. 20, n. 3. Pág. 451-457.
- Valcárcel, M. (1999). La evaluación institucional de la investigación universitaria. *Material Curso de Doctorado*.
- Velho, L. (1986). The "meaning" of citation in the context of a scientific peripheral country. *Scientometrics*, v. 9, n. 1-2. Pág. 71-89.
- Vinkler, P. (1988). An attempt of surveying and classifying bibliometric indicators for scientometrics purposes. *Scientometrics*, v. 13, n. 5-6. Pág. 239-259.
- Vinkler, P. (1998). General Performance Indexes Calculated for Research Institutes of the Hungarian Academy-of-Sciences Based on Scientometric Indicators. *Scientometrics*, v. 41, n. 1-2. Pág. 185-200.