

EVALUACIÓN INSTITUCIONAL DEL SISTEMA UNIVERSITARIO ESPAÑOL
Una visión de la calidad de la evaluación basada en indicadores de
rendimiento

Héctor M. Grad Fuchsel

Héctor M. Grad Fuchsel es Profesor Contratado Doctor en el Departamento de Antropología Social y Pensamiento Filosófico Español de la Facultad de Filosofía y Letras de la Universidad Autónoma de Madrid.

El uso descriptivo de los indicadores estadísticos tiene una tradición relativamente larga en el sistema universitario español (véase, por ejemplo,

Consejo de Universidades, 1987 y 1989, y las órdenes ministeriales de 1988 y 1995 sobre estadística universitaria¹). Esa concepción de la información estadística cambia sustancialmente con la modernización de la gestión universitaria y, especialmente, con el desarrollo de la cultura de la calidad y la rendición de cuentas. Ambos desarrollos han estado relacionados tanto a los proyectos de mejora de la institución como a los proyectos de liberalizar e introducir criterios de mercado en la enseñanza superior desde las últimas décadas del pasado siglo. En ese contexto, la evaluación de la calidad suele servir no sólo para la mejora y la toma de decisiones o el control sino también para gobernar y disciplinar – pues la cultura de la evaluación es usada como mecanismo de transmisión de las políticas universitarias a los valores, las actitudes, la toma de decisiones y la acción de los colectivos y los individuos que forman la institución.

El Plan Nacional de Evaluación de la Calidad de las Universidades (PNECU, 1995)² fue un hito que reflejó y dio impulso a esos procesos, pues promovió el debate sobre los criterios de calidad para las diferentes funciones de la institución universitaria y las metodologías para medirla. Esa experiencia facilitó el desarrollo de los acuerdos técnicos necesarios para decidir indicadores que operativicen los criterios de calidad del sistema universitario. La expresión culminante de esas tenden-

cias puede encontrarse en la Ley Orgánica de Universidades (BOE del 24/12/01), al introducir principios de competición y diferenciación, establecer procesos de evaluación para algunas de las funciones fundamentales de la universidad y crear una agencia independiente de evaluación, la Agencia Nacional de Evaluación de la Calidad de las Universidades (ANECA) a tal efecto.

En ese contexto surge el interés por un sistema de indicadores que, más allá de la mera descripción o monitorización del sistema, sirvan para evaluar los logros (tanto en su evolución histórica como en comparaciones inter-universitarias), para establecer marcos de referencia comunes que faciliten la información y el diálogo con las diferentes audiencias interesadas, y para facilitar la toma de decisiones en la planificación, la gestión y la distribución de recursos en el sistema universitario (ver Sizer, Spee y Bormans, 1992, para una discusión de las posibles funciones de los indicadores de rendimiento en los sistemas universitarios).

Como reseña Escudero (2002), ese interés se refleja en la organización del "Seminario sobre Indicadores en la Universidad: Información y decisiones" (Vidal, 1999) y cristaliza en la elaboración de un Catálogo de indicadores promovida por la Vicesecretaría de Estudios del Consejo de Universidades (Consejo de Coordinación Universitaria, 2001 y 2002). La propuesta recogía indicadores históricos de la estadística universitaria española (ver Consejo de Universidades, 1989, hasta Consejo de Coordinación Universitaria, 2008), adaptaciones de indicadores de rendimiento propuestos por la literatura internacional sobre gestión universitaria (por ejemplo, Cave y otros, 1996; Johnes y Taylor, 1990; Kells, 1993; Taylor y Massy, 1996) y de indicadores establecidos por instituciones internacionales (como *Education at a Glance/Panorama de la Educación* - OCDE, 2008) e indicadores sugeridos por la experiencia y las necesidades de evaluación del PNECU (reflejadas en Vidal, 1999). El sistema propuesto contiene 51 indicadores agrupados en 8 apartados: oferta universitaria, demanda universitaria, recursos humanos, recursos financieros, recursos físicos, procesos, resultados, e información de contexto.

La propuesta de indicadores pudo ser consensuada con la Conferencia de Rectores de las Universidades Españolas (CRUE), que la asume y aplica en los estudios de su Observatorio Universitario (desde Hernández, 2000, hasta Hernández, 2008). Según Escudero (2002), ese proceso permitió definir un sistema flexible de indicadores, comunes para todo el sistema universitario español, que cubren los aspectos relevantes del funcionamiento institucional, tienen claras referencias internacionales, y cuentan con un grado de aceptación que facilitó su implantación sin problemas políticos significativos.

El proceso descrito refleja la existencia de referencias y contribuciones mutuas entre los distintos agentes de gestión y evaluación universitaria, nacionales e internacionales (Consejo de Coordinación Universitaria).

ria, OCDE, ANECA, CRUE) en la elaboración de sus sistemas de indicadores de rendimiento. La redundancia entre las fuentes explica el destacable consenso en los sistemas de indicadores propuestos por esos agentes. El hecho de que esos sistemas de indicadores sirvan de referencia para otros agentes implicados en la evaluación de la enseñanza superior extiende ese consenso a instituciones como la Fundación Conocimiento y Desarrollo - FCYD (Parellada, 2009) e incluso a proyectos de investigación institucional (por ejemplo, Buela-Casal y otros, 2009, o el *Ranking de Universidades del Mundo*, elaborado por el *Institute of Higher Education* de la Shanghai Jiao Tong University, 2008) o periodística (como el ranking *World's Top 200 Universities* elaborado por el Times Higher Education Supplement o *Las 50 carreras* elaborado por el diario El Mundo).

La revisión de los sistemas de indicadores internacionales y nacionales realizada en el *Informe CYD 2005* (Parellada, 2005) muestra que las diferencias entre esos estudios no radica en los indicadores utilizados sino en aspectos como:

Generalidad o especificidad del análisis. Algunos estudios se proponen realizar un análisis integral de las funciones de la institución (como los informes de la Fundación CYD), mientras otros se limitan al análisis de una función específica de la institución (por ejemplo, la investigación en Buela-Casal y otros, 2009, o la enseñanza en el ranking elaborado por El Mundo).

Uso de Indicadores estadísticos y de opinión. Algunos estudios se basan en indicadores estadísticos (como los elaborados por la OCDE o la CRUE), que difícilmente reflejan los aspectos cualitativos de los procesos estudiados. Otros estudios complementan los indicadores de rendimiento con encuestas de opinión que recogen las valoraciones de diferentes audiencias internas o externas (por ejemplo, los elaborados por El Mundo, el *Times Higher Education Supplement*, o el *Wall Street Journal*). Esas valoraciones suelen estar influidas por la imagen social, la reputación y el prestigio de la institución. Por lo tanto, cuando las opiniones reciben un peso significativo (o son la fuente exclusiva, como en el ranking de *Business Week*), estas pseudo-evaluaciones tienden a favorecer a universidades con larga tradición más allá del propio rendimiento de la institución.

Simplificación o complejidad del informe de resultados. Pese a la complejidad de las instituciones universitarias y la multidimensionalidad de sus funciones, los objetivos prácticos de algunos estudios (por ejemplo, facilitar la toma de decisiones por las audiencias interesadas) justifican una presentación resumida de la información en forma de puntuaciones globales únicas (como el *Ranking de Universidades del Mundo* o, en general, los rankings periodísticos).³ Otros estudios optan por reflejar la complejidad de las instituciones y adoptan un enfoque multidimensional⁴ al presentar la información (por ejemplo, los informes *La universidad*

española en cifras publicados por la CRUE o los informes publicados por la Fundación CYD).

Las limitaciones de los sistemas de indicadores utilizados son también compartidas. En primer lugar, los indicadores destacan aspectos cuantitativos del desempeño institucional en detrimento de aspectos cualitativos importantes como la organización de la enseñanza (por ejemplo, el tipo de profesorado en distintos tipos de docencia), el contenido de los programas de enseñanza (por ejemplo, su adecuación a los objetivos o su actualización), o la organización de la investigación.

En segundo lugar, el análisis de los indicadores deja en evidencia serias limitaciones en los sistemas de información universitarios de las universidades españolas, también señaladas por Moneo (2005). El PNECU constató reiteradamente que los sistemas de información de las instituciones universitarias están razonablemente organizados para la gestión cotidiana de los distintos procesos administrativos, pero no están preparados para proporcionar datos útiles para la evaluación, la rendición de cuentas y la toma de decisiones estratégica. Así, la meta-evaluación del PNECU (Consejo de Universidades, 2000) señala la frecuente dificultad para obtener datos suficientes y fiables sobre resultados de la enseñanza y, especialmente, la investigación y la gestión. El informe atribuye esa carencia a la escasa prioridad otorgada desde la gestión universitaria a la eficacia y a la eficiencia de estos procesos y la ausencia de sistemas de información integrados y orientados a la gestión estratégica.

A pesar de los progresos en la materia, los datos sobre algunos aspectos de la universidad aún aparecen sin coincidir en distintas fuentes de información y se carece de información clave sobre el rendimiento en funciones centrales de la institución. En este sentido, es difícil obtener información fiable sobre el rendimiento de la enseñanza, la productividad de líneas o equipos de investigación, o la contribución de las universidades al desarrollo social y cultural de su entorno. Por ejemplo:

a) Aún es posible encontrar datos, no siempre coincidentes, sobre el profesorado de las universidades españolas en cuatro fuentes distintas (Instituto Nacional de Estadística, Consejo de Coordinación Universitaria, Ministerio de Administraciones Públicas y las propias universidades).

b) Los indicadores sobre eficacia de la enseñanza de primer y segundo ciclo no están disponibles o tienen dudosa fiabilidad en un número elevado de casos:

- Para el curso 2002/03 (Parellada, 2005), se carecía de datos sobre tasa de éxito en 25, sobre tasa de rendimiento en 15 y sobre tasa de abandono en 12 de las 47 universidades públicas presenciales estudiadas. En cursos posteriores, la disponibilidad de datos mejora, pero la tasa de graduación deja de ser proporcionada en Parellada (2009).

- En 4 casos de 2002/03 y 6 casos de 2004/05 (Parellada, 2005, 2008), tasas de rendimiento en torno al 60% coinciden con tasas de graduación en el tiempo teórico superiores al 80%. Es decir, más del 80% de los alumnos se graduarían a tiempo a pesar de que no superarían el 40% de los créditos que matriculan. Para un caso extremo se informa una tasa de graduación del 75% con tasa de rendimiento de sólo el 47% (Parellada, 2008, p. 635).

- Siete de las 25 universidades con datos sobre de abandono tendrían una tasa superior al 50% mientras 9 de esas 25 universidades tendrían una tasa inferior al 10%. Estos resultados extremos sugieren la existencia de criterios dispares para contabilizar el abandono más que realidades tan distantes en el sistema universitario.

- En la misma línea, se observan numerosas variaciones interanuales superiores al 10% de las tasas de abandono, rendimiento o graduación (Parellada, 2008, 2009). En tres casos extremos de este tipo, la tasa de graduación se incrementa en más de 20% con respecto a informes anteriores sin que varíe significativamente la tasa de rendimiento (Parellada, 2008, p. 365). Estas mejoras son más probablemente atribuibles a la fiabilidad de los datos que a incrementos tan rápidos y significativos en la eficacia de la enseñanza en instituciones tan complejas.

c) Se carece de datos sobre el impacto de sus actividades de formación continua, inserción laboral y, en general, de extensión y servicios a la comunidad (Parellada, 2007).

Estas observaciones sugieren, en coincidencia con otros distintos estudios (Escudero, 2002; Parellada, 2009), que aún persiste la carencia de bases de datos completas, fiables, homogéneas y homologables sobre las características y los resultados de las instituciones universitarias. En muchas ocasiones, estas limitaciones pueden llegar a poner en cuestión la validez de los indicadores calculados o la viabilidad de esta metodología en las condiciones actuales de los sistemas de información universitaria (cf. Moneo, 2005). Estas limitaciones pueden asimismo explicar la ausencia de análisis evaluativos de los procesos de gestión y los servicios universitarios.

Con esas precauciones, el consenso práctico sobre los indicadores utilizados en los informes *Datos y cifras del Sistema Universitario Español* (Consejo de Coordinación Universitaria, 2008), *La universidad española en cifras* elaborado por la CRUE (Hernández, 2008), *Indicadores bibliométricos de la actividad científica española* elaborado por la Fundación Española para la Ciencia y la Tecnología (Moya Anegón, et. al., 2008) e *Informe CYD* (Parellada, 2009) facilita la descripción de las universidades y su comparación en base a los indicadores de rendimiento en las funciones de enseñanza e investigación. El lector interesado puede encontrar en esos informes abundante información, análisis y opinión

sobre el sistema y las instituciones universitarias en España. En este trabajo nos limitaremos a reseñar algunos aspectos destacados de sus indicadores de rendimiento para las 47 universidades públicas presenciales (cf. Parellada, 2005, 2008, 2009).

En primer lugar, los indicadores disponibles en esos informes permiten caracterizar a las instituciones que conforman el sistema universitario español. De acuerdo su especialización, cabe distinguir entre las universidades politécnicas y generalistas. Las cuatro politécnicas (Madrid, Cataluña, Valencia y Cartagena) centran su oferta de enseñanza en las titulaciones técnicas y destacan por sus recursos en I+D. El resto de las universidades no tiene un perfil especializado (y pueden incluir, en algunos casos, enseñanzas técnicas).

Desde el punto de vista de las magnitudes fundamentales del sistema, seis universidades (Barcelona, Complutense, Granada, País Vasco, Sevilla, y Valencia) concentran una tercera parte del sistema universitario considerando número de matriculados, de estudios ofrecidos, de PDI o PAS, recursos no financieros, ingresos por I+D, etc. Estas instituciones se encuentran entre las universidades tradicionales (su fundación se remonta a los siglos XV y XVI, a excepción de la Universidad del País Vasco, fundada en 1980) y se encuentran localizadas en los mayores centros urbanos y económicos del país. Las universidades autónomas de Barcelona y Madrid destacan en este contexto por la cantidad de estudiantes (autóctonos y extranjeros) de tercer ciclo.

En segundo lugar, los indicadores de rendimiento en la enseñanza disponibles en Parellada (2005, 2008, 2009) muestran una combinación de tasas moderadas de abandono y altas tasas de rendimiento y graduación para el conjunto del sistema. Sin embargo, al margen de las sorprendentes variaciones interanuales antes apuntadas para la misma universidad, esos indicadores presentan una gran variación entre universidades: La tasa de abandono varía entre 4% y 70% en el curso 2002/03, 2% y 30% en el curso 2004/05, y 7% y 33% en el curso 2006/07; la tasa de rendimiento varía entre 47% y 82% en el curso 2002/03, 47% y 81% en el curso 2004/05, y 51% y 82% en el curso 2006/07; la tasa de graduación varía entre 28% y 86% en el curso 2002/03, y 45% y 88% en el curso 2004/05 (finalmente, no se informa el dato para 2006/07). En resumen, estos indicadores reflejan una gran diversidad en la eficacia del proceso de enseñanza entre universidades.

Los indicadores de recursos dedicados a la enseñanza presentan también gran diversidad entre universidades, pero mayor estabilidad interanual. Esta estabilidad se puede atribuir a una larga tradición de criterios que vinculan la dimensión de los recursos físicos y humanos a las necesidades docentes y a una mayor fiabilidad de los datos sobre superficie construida y número de personal y estudiantes. Estos datos tendrían mayor coherencia interna, pese a sus posibles desviaciones de

la realidad, simplemente porque el conteo es un objetivo central de las bases de datos de gestión.

Teniendo en cuenta la existencia de criterios, tanto en la administración como en la toma de decisiones de las propias universidades, que condicionan recursos a las necesidades docentes, es llamativa la gran diversidad entre universidades en los indicadores de recursos humanos dedicados a la docencia. Según Parellada (2005, 2008, 2009), el número de estudiantes por profesor (PDI equivalente a tiempo completo) varía entre 12 y 21 en el curso 2002/03, 11 y 19 en el curso 2004/05, y 10 y 17 en el curso 2006/07. Este rango tan amplio de variación es sorprendente debido a que el desarrollo de la plantilla de profesorado de las universidades fue regida por criterios de financiación homogéneos por décadas, mientras las universidades dependían directamente del MEC, y esos criterios han tenido relativa continuidad tras la transferencia, en fechas no tan lejanas, de las competencias en educación superior a las CCAA. Asimismo, no parece que esa diversidad responda a diferencias en el perfil de estudios ofrecido por las universidades (por ejemplo, que menores ratios de alumnos/PDI respondieran a mayor peso en la institución de enseñanzas que justifiquen dedicación intensiva de profesorado). Por lo tanto, esas diferencias posiblemente reflejen diferentes tradiciones en la política de profesorado de las instituciones.

En la misma línea, el análisis de los indicadores de rendimiento de la actividad investigadora presenta gran estabilidad interanual y amplia diversidad entre universidades. Los principales indicadores para este ámbito (ver Parellada, 2009) se basan en la agregación de los resultados de la evaluación de la productividad científica del profesorado realizada por la Comisión Nacional Evaluadora de la Actividad Investigadora (CNEAI).⁵

Según la memoria de resultados de esa evaluación (CNEAI, 2009b, Tablas 2 y 6), el porcentaje de profesorado funcionario sin haber solicitado u obtenido ningún sexenio desde 1989 se ha reducido levemente a lo largo del tiempo, de 46% hasta 2002 a 42% hasta 2005. Esa tasa global encubre un amplio rango de variación del indicador entre universidades: entre 20% y 73% hasta 2002 y 14% y 68% hasta 2005.⁶ Un indicador del rendimiento investigador complementario del anterior es el número de sexenios por profesor. Este indicador presenta una variación interanual y entre universidades similar al anterior, entre 0,45 y 2,09 hasta 2002 y 0,59 y 2,31 hasta 2005 (Parellada, 2005, 2009). Teniendo en cuenta el medio sexenio transcurrido entre ambas fechas, la mejora parece reflejar el envejecimiento de una plantilla de PDI sin variación sustancial en la productividad reconocida en investigación. La gran diversidad entre universidades, y la relativa estabilidad de los resultados de cada institución parece indicar la existencia de diferencias sustanciales en los recursos físicos y humanos disponibles así como en la cultura institucional en el ámbito de la investigación.

Más allá de factores como la proporción de profesorado doctor y la edad de esa plantilla, el perfil (en este caso, investigador) de la institución nuevamente parece contribuir a las diferencias entre universidades en el rendimiento de investigación. Este efecto surge de las grandes diferencias en el porcentaje de sexenios concedidos entre las comisiones de evaluación de los distintos campos científicos de la CNEAI. Las memorias de resultados de la evaluación de la actividad investigadora del profesorado (CNEAI, 2009a, Gráficos 5.1 a 5.11; CNEAI, 2009d, Tabla 2) muestran diferencias sistemáticas en la concesión de sexenios desde 1989. Por un lado, al profesorado de los campos 01-Matemáticas y Física, 02-Química y 03-Biología Celular y Molecular se les han concedido entre 79%, 80% y 90% de los sexenios solicitados entre 1989 y 2004, y llegan hasta 92%, 96% y 97% de los sexenios solicitados en 2005, respectivamente. Por otro lado, al profesorado de los campos 07-Ciencias Sociales, Políticas, del Comportamiento y de la Educación y 08-Ciencias Económicas y Empresariales se han concedido, respectivamente, 59% y 58% hasta 2004, con una leve mejora de 65% y 68% en 2005. Estos campos se han aproximado a la media de los sexenios concedidos en sólo dos de las 17 convocatorias de la CNEAI desde 1989 (ver CNEAI, 2009a, Gráficos 5.7 y 5.8), mientras el resto de los campos, tanto experimentales (como Ciencias de la Naturaleza) como Jurídicas, de la Salud, Tecnología e Ingeniería y Humanidades oscilan en torno a la media de los sexenios concedidos en esas convocatorias.

La Comisión Nacional de Evaluación de la Actividad Investigadora del profesorado ha realizado ingentes esfuerzos para desarrollar procedimientos de evaluación armonizados y criterios específicos adaptados a cada campo científico.⁷ Sin embargo, la regularidad de las tendencias apuntadas indicaría un ajuste casi perfecto entre la práctica y los criterios aplicados a la evaluación de algunos campos experimentales, por un lado, y una clara brecha entre la práctica de la investigación en Ciencias Sociales y los criterios aplicados a su evaluación. Esta brecha podría atribuirse a la falta de ajuste de la investigación en estos campos a los criterios de calidad requeridos a la productividad científica. La continuidad de la brecha sugeriría una especial rigidez de la comunidad de científicos sociales de nuestro país para adaptar su producción y pautas de comunicación a esos criterios a lo largo de casi dos décadas. La persistencia del desajuste contradice la expectativa usual de adaptación expresada, por ejemplo, en el síndrome de *publicar o perecer* (cf. Moya Anegón et. al., 2008, p. 30). Esto sugiere la posibilidad de que la brecha se mantenga pese a la evolución adaptativa, por el cambio hacia criterios más estrictos.

Una lectura en otro sentido destacaría la especial coincidencia entre cultura y prácticas de investigación y criterios de calidad de la producción científica en los campos experimentales, frente a la relativa inadecuación de la definición y/o la aplicación de los criterios a la cultura y las

prácticas de investigación en las Ciencias Sociales de nuestro sistema universitario. En cualquier caso, el resultado parece ser discriminatorio hacia la investigación en Ciencias Sociales. A nivel institucional, estos sesgos afectarían a los indicadores de rendimiento de la actividad investigadora basados en los complementos de productividad concedidos por la CNEAI dependiendo del peso de cada campo científico (plantilla de PDI y su actividad científica) en cada universidad – en perjuicio de las instituciones con perfil científico más social.

Más allá de las limitaciones técnicas apuntadas, el análisis precedente de los indicadores de rendimiento del sistema universitario público brinda la oportunidad de discutir algunos aspectos problemáticos de su utilización en nuestro país. La primera cuestión relevante es el cambio en los objetivos de la evaluación institucional. Los objetivos, el diseño y la implantación del PNECU (1995) estaban fundamentalmente centrados en la promoción de procesos de mejora auto-regulados, donde la elaboración de indicadores venía a sistematizar datos para la reflexión por las unidades y la institución evaluada.

Grad y Hernández (1997) constataban, ya en aquel momento, que sectores significativos de la administración y las instituciones universitarias compartían tanto los objetivos de innovación y calidad para el sistema universitario como el interés en la debida rendición de cuentas a la sociedad que lo sustenta. La lógica de esos procesos apuntaba al establecimiento de políticas de incentivo de la calidad que vincularan, en algún grado, la financiación con los resultados de evaluación (de Miguel, Mora y Rodríguez, 1991; Mora, 1991 – para procesos similares en nuestro entorno, ver CVCP, 1985; Dahllöf et al., 1990; Kogan, 1989; McGuinness, 1995; Shattock, 1995). Efectivamente, el papel de los indicadores de rendimiento se veía reforzado por la propuesta de modelos de financiación universitaria que incorporaban criterios de eficiencia y calidad como base para la financiación diferencial de las universidades (Consejo de Universidades, 1994; Mora y Villareal, 1995).

El II Plan de Calidad de las Universidades⁸, acompañante de la LOU, significó un alejamiento de las intenciones originales del PNECU hacia objetivos orientados al control y la toma de decisiones por parte de las administraciones competentes. En coherencia con el cambio en los objetivos (ver Kells, 1992, para un análisis de la relación entre distintas posiciones del modelo evaluativo en el continuo control-mejora y las consecuentes opciones metodológicas), los esfuerzos pasan a invertirse en el Catálogo de indicadores y en los planes de financiación por objetivos referenciados en los mismos indicadores acordados por las universidades con sus respectivas administraciones autonómicas.⁹

Otro proceso de deriva de objetivos se ha producido con la evaluación de la productividad científica del profesorado. Esta evaluación fue establecida por el Real Decreto 1086/89, de 28 de agosto (BOE de 9 de septiembre) sobre retribuciones del profesorado. El procedimiento se

implanta en aplicación del artículo 45.3 de la Ley de Reforma Universitaria, que dispuso la evaluación periódica de la actividad docente e investigadora del profesorado universitario. El Real Decreto sobre retribuciones establece que la evaluación de los méritos docentes es competencia exclusiva de cada Universidad, mientras que la evaluación de productividad científica es competencia estatal y se encomienda a la CNEAI. Ambas evaluaciones son estrictamente individuales y voluntarias en tanto que tienen por objeto la concesión de los complementos salariales respectivos. En coherencia con sus fines originales, la evaluación de la actividad investigadora se centra en la productividad individual, tanto en la participación voluntaria de los interesados como en la definición de los datos, en los aspectos analizados y en los criterios aplicados para el reconocimiento del mérito investigador.

Sin embargo, con posterioridad, los resultados agregados de esas evaluaciones se han convertido en fuente de los principales indicadores de rendimiento institucional en investigación. Esta deriva en el uso de esta evaluación implica limitaciones tanto en la información disponible como la ausencia de información sobre otras dimensiones de la calidad de la investigación que deberían ser tomados en cuenta para una evaluación institucional.

Por un lado, la voluntariedad de la evaluación sesga la base de la información dependiendo de la actitud de los interesados con respecto al complemento de productividad y el procedimiento para su concesión. Por esta razón, una parte del profesorado con productividad científica sustancial puede haber optado por no solicitar la evaluación por reticencia al procedimiento de evaluación. En consecuencia, la abstención mezclaría a profesorado que no se presenta por falta de méritos con profesorado contestatario pero productivo desde el punto de vista científico. La tasa de presentación (o de abstención) sería indicador válido de una cultura institucional comprometida con la investigación en el primer caso pero no en el segundo. Por lo tanto, esta mezcla cuestionaría la validez de la tasa de participación (o de abstención) como indicador de la actividad investigadora.

En consecuencia, los indicadores de rendimiento basados en la evaluación de la CNEAI estarían subestimando la productividad científica real de las instituciones. Esta estimación, sin embargo, podría servir para la comparación entre instituciones en caso de que el sesgo se distribuyera homogéneamente en todos los campos científicos. Sin embargo, el porcentaje de profesorado que nunca se ha presentado a la evaluación de la CNEAI hasta 2005 varía del 2% en el campo 02-Química al 34% en el campo 08-Ciencias Económicas y Empresariales (sin incluir TEU, CNEAI, 2009c, Tabla 8). Es decir, la reticencia y la disposición a participar en este proceso estarían relacionadas con diferencias en cultura y prácticas de investigación entre los distintos campos científicos. En conclusión, las comparaciones en este tipo de indicadores podrían

perjudicar, de forma similar a la (in)adecuación entre práctica científica y criterios para su evaluación, a las instituciones con mayor peso de campos reticentes a ser evaluados.

Por otro lado, una evaluación centrada en la productividad individual puede ser apropiada y válida para la decisión del incentivo salarial, pero puede tener efectos perversos sobre la calidad de la investigación al nivel de la institución. Por ejemplo, ese incentivo puede reforzar las tendencias individualistas y la competencia entre investigadores. Este refuerzo puede mejorar la productividad en campos donde la investigación sea una actividad más individual pero minaría la colaboración en campos que la requieran. De hecho, los informes de la Fundación Española para la Ciencia y la Tecnología muestran que desde principios de los años noventa hasta 2006, la producción del sistema científico español se ha duplicado con una tasa promedio anual de incremento cercana al 9% en el número de documentos producidos. En línea con este crecimiento, España ocupa el puesto décimo en el *ranking* de número de documentos y undécimo en el número de citas pero sólo el puesto 36 en la media de citas por documento, en la *Web of Science* del Institute for Science Information (Moya Anegón et. al., 2008). Sin embargo, la productividad medida por el ratio de documentos producidos ha descendido porque el número de investigadores crece más rápidamente que el número de documentos producidos (Moya Anegón et. al., 2007). Este descenso puede atribuirse a la incorporación de PDI joven (Moya Anegón et. al., 2007), pero no debería descartarse el efecto de un agotamiento del esfuerzo (voluntarista) de una generación de investigadores por incorporarse al *mainstream* internacional de su campo, o los obstáculos estructurales para consolidar grupos de investigación con masa crítica y apoyos institucionales que faciliten su continuidad.

En la misma línea, el incentivo a la productividad puede inducir al abandono de tareas no recompensadas pero necesarias para el funcionamiento del sistema, como la formación de investigadores (dirección de tesis), la responsabilidad editorial o la participación en procesos de evaluación por pares. Finalmente, al centrarse en la actividad individual, los indicadores derivados de la evaluación de la CNEAI no toman en consideración aspectos colectivos y organizacionales relevantes para el rendimiento y la toma de decisiones institucionales, como la construcción de grupos y redes de investigación, y la productividad de grupos y líneas de investigación. En definitiva, los efectos institucionales de los indicadores diseñados para medir la productividad individual deberían ser examinados con mayor atención y estos indicadores ser complementados por otros diseñados para medir las dimensiones colectivas y organizacionales del rendimiento en investigación.

En conclusión, el consenso en los sistemas de indicadores de rendimiento parece facilitar una evaluación comparativa de las instituciones del sistema universitario español. Sin embargo, este análisis ha muestra-

do que debilidades básicas de los sistemas de información condicionan la fiabilidad y la validez de los indicadores obtenidos y dejan aspectos centrales del funcionamiento institucional fuera del ámbito de la evaluación. Este análisis ha mostrado, asimismo, algunos aspectos como la evaluación de la productividad científica donde los sesgos del proceso evaluativo pueden generar situaciones discriminatorias entre campos e instituciones científicas. Finalmente, se han presentado algunos ejemplos donde el cambio en los usos de la evaluación hace que procedimientos coherentes con sus fines originales puedan tener efectos perversos, o incluso contradictorios, en su nueva aplicación (una crítica más integral de estos efectos puede encontrarse en Bruneau y Savage, 2002). Por lo tanto, aceptando su potencial contribución a la transparencia informativa, la rendición de cuentas y la toma de decisiones, muchas son las mejoras y los ajustes aún pendientes para lograr un sistema de indicadores de rendimiento útil, fiable y válido para la evaluación institucional de las universidades en España.

Nota: Una versión previa de este artículo fue presentada en el Seminario "Evaluación de la Calidad vs. Calidad de la Evaluación de la Universidad", Departamento de Psicología Social, Universidad del País Vasco, San Sebastián, 5 de Diciembre de 2007.

Referencias

- Bruneau, W.A.-Savage, D.C. (2002): *Counting out the scholars: how performance indicators undermine universities and colleges*. Toronto: James Lorimer & Company, Canadian Association of University Teachers
- Buela-Casal, G.-Bermúdez, M.P.-Sierra, J.C.-Quevedo-Blasco, R.-Castro, A. (2009): Ranking de 2008 en productividad en investigación de las universidades públicas españolas. *Psicothema*, 21, 309-317.
- Cave, M.-Hanney, S.-Henkel, M.-Kogan, M. (1996): *The use of performance indicators in higher education: The challenge of the quality movement* (3rd Ed.). London: Jessica Kingsley Publishers (Higher Education Policy Series 34).
- Comisión Nacional Evaluadora de la Actividad Investigadora (2009a): *Memoria de los resultados de las evaluaciones realizadas de 1989 a 2005 (Profesores de Universidad)*. Madrid: Ministerio de Ciencia e Innovación. [Disponible en <http://www.micinn.es/ciencia/cneai/>].
- Comisión Nacional Evaluadora de la Actividad Investigadora (2009b): *Memoria sobre la situación del profesorado numerario con respecto a los sexenios de investigación. Año 2005*. Madrid: Ministerio de Ciencia e Innovación. [Disponible en <http://www.micinn.es/ciencia/cneai/>].
- Comisión Nacional Evaluadora de la Actividad Investigadora (2009c): *Memoria sobre la situación del profesorado numerario con respecto a los sexenios de investigación. Año 2005 (sin contabilizar los Profesores Titulares de Escuela Universitaria)*. Madrid: Ministerio de Ciencia e Innovación. [Disponible en <http://www.micinn.es/ciencia/cneai/>].
- Comisión Nacional Evaluadora de la Actividad Investigadora (2009d): *Resultados de la evaluación de 2005 (Profesores de Universidad)*. Madrid: Ministerio de Ciencia e Innovación. [Disponible en <http://www.micinn.es/ciencia/cneai/>].

- Consejo de Coordinación Universitaria(2001): *Borrador del Catálogo de Indicadores*. Madrid: Ministerio de Educación, Cultura y Deporte.
- Consejo de Coordinación Universitaria(2002): *Catálogo de Indicadores del Sistema Universitario Público Español*. Madrid: Ministerio de Educación, Cultura y Deporte.
- Consejo de Coordinación Universitaria(2008): *Datos y cifras del Sistema Universitario Español*. Curso 2008/2009. Madrid: Ministerio de Educación, Política Social y Deporte, Subdirección General de Información y Publicaciones.
- Consejo de Universidades(1987): *Plan de Estadística Universitaria*. Madrid: Secretaría General del Consejo de Universidades, Vicesecretaría de Estudios.
- Consejo de Universidades(1989): *Anuario de estadística universitaria: 1988*. Madrid: Ministerio de Educación y Ciencia, Centro de Publicaciones.
- Consejo de Universidades(1994): *Informe sobre la financiación de las Universidades*. Madrid: Secretaría General del Consejo de Universidades.
- Consejo de Universidades(2000): *Informe Anual de la Segunda Convocatoria del Plan Nacional de Evaluación de la Calidad de las Universidades*. Madrid: Consejo de Universidades.
- CVCP(1985): *Report of the Steering Committee for Efficiency Studies - Jarrat Report*. London: Committee of Vice-Chancellors and Principals.
- Dahllöf y otros (June, 1990): *Report of the IMHE Study Group on Evaluation in Higher Education*. Paris: OCDE Tenth General Conference of Member Institutions.
- De Miguel, M; Mora, J-G; Rodríguez,S. (Eds.)(1991): *La evaluación de las instituciones universitarias*. Madrid: Consejo de Universidades.
- Escudero,T.(2002): *Indicadores: Información estratégica para el sistema universitario español*. En J. Hernández (Dir.), *La universidad española en cifras: Información académica, productiva y financiera de las universidades españolas, Año 2000. Indicadores Universitarios, Curso Académico 2000/2001* (pp. 177-199). Madrid: Conferencia de Rectores de las Universidades Españolas (Observatorio Universitario).
- Grad,H.M.-Hernández,J.M.(1997): *Prospectiva de la Evaluación Institucional y de los Planes de Calidad*. En P. Apodaca y C. Lobato (Eds.), *Calidad en la Universidad: Orientación y Evaluación* (pp. 252-270). Barcelona: Laertes.
- Hernández,J.(Dir.)(2000): *La universidad española en cifras: Información académica, productiva y financiera de las universidades españolas. Curso Académico 1998-1999*. Madrid: Conferencia de Rectores de las Universidades Españolas.
- Hernández,J.(Dir.)(2008): *La universidad española en cifras(2008): Información académica, productiva y financiera de las universidades españolas, Año 2006. Indicadores universitarios, Curso Académico 2006-2007*. Madrid: Conferencia de Rectores de las Universidades Españolas.
- Johnes,J.-Taylor,J.(1990): *Performance indicators in higher education*. London: SRHE and Open University Press.
- Kells,H.R.(1992): *Purposes and means in higher education evaluation. Higher Education Management, 4 (1), 91-102*.
- Kells,H.R.(Ed.)(1993): *The development of performance indicators for higher education. A compendium for twelve countries (2nd Ed.)*. Paris: Organization for Economic Co-Operation and Development.
- Kogan,M.(1989): *The evaluation of higher education: An introductory note*. En M. Kogan (Ed.), *Evaluating higher education* (p. 11-25). London: Jessica Kingsley (Higher Education Policy Series 6).
- McGuinness,A.C.(1995): *The changing relationships between the states and universities in the United States. Higher Education Management, 7, 263-279*.

- Moneo, C. (2005): Los retos estadísticos del sistema universitario español. En M. Parellada (Dir.) (2005). *Informe CYD 2005: La contribución de las universidades españolas al desarrollo* (Cap. 5: Indicadores y rankings universitarios. Panorámica internacional y universidades públicas españolas, pp. 288-290). Barcelona: Fundación Conocimiento y Desarrollo.
- Mora, J.G. (1991): *Calidad y rendimiento en las instituciones universitarias*. Madrid: Consejo de Universidades.
- Mora, J.-G.-Villareal, E. (1995): Un modelo para la financiación de las universidades públicas. En E. Oroval (Ed.): *Planificación, evaluación y financiación de sistemas educativos* (pp. 175-199). Madrid: Asociación de Economía de la Educación y Editorial Civitas.
- Moya Anegón et. al. (2007): *Indicadores bibliométricos de la actividad científica española (1990-2004)*. Madrid: Fundación Española para la Ciencia y la Tecnología.
- Moya Anegón et. al. (2008): *Indicadores bibliométricos de la actividad científica española (2002-2006)*. Madrid: Fundación Española para la Ciencia y la Tecnología.
- OCDE (2008): *Education at a Glance 2008: OECD Indicators*. Paris: Organization for Economic Co-Operation and Development, Programme on Institutional Management in Higher Education (IMHE).
- Parellada, M. (Dir.) (2005): *Informe CYD 2005: La contribución de las universidades españolas al desarrollo* (Cap. 5: Indicadores y rankings universitarios. Panorámica internacional y universidades públicas españolas, pp. 253-294). Barcelona: Fundación Conocimiento y Desarrollo.
- Parellada, M. (Dir.) (2007): *Informe CYD 2006: La contribución de las universidades españolas al desarrollo* (Cap. 4: La contribución de las universidades al desarrollo. Una panorámica de los indicadores existentes, pp. 230-265). Barcelona: Fundación Conocimiento y Desarrollo.
- Parellada, M. (Dir.) (2008): *Informe CYD 2007: La contribución de las universidades españolas al desarrollo*. Barcelona: Fundación Conocimiento y Desarrollo.
- Parellada, M. (Dir.) (2009): *Informe CYD 2008: La contribución de las universidades españolas al desarrollo*. Barcelona: Fundación Conocimiento y Desarrollo.
- Shanghai Jiao Tong University (2008): *Academic Ranking of World Universities*. Shanghai: Institute of Higher Education, Shanghai Jiao Tong University [Disponible en <http://www.arwu.org/>].
- Shattock, M.L. (1995): The university of the future. *Higher Education Management*, 7, 157-163.
- Sizer, J.-Spee, A.-Bormans, R. (1992): The rôle of performance indicators in higher education. *Higher Education*, 24, 133-155.
- Taylor, B.E.-Massy, W.F. (1996): *Strategic indicators for higher education*. Princeton: Peterson's.
- Vidal, J. (Coord.) (1999): *Indicadores en la Universidad: Información y decisiones*. Madrid: Consejo de universidades (Plan Nacional de Evaluación de la Calidad de las Universidades), Ministerio de Educación y Cultura, Centro de Publicaciones.

¹ Orden del Ministerio de Educación y Ciencia de 25 de mayo de 1988 por la que se establece la información estadística que sobre el alumnado deben suministrar las universidades (BOE de 28 de mayo), cuya modificación por Orden de 1 de agosto de 1995 (BOE de 9 de agosto).

² Real Decreto 1947/1995, de 1 de diciembre (BOE de 9 de diciembre), por el que se establece el Plan Nacional de Evaluación de la Calidad de las Universidades.

³ Estas simplificaciones necesariamente engloban aspectos diversos o incluso claramente heterogéneos de la institución. Por lo tanto, su elaboración plantea los problemas metodológicos derivados de la integración de datos mediante la ponderación de aspectos. Esta ponderación plantea la cuestión de qué pesos asignar a funciones disímiles e, incluso, qué estrategias metodológicas usar para resolver ese problema (desde la decisión por expertos en el US News & World Report, hasta encuestas de opinión sobre la centralidad o importancia de los aspectos integrados en Buena-Casal y otros, 2009). En consecuencia, las puntuaciones elaboradas pueden carecer de significado intrínseco. Con esa precaución, su valor es instrumental y su validez depende de que la aditividad de los aspectos permita ordenar instituciones.

⁴ Este enfoque no es necesariamente contradictorio con el anterior pues podría incluir ranking por aspectos que tendrían mayor validez y serían más informativos que los ranking globales, como en el *Informe CYD 2008* (Fundación CYD, 2009).

⁵ Esta evaluación fue establecida por el Real Decreto 1086/89, de 28 de agosto (BOE de 9 de septiembre) sobre retribuciones del profesorado.

⁶ Este indicador puede ser apropiado para representar el rendimiento global de la universidad en investigación. Sin embargo, parte de los resultados más negativos y de la variación entre universidades es debida a la inclusión en el indicador del profesorado Titular de Escuela Universitaria, que puede no tener reconocida ni requerida la función investigadora. Un indicador más ajustado de la productividad científica del profesorado se puede obtener excluyendo a esa categoría del cálculo. De esta manera, el porcentaje de profesorado funcionario (excluyendo TEU) sin haber solicitado u obtenido ningún sexenio se reduciría de 32% en 2002 al 28% en 2005 mientras los valores superiores del indicador se recortan, por lo que los rangos entre universidades serían 18-52% en 2002 y 15-47% en 2005 (CNEAI, 2009c, Tablas 2 y 6). Se produce un recorte significativo en los valores con mayor porcentaje de profesorado sin sexenios. Este recorte reflejaría, en parte, el mayor peso del profesorado TEU en las universidades con menor productividad investigadora.

⁷ Ver, por ejemplo, la Resolución de 11 de noviembre de 2008 (BOE del 22 de noviembre) de la Presidencia de la Comisión Nacional Evaluadora de la Actividad Investigadora, por la que se establecen los criterios específicos en cada campo de evaluación.

⁸ Real Decreto 408/2001 del 20 de abril 2001 por el que se establece el II Plan de Calidad de las Universidades.

⁹ Por otra parte, con la implantación del Espacio Europeo de Educación Superior, la responsabilidad sobre la evaluación institucional es transferida hacia las propias universidades. En ese contexto, la cultura de la evaluación adquiere la función de gobernar y disciplinar – pues es usada como mecanismo para imponer las políticas universitarias derivadas de la interpretación de los Acuerdos de Bolonia a los valores, las actitudes, las decisiones y la acción de los colectivos y los individuos que forman la institución.

En este escenario, la función básica de la ANECA pasa a ser de control: Por un lado, la certificación de los sistemas internos de garantía de la calidad (mediante el Programa AUDIT). Por otro lado, la verificación y acreditación de los títulos universitarios oficiales (Programa VERIFICA) en aplicación del Capítulo 6 del Real Decreto 1393/2007, de 29 de octubre, que establece el marco normativo para la ordenación de las enseñanzas universitarias oficiales, de conformidad con lo previsto en el Título VI de la Ley Orgánica 4/2007 de Universidades.