

LA EVALUACIÓN DE LA ACTIVIDAD INVESTIGADORA DEL PROFESORADO UNIVERSITARIO: CONTRASTE ENTRE INTENCIONES Y HECHOS

Francisco M. Aliaga Abad*

Uno de los intentos más importantes en la modernización de la Universidad española a finales del siglo XX fue el intento de implicarla en la corriente del *accountability*, es decir, en la necesidad de dar cuenta de sus actividades a la sociedad de la que forma parte, a la que sirve y de la que, finalmente, proceden los recursos que la mantienen. El objetivo era doble. Por un lado, permitir evaluar sus resultados de manera directa y, por otro, la de posibilitar los incentivos individualizados. La rígida estructura funcional propia de las universidades públicas en España no ofrecían incentivos relevantes para su personal, ya que los emolumentos de los profesores no dependían en absoluto de la actividad realizada ni mucho menos del rendimiento alcanzado.

Dentro de este modelo general se desarrolló un modelo de evaluación de las actividades del profesorado universitario, al menos del funcionariado (aunque en los últimos años este requisito se ha flexibilizado ligeramente para adaptarse a las nuevas modalidades contractuales). Vamos a intentar presentar, a modo de evaluación, un análisis global de este proceso, haciendo particular hincapié en la vertiente investigadora. Desgraciadamente la parte de evaluación docente del profesorado universitario se ha minusvalorado y con frecuencia se ha convertido en un «café para todos» que la aleja tanto de su intención original como de su función incentivadora y, en última instancia, prestigiadora.

Ya en la Ley de Reforma Universitaria (LRU) del año 1983 se señalaba (en su artículo 45) que las universidades debían incluir en sus estatutos disposiciones sobre «los procedimientos para la evaluación periódica del rendimiento docente y científico del profesorado». Sin embargo, dicha disposición apenas se desarrolló en la práctica, con al-

* Francisco Aliaga Abad es Dr. en Psicología. Trabaja como profesor titular en el Departamento de Métodos de Investigación y Diagnóstico en Educación de la Universitat de València. Sus líneas de investigación son: Metodología de investigación, Evaluación educativa y Tecnologías de información en educación. E-mail: Francisco.Aliaga@uv.es

gunas pocas excepciones, entre las que cabe destacar las de las universidades Autónoma de Madrid (bajo la dirección del profesor Tejedor) y la de Valencia (desarrollado por Jornet y Suárez), que ya en 1987 pusieron en marcha unos primeros intentos de evaluación de la docencia de profesorado.

Para dar un nuevo impulso al tema de manera global se publicó en 1989 el Real Decreto 1.086, que regulaba las retribuciones del profesorado universitario, incluyendo como elemento estructural de las mismas dos complementos ligados a los resultados de la evaluación realizada a los profesores, uno de ellos vinculado a los méritos docentes (que pasaron a denominarse en el argot universitario *quinquenios*, pues éste era el período evaluado cada vez) y el otro, a la productividad investigadora (los famosos *sexenios*). El objetivo declarado era «incentivar la labor docente e investigadora del personal docente universitario, mediante la evaluación individualizada de estas actividades». Así pues, los dos principios rectores eran los de procurar la mejora de la docencia y la investigación, por un lado, y por el otro individualizar el trato a dicho personal, ambas pretensiones ciertamente loables.

El ámbito de la evaluación docente se adjudicó a las propias universidades, que por su cuenta han ido desarrollando métodos más o menos semejantes basados fundamentalmente en las opiniones del alumnado (Muñoz Cantero, Ríos de Deus y Abalde, 2002). El que dicho tipo de procedimiento esté procurando la pretendida «mejora de la docencia» es una cuestión cuanto menos discutible. En muchos casos se ha convertido en un trámite con enfoque fundamentalmente burocrático (y con frecuencia indiscriminado) para conseguir un pequeño aumento de sueldo para el profesorado universitario funcionario. Así, por ejemplo, la normativa actualmente vigente en la Universitat de València, que data del año 2001, añade un principio general básico para configurar el proceso de evaluación de la docencia de los profesores: «el ahorro en el proceso de la información», ya que evaluar a todos los profesores en todas las asignaturas de manera anual supone, al parecer, un esfuerzo excesivo para la institución. Por ello se restringe la necesidad de superar evaluaciones positivas a un año de cada cinco, con tal de que haya un historial previo de haber superado otras, lo que es casi universal. Se asume entonces, de manera quizá algo atrevida, que si alguien es un docente esforzado y reconocido durante uno de los cursos de cada quinquenio, también lo será el resto del tiempo. A nuestro entender, esta burocratización extrema del proceso ha hecho perder buena parte del sentido (y el impulso) que en su momento pudiera tener la normativa para procurar la mejora de la actividad docente. El resultado ha sido que la actividad docente se percibe, de manera generalizada, como menos valorada que otras actividades propias del profesorado universitario, ya que su reconocimiento va vinculado sólo de manera muy marginal con la actividad efectivamente realizada de los profesores. Un análisis detallado de los principios, la casuística, las aplicaciones y las consecuencias de este proceso de evaluación de la docencia exceden las pretensiones de este escrito, por lo que nos centraremos más bien en la otra dimensión, la de la evaluación de la actividad investigadora.

A diferencia del proceso de evaluación de la docencia, el Real Decreto 1.086/1989 establece que la evaluación de la actividad investigadora del profesorado universitario (y también del personal investigador del CSIC) habrá de llevarla a cabo una instancia externa a las universidades, la Comisión Nacional de Evaluación de la Actividad Investigadora (CNEAI), dependiente directamente del Ministerio de Educación. Resulta cuando menos llamativa esta distribución de atribuciones, ya que parece mostrar una cierta desconfianza del legislador a la capacidad de cada universidad de organizar y gestionar un proceso de evaluación de la investigación de suficiente calidad, a diferencia de lo que ocurre con la evaluación de la docencia. De hecho se producía, al asignar las competencias de evaluación de la actividad investigadora a la CNEAI, una clara contradicción con lo indicado expresamente en el artículo 45 de la LRU, norma de rango superior que, teóricamente, debería haber prevalecido.

La regulación de la evaluación de la investigación de los docentes universitarios tiene una enorme trascendencia en España, ya que son ellos los que realizan el grueso de la investigación que se realiza en nuestro país. Se trata por tanto de una herramienta muy poderosa en manos de la administración que puede permitir orientar, e incluso condicionar, la actividad no sólo de muchos profesionales implicados, sino incluso la eficacia o dirección de la investigación a nivel nacional. A nuestro entender es ésta una medida sana que somete la actividad del profesorado universitario a una cierta «dación de cuentas» de su actividad (si bien voluntaria) a cambio de unos ciertos, aunque modestos, incentivos. Además, y aunque no era en absoluto ésa la finalidad inicial de todo este proceso de evaluación, en tiempos recientes se ha hecho depender, en la práctica, el desarrollo de la carrera profesional o las posibilidades de participar en una gran cantidad de actividades académicas, de los resultados que cada profesor vaya obteniendo en estos procesos evaluativos. Por todo ello los objetivos iniciales se han ido viendo desbordados y sobrepasados por la propia dinámica institucional.

La normativa específica que regula el proceso de evaluación de la actividad investigadora ha ido desarrollándose paulatinamente a lo largo de cada una de las convocatorias, estableciendo con frecuencia modificaciones relevantes que denotan la falta de estabilidad del sistema. Sin embargo, sí hay algunas pocas características que se han mantenido estables a lo largo del tiempo. La principal de todas ellas es que la evaluación se produce por parte de expertos (aunque los criterios para determinar qué sea tal cosa han variado con el tiempo), siguiendo la modalidad de evaluación que se suele conocer como *peer review*. Ciertamente no era ésta la única opción disponible para el legislador, pero sí la única que finalmente se ha desarrollado y aplicado.

Otra característica relevante y que se ha mantenido a lo largo del tiempo en el proceso es que la evaluación por expertos se realiza tomando como base indicios indirectos o secundarios de calidad de las investigaciones. Por tanto, la evaluación que se realiza es de carácter casi exclusivamente formal, dado que no se evalúan (ni siquiera se requiere que se presenten) las publicaciones que han de someterse para su evaluación. Por el con-

trario, se suele analizar más la solvencia del medio en el que se publican los trabajos de investigación que los trabajos en sí. Ha de señalarse además que este diseño se realizó en una época en la que los estudios sobre la calidad, solvencia o difusión de los medios de difusión de información científica españoles, europeos o latinoamericanos eran prácticamente inexistentes, lo que dio lugar a un marcado subjetivismo en las decisiones iniciales de los comités de la CNEAI. Ni el legislador ni las instancias del poder ejecutivo encargadas de gestionar este proceso han considerado importante desarrollar herramientas que evalúen dichas características de calidad y difusión de las publicaciones, y han tenido que solventar parcialmente el problema en épocas muy recientes con iniciativas particulares, tales como los estudios de prestigio entre los investigadores funcionarios realizados por personal del CINDOC (con Alcaín y Román a la cabeza) o el reciente desarrollo, en condiciones muy precarias, de proyectos como In-RECS o In-RECI. Por ello es el propio profesor-investigador el que, apenas sin apoyo instrumental y con muy pocas referencias de qué es lo que se valora realmente, ha de acreditar la «calidad» de sus publicaciones.

En los primeros años la mayoría de críticas al proceso realizado por la CNEAI (De Miguel, 1997) provenía de la falta de contextualización de la evaluación: se establecieron normas generales que debían cumplir todas las aportaciones, independientemente del área en el que se hubiesen desarrollado. Encontramos así la exigencia de que algunas de las aportaciones sometidas por los investigadores a evaluación dentro de cada sexenio fuesen artículos publicados en revistas indexadas en las bases de datos del *Institute of Scientific Information* (ISI), creado por el padre de la Bibliometría, Eugene Garfield. Por lo tanto, estas bases de datos se convertían en el único referente explícitamente señalado de la calidad de las publicaciones científicas. Sin embargo, las bases de datos del ISI han sido reiteradamente criticadas por diversos autores por muy variadas razones (Cf. Aliaga y Orellana, 1999):

- Sus bases de datos tienen un porcentaje de errores y confusiones muy alto que se ha calculado en torno al 25% del total.
- Los temas de interés «local» (característicos de las Ciencias Sociales) están extraordinariamente muy poco representados, incluso en aquellos casos en los que generan investigación de gran calidad y fuerte impacto social. Hay que tener en cuenta, además, que el concepto de «local» sólo se aplica, en las bases de datos del ISI, a investigaciones de fuera de los países anglosajones, lo que fomenta un evidente «colonialismo cultural».
- Hay un sesgo lingüístico muy fuerte en las bases de datos del ISI a favor de las publicaciones editadas en inglés, siendo claramente marginadas las publicadas en otros idiomas.
- Sólo se tienen en cuenta en las bases de datos del ISI las publicaciones periódicas, a pesar de que una parte fundamental de la comunicación científica en

Ciencias Sociales y Humanas se produce por otros medios, fundamentalmente libros (entre el 50 y 60%) u otro tipo de informes (entre el 15 y 25%), frente a un modesto 10-35% que representa las revistas.

- Puesto que las disciplinas aplicadas se citan menos que las básicas, las investigaciones de este tipo (como son la mayoría de las educativas) tienden a ser menos citadas y, por tanto, a estar menos representadas.
- El número de veces que se cita una publicación no es ciertamente una buena medida de su calidad.
- Existe un sesgo notable a favor de revistas generalistas. Las revistas especializadas, sobre todo si lo son en temas muy novedosos, tienen poca probabilidad de aparecer, independientemente de la calidad de sus publicaciones.
- Temas de investigación que incumban a un grupo relativamente reducido de investigadores (ya sea la eficacia de adaptaciones curriculares para personas con síndromes raros, metodología de investigación o técnica aeroespacial, pongamos por caso) tendrán un número de citas muy bajo, con independencia de su calidad o del avance que supongan en su disciplina.

Incluso el legislador es consciente de las deficiencias que plantea el considerar las bases de datos del ISI como criterio fundamental de calidad. Así, durante muchos años ha mantenido (resolución de 6 de noviembre de 1996, BOE 280/96) que: «No obstante, en diversos ámbitos científicos correspondientes a este Campo [el campo 7, relativo a las Ciencias Sociales], la base de datos mencionada es incompleta en el entorno europeo y la Comisión Nacional podrá ampliar estos listados reconociendo la calidad de otras revistas». Sin embargo, resulta que no había en el ISI, hasta hace apenas un par de años, ninguna revista española sobre educación o pedagogía, ni francesa, ni italiana, ni portuguesa, ni latinoamericana, es decir, de nuestro entorno cultural. En estas circunstancias decir que la JCR (la base de referencia entre las del ISI) es «incompleta», como mantiene la normativa, nos parece sencillamente un eufemismo destinado a no reconocer la realidad. Como plantean algunos autores, un instrumento que ignora una realidad, una sociedad, una cultura, no puede tomarse como criterio de referencia de la calidad de la investigación que se realiza *en y sobre* ella.

Desde nuestro punto de vista, este proceso de falta de contextualización es un grave inconveniente. Aplicar los mismos principios para evaluar todas las disciplinas académicas provoca grandes perjuicios y tiene poca base científica. Disentimos en este sentido profundamente de otros autores (como De Miguel, 1997) que han tenido un gran peso en la configuración del proceso. Desde su punto de vista, todas las áreas académicas deberían evaluarse según unos mismos criterios y procedimientos. Nosotros, por el contrario, mantenemos que no se puede pretender tratar igual lo que es diferente y decir que eso es igualdad de oportunidades. A nadie se le ocurre considerar «fuertes» a hombres y mujeres, a niños y a adultos, utilizando un único y mismo criterio. En el caso

de la investigación científica es evidente que existe una clara dicotomía entre los distintos tipos de disciplinas. Así, en el caso de las Ciencias Naturales su objeto de estudio es común en todo el mundo y los hallazgos son válidos para todos los contextos. Se trata por tanto de una «ciencia internacional», que puede utilizar canales de comunicación comunes en todo el mundo. Sin embargo, no sucede lo mismo en el caso de las Ciencias Sociales. En este caso, el objeto de estudio depende claramente de aspectos culturales y el interés de los hallazgos se restringe, habitualmente, al contexto en el que se realiza. Ciertamente hay matices según la disciplina concreta, pero en el frecuente caso de que el objeto de interés y los objetivos estén determinados con claridad por la sociedad concreta en la que se realizan (pongamos por caso la eficacia de la reforma educativa en España), el interés de dicha investigación es evidentemente «local». Ello no significa demérito alguno. El hecho de que sea «local» no implica que no pueda tener un gran impacto social. Antes al contrario: puede tratarse de investigaciones orientadas precisamente a analizar o reconducir fenómenos socioculturales de un país concreto (como en el caso de la evaluación de la reforma educativa). Pero lo que es indudable es que el interés y la aplicabilidad de tales hallazgos no forma parte de esa «ciencia internacional», sino de una «ciencia social», o si se prefiere «local». Por supuesto, las vías de comunicación y difusión de esa investigación tendrán, necesariamente, condicionantes y alcances muy distintos a los de las Ciencias Naturales. No tiene por tanto sentido alguno utilizar los mismos criterios para evaluar un tipo de investigación y otro. No tiene sentido ni por el objeto, ni por la tradición, ni por entorno cultural, ni por las prioridades sociales, ni por la propia lógica del procedimiento.

El segundo gran grupo de críticas que ha recibido el proceso evaluador de la actividad investigadora llevado a cabo por la CNEAI es el relativo a la falta de transparencia sobre los criterios utilizados. Se trata de un defecto grave, ya que no hemos de olvidar que el *Joint Committee on Standards for Educational Evaluation* definió en 1981 la evaluación como el «enjuiciamiento sistemático de la valía o mérito de algo en base a cierto *criterio explícito* de valoración [el énfasis es nuestro]». Eso significa que si los criterios no son explícitos podemos concluir, lisa y llanamente, que se está haciendo mal el proceso de evaluación.

Esta falta de criterios explícitos provoca, por un lado, un efecto de indefensión y desamparo para los profesores-investigadores que se someten al proceso. Al no saber en concreto cuáles son los criterios con los que se va a evaluar su producción investigadora es frecuente que se presenten trabajos que luego son descartados directamente. Así, por ejemplo, hay casos en los que el interesado ha incluido trabajos de investigación que han recibido premios nacionales por su calidad por parte del Ministerio de Educación (que incluso se ha podido encargar de publicarlos) y el comité evaluador respectivo lo ha rechazado argumentando una difusión escasa, a pesar de que las diversas normativas han contemplado la posibilidad de tener en cuenta, incluso, aportaciones que no tienen en cuenta el criterio de «publicidad». También han argumentado algunos miem-

bros de estas comisiones (De Miguel, 1987) que las comunicaciones o ponencias en congresos «sólo pueden ser estimadas cuando son excepcionales». Sin embargo, qué significa en la práctica ese concepto de excepcionalidad nunca se ha aclarado, lo que en la práctica deja la valoración global al albur de los cambiantes criterios de los miembros de las comisiones.

Es un axioma conocido que, en todo tema de interés generalizado, la falta de información provoca rumores. En este caso no ha habido ninguna excepción y la falta de transparencia sobre los criterios utilizados ha provocado una persistente rumorología. Forman parte de la misma diversos listados que supuestamente se habían filtrado desde la CNEAI y que calificaban la importancia de diversas revistas, corriendo de mano en mano, aunque siempre sin autor reconocido. Esos listados, para colmo de males, eran cambiantes de una temporada a otra, lo que dejaba a los investigadores que querían planificar con tiempo la publicación de sus investigaciones en revistas «reconocidas» en una situación de perplejidad muy poco recomendable.

Un corolario que se deriva de esta falta de transparencia, y que viene a agravar sus efectos, es la variación de criterios. Al no existir criterios explícitos y transparentes, conocidos por todos y sometidos a crítica para que puedan mejorarse (que éste es el funcionamiento habitual de la ciencia que pretende evaluarse) las sucesivas comisiones no se limitan a aplicarlos, sino que los interpretan a su particular manera, que desde luego es cambiante hasta límites exasperantes. Podemos citar como ejemplo palmario el de un joven profesor-investigador que nada más ganar la oposición envió sus aportaciones para ser evaluado por la CNEAI y cuya solicitud no obtuvo el aprobado. Reponiéndose de la intensa frustración, el año siguiente ese profesor volvió a solicitar la evaluación para un nuevo período sexenal, que en este caso incluía algunos años del período que no había conseguido superar. Recuperó esos años porque en el informe que pretendía justificar su evaluación negativa alguno de los artículos de esa época sí había sido bien valorado, en concreto con siete puntos (recordemos que se requiere alcanzar un seis de media), considerando que se había publicado en una revista de gran calidad y prestigio. Bueno, pues en la evaluación de la segunda solicitud también fue suspendido y cuál no sería su sorpresa al indicársele en la justificación de la CNEAI que ese mismo artículo que había obtenido anteriormente siete puntos había pasado a ser valorado con sólo cuatro: ¡resulta que el mismo artículo, publicado en la misma revista de «calidad y prestigio», había perdido casi la mitad de su valor de un año a otro, y había de ser considerado una aportación notable a considerarse insuficiente! En la práctica ese cambio en la valoración le impedía al joven profesor-investigador alcanzar la media requerida para aprobar. Puede comprenderse fácilmente los efectos de perplejidad, desmoralización ante la arbitrariedad e incluso ira que provocan actuaciones de ese tipo, debidas fundamentalmente a una inadecuada forma de planificar un tipo de evaluación como la que estamos tratando.

Sin duda, el hacer explícitos los criterios detallados permitiría aclarar la situación, orientar a los investigadores en la línea deseada por los responsables de la investigación en nuestro país y motivar al personal investigador. Un paso notable en este sentido se ha producido a partir del año 2005, en las resoluciones de la presidencia de la CNEAI que regulan las convocatorias para presentar la solicitud de evaluación de los sexenios y establecen los criterios específicos en cada uno de los campos de evaluación. En la citada normativa, además de considerar de calidad contrastada las revistas que aparezcan en determinadas bases de datos (y vuelven a citarse las del ISI), se considera que «para que un medio de difusión de la investigación pueda ser reconocido a priori como de suficiente garantía» ha de cumplirse «una lista de criterios formales mínimos» (que aparecen como *Apéndice 1*). Se trata de «los requerimientos mínimos» que han de cumplir las revistas «para obtener una evaluación positiva», y que «tratan de reglar lo más posible la evaluación», aunque si bien es cierto que permitiendo a los diferentes comités de expertos «la modulación de estos mínimos, en cada caso particular». El caso es que revistas que cumplen con todos y cada uno de los requisitos del *Apéndice 1* son luego consideradas, durante el proceso de evaluación por parte de los comités de la CNEAI, revistas «de escasa repercusión científica», aunque simultáneamente sean premiadas por su calidad por otros organismos del mismo ministerio al que pertenece la CNEAI y estén entre las de mayor impacto según los estudios empíricos de In-RECS. Así las cosas, el problema ya no parece ser sólo de explicitación de criterios, faceta en la que aún quedaría mucho por avanzar, sino también, directamente, del cumplimiento de los mismos por parte de los integrantes de los comités. El resultado resulta, por ello, mucho más demoralizador para los implicados, que someten las decisiones sobre el reconocimiento de su valía profesional, de sus opciones de progreso en su carrera y de su mejora salarial a toda clase de subjetivismos y arbitrariedades.

ALGUNOS RESULTADOS

Los datos resultantes de las evaluaciones de la actividad investigadora están sometidos, por un lado, a la Ley de Protección de Datos y, por otro, al propio secretismo que ha impregnado buena parte del proceso, incluso desde antes de la aparición de la citada ley. Ello hace difícil realizar un análisis empírico detallado de los procesos. Sin embargo, entre el año 2002 y el año 2005 se realizaron y publicaron en la página web de la CNEAI diversos estudios estadísticos parciales sobre los resultados del proceso evaluador. Son especialmente interesantes los análisis de los datos acumulados entre 1989 y 2005, principal fuente para conocer algunas de las informaciones más relevantes sobre este proceso evaluador (tabla 1).

Como podemos comprobar, el único de los cuerpos docentes universitarios que consigue superar, eso sí, mínimamente, el 50% de integrantes con todos los sexenios posibles es el de catedráticos. Ha de tenerse en cuenta, sin embargo, que el 10% de ellos no

posee ningún sexenio (un 4% porque no lo ha solicitado nunca, mientras que el otro 6% corresponde a catedráticos que sí lo han solicitado, pero que lo han visto denegado en todos los casos), y todo ello a pesar de que es la máxima categoría administrativa, la cima de la carrera del profesor-investigador. El resultado global indica que en el año 2005, el 42% de los profesores funcionarios de universidades públicas no poseían ningún sexenio (el 31% ni siquiera se había presentado jamás al proceso evaluador). En estas condiciones parece obligado concluir que o bien el profesorado universitario español tiene, como globalidad, una calidad investigadora mínima (según la propia administración), o bien que los procesos de evaluación no son plenamente adecuados. Es cierto que algunos hábitos de publicación y de transferencia de resultados de la investigación son claramente mejorables. Pero es dudoso que, a la vista de los resultados, una herramienta de incentivación tan poderosa como los sexenios haya cumplido con las expectativas que despertaron.

TABLA 1
Datos acumulados entre 1989 y 2005

Cuerpo	% con todos los sexenios	% sin ningún sexenio
Catedráticos	52%	10%
Titulares de Universidad	47%	31%
Catedráticos Escuela Univ.	25%	57%
Titulares de Escuela Univ.	Sin datos	82%

Fuente: CNEAI

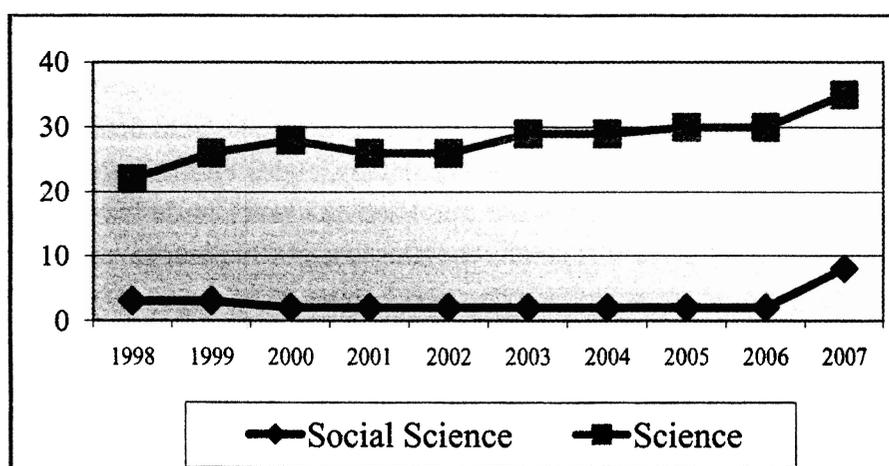
Si analizamos el resultado de las evaluaciones por grandes áreas podemos empezar a atisbar la respuesta a la dicotomía anteriormente planteada. Como se muestra en la tabla 2, más de la mitad de los profesores funcionarios de las áreas de Tecnología e Ingeniería y de las de Ciencias Sociales y Jurídicas no tiene ningún sexenio. Obsérvese que en el primer caso prácticamente la mitad de ellos (46%) ni siquiera lo ha solicitado. Probablemente, esto tiene que ver con el hecho de que se trata de áreas muy aplicadas que poseen un fuerte vínculo con proyectos aplicados en contextos económicos de alta demanda y bien recompensado, lo que hace que muchos de los profesores opten más por atender ese tipo de demandas sociales que por realizar investigación.

TABLA 2
*Porcentaje de profesores funcionarios sin tramos de investigación,
distribuidos por grandes áreas (1989-2005)*

Área	% No presentados nunca (A)	% Presentados pero sin obtenerlo (B)	A + B
Humanidades	22%	10%	32%
CC. Sociales y Jurídicas	39%	14%	53%
CC. Experimentales	17%	7%	24%
CC. de la Salud	27%	13%	40%
Tecnología e Ingeniería	46%	10%	56%

Puede observarse también en la tabla 2 que, con notable diferencia, el área en la que hay menos profesores sin sexenios es en la de las Ciencias Experimentales. A nuestro entender eso se debe a que los criterios de evaluación de ese tipo de ciencias «duras» son «internacionales», y por tanto todos los investigadores conocen el contexto y las normas de evaluación, que pasan por estándares internacionales (como, por ejemplo, el de revistas de impacto internacional incluidas en las bases de datos del ISI). La gran distancia con respecto a los resultados de los profesores de Ciencias Sociales nos indica que o bien estos últimos son investigadores mucho peores que sus compañeros de Ciencias Experimentales, o bien que la aplicación de criterios propios de las ciencias «internacionales» a las que tienen un desarrollo más vinculado a las sociedades «locales» es inadecuada. De hecho, podemos visualizar un indicio de la respuesta en el gráfico 1.

GRÁFICO 1
Revistas españolas incluidas en el JCR del ISI



Como podemos comprobar, tal como habíamos indicado con anterioridad, las bases de datos del ISI (después denominadas ISI-Thompson, más tarde Thomson y actualmente Thomson-Reuters) prácticamente ignoran la realidad cultural (social) española (de hecho una de las dos revistas que recoge en Ciencias Sociales es realmente de una especialidad médica). Por tanto, parece bastante evidente que los criterios de referencia sobre la calidad en las revistas son mucho menos adecuados para ser aplicados a los de Ciencias Sociales.

Hay un detalle que no quisiéramos dejar de resaltar. Se trata del notable incremento de revistas españolas que se han introducido en las bases de datos del ISI, particularmente llamativo en el caso de las de Ciencias Sociales. Existen dos factores que contribuyen a ello. Por un lado, la adquisición de esas bases de datos por grandes multinacionales que cotizan en bolsa y que han de incrementar sus niveles de negocio para satisfacer a sus accionistas, junto con la aparición de nuevos competidores en el área (Scopus, Scholar Google, etc.) parecen haber convencido a los gestores de las bases de datos para ampliar su presencia en diversos mercados. Por otro lado, la unificación de todos los contratos de acceso españoles a las bases de datos del ISI bajo una única firma, la FCYT (Fundación de Ciencia y Tecnología, dependiente del mismo ministerio que la CNEAI), ha incrementado notablemente el poder de negociación y la influencia como cliente de nuestro país. Todo ello se plasmará en los próximos años, sin duda, en un notable incremento de la presencia de revistas españolas en las citadas bases de datos, aunque por razones que tienen más que ver con la economía que con la ciencia.

SOBRE LOS EFECTOS DE LAS EVALUACIONES

Es muy difícil medir el impacto que la introducción de la evaluación de la actividad investigadora ha supuesto en la Universidad española en los últimos años. Se trata de un fenómeno muy complejo y de gran amplitud, para cuyo conocimiento nos falta perspectiva y mucha investigación al respecto. Hay sin embargo un hito que ha significado una cierta confirmación ante el público internacional. Se trata del escrito de Jiménez Contreras *et al.* (2002) en la revista *Nature*. Defienden los autores que «las consecuencias de esta ley han sido, en primer lugar, un cambio en hábitos de publicación de los científicos españoles». Mantienen también que la normativa ha supuesto «un aumento en el número de elementos de origen español en las bases de datos del ISI». Creemos que ambas afirmaciones son ciertas de manera global y son muestras de un cierto efecto de la reorientación de la difusión española que pretendía originalmente el legislador al introducir los sexenios. Jiménez Contreras y sus colegas (los mismos que han desarrollado In-RECS e In-REJS) mantienen igualmente que «la productividad científica española se ha duplicado: más del doble de artículos fueron publicados entre 1991 (inmediatamente después de que el sistema de incentivos por las publicaciones fuese aprobada por la ley) y 1998 en comparación con el período 1982-1990». En términos cuan-

titativos la afirmación última respecto al ritmo del incremento es cierta. Sin embargo, atribuir dicho aumento a los efectos de la introducción de los sexenios es sin duda excesivo. La mera aplicación de conocidas leyes cuantitativas nos da una visión muy distinta de esa realidad. Así, la Ley de Price pronostica que el crecimiento de la ciencia publicada crece de manera exponencial, de tal modo que se duplica aproximadamente cada diez años (aunque hay variaciones notables en el ritmo según las disciplinas). Por tanto, también es de esperar que se haya producido un cambio de semejante magnitud en el incremento de publicaciones de otros países, sin que ello pueda achacarse, lógicamente, a la introducción del sistema de sexenios en España. Se necesitarían estudios comparados e históricos de mucho mayor calado para poder afirmar con fundamento el efecto real que haya podido tener sobre la Universidad española la creación del sistema de incentivos basados en la evaluación de la actividad investigadora.

Por otro lado, existe también otro aspecto, poco estudiado pero no por ello menos conocido, respecto a los efectos secundarios de los sexenios. Se trata de las consecuencias negativas que ha implicado el «suspense» atribuido a muchas de las actividades de investigación. No son pocos los profesores-investigadores que tras una denegación del sexenio solicitado se han replegado sobre sí mismos, se han replanteado la finalidad de su actividad y, viendo comprometido hasta su amor propio, han optado por reorientarse hacia actividades más lucrativas que la investigación. Algunos otros, frustrados ante un proceso que perciben como injusto y arbitrario, sencillamente han decidido desconectar de la investigación (y a veces de más cosas) y languidecer con calma hasta la jubilación. Muchos conocemos casos cercanos y a mí, personalmente, algún compañero de edad próxima me ha confesado que al recibir una denegación de un sexenio solicitado decidió dejar sin acabar los tres libros en los que estaba trabajando, convencido de que el esfuerzo no iba a ser recompensado. Se trata de los «daños colaterales» de los sexenios, de los investigadores que no investigan y de los textos que no se escriben, muy difíciles de cuantificar y bastante ocultos, pero no por ello menos ciertos.

EN CONCLUSIÓN

Es difícil encontrar una conclusión definitiva en un proceso tan cambiante y de aún tan corto recorrido como el de la evaluación de la actividad investigadora. Se trata a nuestro entender de un procedimiento muy interesante de individualización del trato hacia los profesores-investigadores, que asegura que a distinto trabajo y mérito no habrá igual sueldo. Ello es un incentivo, si bien económicamente modesto, poderoso en el mundo académico. Los efectos prácticos de la introducción de los sexenios sobre el sistema de investigación español están aún por estudiar con detalle, pero sin duda han ido creando un nuevo ambiente, una nueva cultura, que hace que muchos profesores planifiquen con mucho más cuidado sus publicaciones.

Los procedimientos para la realización de la evaluación no son ciertamente perfectos, por lo que hace referencia a la claridad y rigor del proceso. Tenemos aún pendiente una muy necesaria mejora si queremos conseguir que los efectos de la introducción de los sexenios se conviertan en un estímulo continuo para el trabajo bien hecho y no, en algunas ocasiones al menos, en una fuente de frustración y resentimiento. Se trata de un instrumento poderoso que debe utilizarse mejor cada día, en bien de los investigadores que realizan su trabajo, de las universidades y de la sociedad a la que sirven.

REFERENCIAS BIBLIOGRÁFICAS

- ALIAGA, F. y N. ORELLANA (1999): *La utilización de las bases de Datos del I.S.I. para la evaluación de la calidad de las publicaciones sobre investigación educativa en España: Argumentación para un debate*. Comunicación presentada en el IX Congreso Nacional de Modelos de Investigación Educativa, Málaga, octubre. Publicado en AIDIPE, *Nuevas realidades educativas, nuevas necesidades metodológicas*, pp. 29-34, Málaga, Centro de Ediciones de la Diputación Provincial de Málaga. Disponible en línea <<http://www.uv.es/~aliaga/curriculum/Revistas-ISI.htm>>.
- DE MIGUEL, M. (1997): «La evaluación de la actividad investigadora del profesorado universitario en el ámbito de las Ciencias de la Educación», *Revista de Investigación Educativa*, vol. 15 (1), pp. 171-186.
- JIMÉNEZ CONTRERAS, E.; E. DELGADO LÓPEZ-CÓZAR; R. RUIZ PÉREZ y Víctor M. FERNÁNDEZ (2002): «Impact-factor rewards affect spanish research», *Nature*, 417(6.892), p. 898.
- LÓPEZ PIÑERO, J. M. y M. L. TERRADA (1992a): «Los indicadores bibliométricos y la evaluación de la actividad médico científica» (II). «La comunicación científica en las distintas áreas de las ciencias médicas», *Medicina Clínica*, vol. 98, pp. 101-106.
- MUÑOZ CANTERO, J. M.; M. P. RÍOS DE DEUS y E. ABALDE (2002): «Evaluación Docente vs. Evaluación de la Calidad», *RELIEVE*, vol. 8, n.º 2. Disponible en línea <http://www.uv.es/RELIEVE/v8n2/RELIEVEv8n2_4.htm>. Consultado el 30 de octubre del 2008.
- STUFFLEBEAM, D. L.; W. J. FOLEY; W. J. GEPHART; E. G. GUBA; R. L. HAMMOND; H. O. MERRIMAN y M. M. PROVUS (1971): *Educational Evaluation and Decision-making*, Itasca, Illinois, F. E. Peacock Publishing.
- UNIVERSITAT DE VALÈNCIA (2001): *Normativa de l'Avaluació de la Docència*. Documentación interna.