

Ante la reforma del sistema español de evaluación académica: recordando el pasado para no errar en el presente

On the brink of the reform of the Spanish academic evaluation system: Recalling the past to avoid mistakes in the present

Emilio Delgado-López-Cózar

Delgado-López-Cózar, Emmilio (2022). "Ante la reforma del sistema español de evaluación académica: recordando el pasado para no errar en el presente". *Anuario ThinkEPI*, v. 16, e16a37.

<https://doi.org/10.3145/thinkepi.2022.e16a37>

Publicado en *IweTel* el 14 de noviembre de 2022

Emilio Delgado-López-Cózar

<https://orcid.org/0000-0002-8184-551X>

Universidad de Granada

Facultad de Comunicación y Documentación

Campus Cartuja. 18071 Granada, España

edelgado@ugr.es



Resumen: Reflexiones sobre el origen del sistema de evaluación científica en España, centrando el análisis en el sistema de selección y acreditación del profesorado universitario. Se recuerda cómo surgió para evitar la endogamia universitaria. Se desmonta la tesis de que estemos ante un sistema centralizado: la acreditación sí lo está, pero la selección del profesorado depende de las universidades. Se defiende que no podemos confiar plenamente en las universidades para que instauren políticas inmunes al localismo y el nepotismo si no se les hace plenamente responsables de sus actos. Se destaca el desequilibrio funcional que existe en el profesorado universitario español. Dentro de las actividades que actualmente se atribuyen al profesor universitario (docencia, investigación, transferencia, gestión) son las de investigación las más valoradas, especialmente las publicaciones que son claves para ingresar y progresar en la carrera académica. A ello ha contribuido el éxito de los sexenios de investigación no sólo como incentivo de productividad sino de reputación científica y jerarquización académica. Otros rasgos de nuestro sistema de evaluación son el predominio de la evaluación individual y el uso de indicadores bibliométricos como supremos valores de calidad. El sistema español de evaluación académica optó por dar preferencia a la evaluación individual. De nuevo el origen de esta inclinación hay que buscarla en la introducción de los complementos de productividad docente e investigador en 1989. Los académicos españoles crecieron en una cultura de evaluación individual donde debían obtener sus plazas por concursos-oposición y ser premiados con gratificaciones basadas en la consecución de méritos personales. Los criterios cuantitativos aparecieron como los antídotos necesarios para frenar el nepotismo. En un país sin infraestructura y cultura evaluativa aseguraban la objetividad y eficiencia en los procesos de selección y valoración de las publicaciones e indirectamente un mínimo rigor y relevancia de las propias investigaciones. Los pensamientos expuestos se acompañan de sugerencias que ayuden a orientar la necesaria reforma del sistema de evaluación científica.

Palabras clave: Evaluación científica; Evaluación del rendimiento de los investigadores; Política científica; Indicadores bibliométricos; Profesorado universitario; Enseñanza universitaria; Sexenios de investigación; *Aneca*; España.

Abstract: This essay provides some reflections on the origin of the current scientific evaluation system in Spain, focusing on an analysis of the system of selection and accreditation of university faculty. This system was put in place to prevent academic inbreeding. The thesis that this is mainly a centralized system is rebutted: although accreditation is indeed centralized, faculty selection is carried out by universities themselves. It is argued that we cannot fully trust universities to establish policies that are immune to localism and nepotism if they are not held fully responsible for their actions. The functional imbalance that exists in Spanish university teaching staff is highlighted. Among the activities currently carried out by university professors (teaching, research, knowledge transfer, and management), research is the most highly valued in evaluation programs. More specifically, publications are key to launch and advance careers in the academic environment. The success of the program of *sexenios* (six-year assessment periods), which serves not only as an incentive for productivity but also as an instrument to establish scientific reputation and academic hierarchy, has contributed to this situation. Other features of the current evaluation system in Spain are the predominance of individual-level evaluation and the use of bibliometric indicators as the final quality values. The Spanish academic evaluation system chose to give preference to individual-level evaluation. Once again, the origin of this approach can be found in the introduction of teacher and researcher productivity salary supplements in 1989. Spanish academics grew up in a culture of individual-level evaluation where they had to obtain their jobs through competitive examinations and were rewarded with bonuses on the basis of personal achievements. Quantitative criteria appeared as the necessary antidotes to curb nepotism. In a country lacking infrastructure and a culture of evaluation, quantitative indicators ensured objectivity and efficiency in the processes of selection and evaluation of publications, and, indirectly, a minimum amount of rigor and relevance of the research itself. The thoughts exposed here are accompanied by suggestions that might help guide the necessary reform of the scientific evaluation system.

Keywords: Research evaluation; Performance assessment of researchers; Research policy; Bibliometric indicators; University faculty; Higher education; Universities; *Aneca*; Spain.

Estos dos últimos años en España hemos asistido a intensos debates sobre el rumbo que debe seguir el sistema de evaluación académica (Delgado-López-Cózar; Ràfols; Abadal, 2021; 2022; Ràfols; Molas-Gallart, 2022), tanto en lo que respecta al modelo de acreditación del profesorado como a la forma de evaluar la dimensión investigadora con especial énfasis en el uso de indicadores bibliométricos (Aguillo, 2021; Codina, 2021; Rodríguez-Navarro, 2022). Esto se produce en el contexto de la introducción de cambios en el sistema de educación superior en España (Arenilla-Sáez, 2021), con la aprobación de una nueva ley orgánica (Ahedo-Gurrutxaga; Martínez-Palacios; Ormazabal-Gaston, 2022) y de la germinación de un movimiento impulsado por la Comisión Europea para cambiar los sistemas de evaluación de la investigación en Europa (European Commission, 2021; 2022).

A fin de enriquecer la discusión y no perder el norte creo que no está de más realizar un breve ejercicio de memoria histórica que refresque el porqué del sistema de evaluación científica en España que tenemos en la actualidad, centrando el análisis en el sistema de selección y acreditación del profesorado universitario.

1. ¿Por qué implantamos la acreditación nacional del profesorado universitario?

El actual sistema de acreditación nacional surge, a comienzos de este siglo, fundamentalmente como un intento de acabar con la endogamia practicada en la selección y promoción del profesorado universitario. La *Ley de Reforma Universitaria (LRU)* (1983), al conceder plena autonomía a la universidad en la contratación de personal, condujo automáticamente a que las universidades primaran a los candidatos formados y reclutados en la propia universidad. Durante casi dos décadas, los profesores nos presentábamos a “nuestra plaza”. Lo era porque la convocaba “nuestra universidad”, con un perfil docente e investigador ajustado al perfil del candidato de la casa, con un tribunal propuesto por “nuestro departamento”, a instancias y siguiendo las indicaciones sugeridas por el “propio candidato”. Los ejercicios consistentes en la defensa del historial académico, de un proyecto docente de la asignatura a impartir y de un tema elegido libremente por el o la concursante se ajustaban como anillo al dedo del aspirante casero. Nada se dejaba al azar y si bien es verdad que tres de los cinco miembros de

“El actual sistema de acreditación nacional surge, a comienzos de este siglo, fundamentalmente como un intento de acabar con la endogamia practicada en la selección y promoción del profesorado universitario”

los tribunales se elegían por sorteo por parte del *Consejo de Universidades*, éstos normalmente estaban a lo que señalaran el presidente y secretario, representantes y altavoces del departamento convocante.

En fin, todo quedaba en casa, *ad libitum* de los opositores como describió con sorna en un estilo cáustico el profesor **Sosa-Wagner** (2016); más mordaz y ácido era el chiste que circulaba por aquellos entonces en los campus universitarios: para ganar un concurso-oposición a una plaza de profesor universitario se requieren tres requisitos: tener posición “bipedestante”, poseer lenguaje articulado y contar con tres votos; los dos primeros no son indispensables.

El intento de acabar con este régimen de cosas vino de la mano de la *Ley Orgánica de Universidades* (LOU) en 2001 (España, 2001) con la creación de las habilitaciones, primero, y de las acreditaciones, después. Las habilitaciones, que con sus ejercicios se aproximaron a lo que son unas “auténticas oposiciones” pues fomentaron la competencia, fracasaron porque el procedimiento y sistema de gestión de los concursos era disparatado a la par que dispendioso; la intención era buena pero su implementación, no.

La introducción del sistema de acreditación nacional ha atemperado y mitigado esta anormalidad, pero no ha logrado desgraciadamente orillarla definitivamente.

“Ciertamente los profesores deben acreditarse en el nivel nacional o autonómico en determinadas figuras, pero una vez conseguido la universidad vuelve a contratar a los suyos, esto es, a sus acreditados. Los candidatos acreditados pasan después en su propia universidad plácidamente por ejercicios donde ellos siguen defendiendo su curriculum, su programa docente, su proyecto investigador eligiendo la lección o el tema de investigación sobre el que quieren disertar. Los tribunales convenientemente elegidos, (todos lo son por la universidad convocante a propuesta del departamento respectivo y a instancia del defensor “in pectore” de la plaza, esto es el, concursante doméstico), no suelen ser muy inquisitivos; si *Aneca* lo ha acreditado, ¿quiénes somos nosotros para suspender al candidato!, se puede oír frecuentemente de boca de los miembros de los tribunales. Por tanto, el nuevo profesor entra y va pasando por las escalas sin tener que demostrar sus saberes y su capacidad para transmitirlos. Este sistema podría ser eficiente si la propia universidad y el departamento sufrieran las consecuencias de una deficiente selección y contratación del profesorado. Si su futuro dependiera de ello tal vez las universidades se tentarían la ropa y competirían por seleccionar a los mejores, sean del terruño o de fuera” (Delgado-López-Cózar, 2016).

2. ¿De verdad estamos ante un sistema centralizado o ante el ejercicio fallido de la autonomía universitaria consagrada en la *Constitución Española*?

Cuando hoy se señala como uno de los causantes de los males del sistema de evaluación académica española su centralización, encarnada en la *Aneca*, y su excepcionalidad en Europa, convendría preguntarse si de verdad estamos ante un sistema centralizado. Pues bien, desde el punto de vista de la selección y contratación del profesorado el sistema español no es centralizado; posee un carácter híbrido, aunque con un decisivo poder autónomo de las universidades. Si bien la acreditación del profesorado es ejercida bajo directrices nacionales (*Aneca*) los criterios efectivos para la selección y contratación del profesorado dependen de las universidades. Ciertamente los criterios de evaluación fijados por *Aneca* son y han sido decisivos para la orientación del perfil del profesorado que tenemos. Pero no es menos verdad que las universidades disponen de autonomía para determinar cuál es el perfil de profesor que más le interesa. Conviene recordar que los baremos que rigen los concursos-oposición para la selección y contratación de ayudantes y contratados doctores, de profesores titulares y catedráticos, así como del resto de personal (colaboradores, asociados, sustitutos interinos, etc..) son elaborados y aprobados por las propias universidades sin cortapisas del poder central o autonómico. Cuentan, por tanto, con un absoluto margen para modelar y elegir dentro del profesorado acreditado aquél que mejor se adapte a sus intereses estratégicos. En definitiva, las universidades españolas cuentan *de iure* y *de facto* con la potestad de seleccionar al profesorado con total autonomía.

Lo sorprendente es que cuando uno compara los baremos empleados en las universidades se detecta un extraordinario parecido entre ellos. La diversidad teóricamente esperable de la autonomía universitaria en la selección del profesorado se traduce en una práctica homogeneización (el tipo de profesor que se privilegia posee un perfil idéntico), pues la naturaleza y el peso de los méri-

“Desde el punto de vista de la selección y contratación del profesorado el sistema español no es centralizado. Los baremos que rigen los concursos-oposición para la selección y contratación de ayudantes y contratados doctores, profesores titulares y catedráticos son elaborados y aprobados por las propias universidades sin cortapisas del poder central o autonómico”

tos que se contemplan son extraordinariamente similares. Muchos baremos parecen el resultado de un inmisericorde copia y pega. Ahora bien, lo más inverosímil de este sistema es que conduce siempre a un resultado idéntico: normalmente es el profesor de la casa, esto es, el formado en la propia universidad y el que ha trabajado en ella es el que suele obtener la plaza. Así, resulta que el 73,3% del profesorado de las universidades públicas españolas trabaja en el mismo centro en el que se doctoró, un 87,4% lo hace en la misma comunidad autónoma, y casi un 98% trabaja en el mismo país, España¹.

La pregunta es ¿por qué ocurre esto en un sistema supuestamente autónomo?

Me temo que la respuesta haya que buscarla en explicaciones más prosaicas. Parece que la acreditación se percibe como un obstáculo a superar, como el marchamo que certifica la idoneidad de un profesor, la condición necesaria para que la universidad de origen (*alma mater*) del profesor confirme su definitivo destino. “Si *Aneca* ha dicho que es buen profesor, yo –la universidad, se entiende– no tengo que decir nada en contra”, se oye decir con frecuencia entre los miembros de los tribunales que pueblan los campus universitarios. La endogamia asegurada, pervirtiendo el inicial objetivo del sistema.

Por todo ello, antes de abolir el sistema de acreditación nacional actual debemos pensar seriamente sobre sus posibles efectos en el contexto español, pues podría ser un craso error. La experiencia pasada nos enseña que cuando se deja en manos de la propia universidad la valoración del rendimiento del profesorado este se corrompe o termina en hacerse inservible. A la vista está en las cifras que acabamos de reseñar y recuérdese lo ocurrido con los quinquenios docentes y los sexenios de investigación. Los dos se implantaron al mismo tiempo, pero con muy desiguales modos de aplicación y efectos. Mientras los quinquenios se evalúan por las universidades (café para todos, con minusvaloración y efecto negativo para la buena consideración de la docencia) los sexenios se entregan a una comisión nacional con evaluadores externos (mecanismo selectivo que inmediatamente se convierte en elemento reputacional: llave que abre el progreso en la carrera académica).

¿Podemos confiar plenamente en las universidades para que instauren políticas inmunes al localismo, el nepotismo y la corrupción? Desgraciadamente, me temo que no. Si el destino de una universidad dependiera realmente de sus actos, y entre ellos, el crucial sistema de selección y promoción del profesorado, el resultado sería otro. Con el actual sistema universitario nadie se “juega” el futuro de su departamento, centro, universidad. No se genera un incentivo a seleccionar al profesorado más excelso para el éxito de la propia universidad sino al más dócil, fiel, partidario e incondicional adepto a los miembros de la propia casa. Solo si se instauraran mecanismos que aseguren que si una universidad no lo hace bien podría extinguirse; solo si se fuera consciente de que el futuro está en las propias manos se podría avanzar por esa línea de confianza en las universidades: exigiendo plena responsabilidad sobre sus actos y, por supuesto, sobre sus consecuencias. Si los actos de las universidades no tienen efectos seguiremos mejor con el sistema actual.

Creo que, sin plena autonomía financiera, no existe plena responsabilidad en sus decisiones. Y es que las universidades españolas, aunque formalmente gozan de una gobernanza autónoma realmente son dependientes de manera efectiva y muy uniforme de los presupuestos públicos. La clave está pues, como siempre, en la economía, en la forma de financiar las universidades. Ese es el factor que determina la auténtica autonomía. Se hace a escote, sin mecanismos de rendición de cuentas y competencia por los recursos. Si parte de la financiación de la universidad dependiera de su actuación, de la toma de decisiones autónoma, creo que empezariamos a cambiar realmente la universidad. Otra cuestión es si eso es posible en España. Responder a esta pregunta nos llevaría a elucubraciones escabrosas; en nuestro país somos especialistas en el gatopardismo, en el lampedusiano adagio de “cambiar todo para que nada cambie”. Permítanme, por tanto, que sea escéptico.

“Normalmente es el profesor de la casa, el formado en la propia universidad y el que ha trabajado en ella es el que suele obtener la plaza. El 73,3% del profesorado de las universidades públicas españolas trabaja en el mismo centro en el que se doctoró, un 87,4% lo hace en la misma comunidad autónoma, y casi un 98% trabaja en el mismo país, España”

“Antes de abolir el sistema de acreditación nacional actual debemos pensar seriamente sobre sus posibles efectos en el contexto español, pues podría ser un craso error. La experiencia pasada nos enseña que cuando se deja en manos de la propia universidad la valoración del rendimiento del profesorado este se corrompe o termina en hacerse inservible”

3. El modelo de profesorado universitario español: un desequilibrio funcional

Desde el nacimiento de las universidades en el siglo XI la función primigenia del profesor es la docente (formar y educar en el ejercicio de profesiones de nivel superior) a la que se añadió, ya en el siglo XX, la investigación (la capacidad de crear y generar nuevo conocimiento a través de la indagación intelectual). Más recientemente, se ha agregado el cometido de transmisión/transferencia del conocimiento a la sociedad en todas las vertientes posibles: económicas, sociales, culturales. Siempre, y de forma inherente a la condición universitaria, el profesorado también ha tenido que asumir funciones de gestión de las instituciones docentes y científicas donde desarrolla su labor.

Si embargo, en España, nunca ha existido un equilibrio entre las obligaciones que actualmente se atribuyen al profesor universitario. Hasta la llegada de la *LRU* (1983) al profesor español se le juzgaba y premiaba fundamentalmente por sus actividades docentes. El impulso dado a la investigación por las distintas políticas emprendidas para mejorar nuestro sistema de I+D+I (*Ley de la ciencia* (1985)), junto a la aceptación y prestigio conseguido entre la comunidad científica española por los sexenios de investigación no solo como sistema de recompensa sino de reputación científica, hizo que fuera la dimensión investigadora el aspecto más valorado.

Lo que inicialmente se pretendía con los sexenios otorgados por la *Cneai* (1989), que era introducir un complemento de productividad que mejorara el sueldo del profesorado y que incentivara y premiara la actividad investigadora, se transmutó, especialmente a partir de la *LOU* en 2001, en la llave que abrió todas las puertas del desarrollo, promoción y ascenso en la carrera profesional universitaria (acceso a plazas, obtención de proyectos de investigación, participación en enseñanza de máster oficial, doctorado, dirección tesis, presencia en tribunales, etc.), y, desde el famoso *Decreto Wert* de reorganización de las cargas docentes (los famosos 32 créditos), un elemento de distribución de las cargas de trabajo en la universidad y de jerarquización social de los cuerpos docentes. Esta metamorfosis fue un claro síntoma del éxito del modelo de evaluación de la *Cneai*, que a pesar de los titubeos iniciales y los dimes y diretes ha sido aceptado e interiorizado por la comunidad científica española. ¡Qué diferencia con los complementos creados en el mismo decreto de 1989 para evaluar la docencia -los llamados quinquenios-, cuya devaluación tantos efectos nocivos ha causado en la buena estima de la docencia! Mientras que los sexenios de investigación fueron gestionados por una comisión nacional (la *Cneai*), pilotada por expertos en los distintos campos de conocimiento, los quinquenios docentes se administraron por las propias universidades. Los efectos han sido mayúsculos para todo el sistema universitario, ya que el disponer de sexenios, un indicador exclusivamente de suficiencia investigadora, se convirtió en la llave que abrió todas las puertas. El mensaje que se retiene es que, si se investiga y además se obtiene buen nivel de publicaciones, se progresa. No es de extrañar que este dicho se haya hecho tan popular: “se nos paga por lo que enseñamos y se nos premia por lo que investigamos”... quiero decir, publicamos.

Hasta 2017 en *Aneca* es la investigación (figura 1) y dentro de ella la publicación (tabla 1) el elemento determinante para la acreditación en todas las categorías profesionales que se aseguran el ingreso y progreso en la carrera académica.

“En España, nunca ha existido un equilibrio entre las obligaciones que actualmente se atribuyen al profesor universitario. En *Aneca* es la investigación y dentro de ella la publicación el elemento determinante para la acreditación en todas las categorías profesionales que se aseguran el ingreso y progreso en la carrera académica”

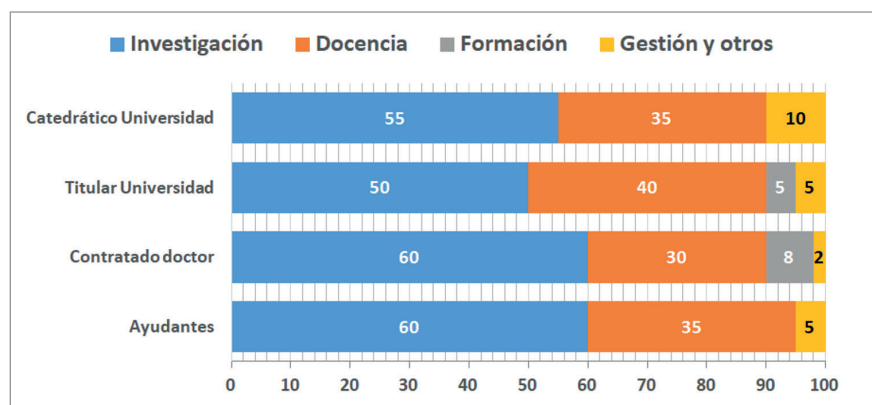


Figura 1. Puntuaciones máximas otorgadas por *Aneca* a las distintas dimensiones de la actividad académica desde 2007 hasta 2017

Tabla 1. Puntuaciones máximas y porcentajes otorgados por Aneca a las publicaciones respecto del total de puntos necesarios para acreditarse (2007-2017)

	Máximo en publicaciones puntos	Puntuación necesaria para obtener la acreditación	
	Puntos	Puntos	Porcentaje
Ayudantes doctor	42	55	76,4
Contratado doctor	42	55	76,4
Titular universidad	38	65	58,5
Catedrático universidad	35	80	43,8

La reforma introducida en 2017, revisada parcialmente en 2019, empezó a flexibilizar las posibilidades, ofreciendo otras vías (tabla 2).

Tabla 2.

Catedrático					Titular					
	Investigación	Docencia	Transferencia/ Actividad profesional	Gestión		Investigación	Docencia	Transferencia/ Actividad profesional	Gestión	Formación
Calificación mínima	B	B			Calificación mínima	B	B			
Calificación mínima	A	C, E ¹			Calificación mínima	A	C, E ¹			B
Calificación mínima	B	C	B		Calificación mínima	B	C	B		B
Calificación mínima	B	C		B	Calificación mínima	B	C		B	B
Calificación mínima	C	B	A		Calificación mínima	C	B	A		B
Calificación mínima	C	B		A						

¹ El nivel E solo es válido para la acreditación cuando un solicitante haya desarrollado su carrera principalmente en una institución no universitaria o en una universidad no española donde el cómputo y los instrumentos de medición de la calidad de la actividad docente resulten difíciles de trasladar al sistema español.

No obstante, el peso atribuido al componente investigador, y en concreto a la publicación, es abrumador ya que los méritos obligatorios a cumplir indefectiblemente para obtener las mejores calificaciones (A, B) se componen solo de 4 aportaciones significativas (normalmente publicaciones) y un número variable de publicaciones, fundamentalmente artículos publicados en revistas con posiciones relevantes en los rankings de revistas por índice de impacto (*JCR*, *SJR* y otros). Entre los méritos específicos y complementarios que se pueden usar para compensar calificaciones insuficientes figuran otras ligadas a la publicación científica.

Es entendible y más que justificado que en los años ochenta del siglo pasado en aras de acabar con el aislamiento y sempiterno retraso científico español fuera prioritario reorientar el perfil exclusivamente docente de las universidades incentivando la investigación; la publicación de sus resultados en espacios internacionales; la introducción de la cultura de la evaluación por pares para la asignación de los recursos del sistema y para el control y certificación de los contenidos científicos (revistas, editoriales); la evaluación de resultados tangibles.

Conseguida hace años ya la homologación del quehacer científico de los investigadores españoles e interiorizada ya la misión investigadora como propia del ADN universitario debemos reequilibrar y alinear las funciones del profesorado con lo que son las misiones actuales de la universidad. El profesorado debe atenerse fundamentalmente a la docencia, la investigación y la transferencia asegu-

“Es urgente que la publicación y los indicadores bibliométricos asociados a ella dejen de ser el eje sobre el que pivota la evaluación. Deben valorarse adecuadamente otras actividades que permitan a los profesores poder desarrollar su carrera académica sin la presión permanente del “publica o perece”. Debe incrementarse la flexibilidad para enriquecer la variedad de perfiles de profesorado. Posibilitar diversos caminos y trayectorias basados en cualquiera de los tres ejes centrales de docencia, investigación y transferencia”

rando que los académicos puedan elegir distintas rutas en su carrera académica. Hay que permitir con flexibilidad que los profesores puedan centrarse preferentemente en aquellas dimensiones que más le satisfagan o se adapten a sus circunstancias, eso sí, siempre asegurando unos mínimos a cumplir en todas las dimensiones.

Creo que es urgente que la publicación y los indicadores bibliométricos asociados a ella dejen de ser el eje sobre el que pivota la evaluación. Deben valorarse adecuadamente otras actividades que permitan a los profesores poder desarrollar su carrera académica sin la presión permanente del “publica o perece”. Debe incrementarse la flexibilidad para enriquecer la variedad de perfiles de profesorado. Posibilitar diversos caminos y trayectorias basados en cualquiera de los tres ejes centrales de docencia, investigación y transferencia de manera que el profesorado pueda, ajustar su carrera académica según sus preferencias y que las universidades puedan elegir figuras laborales diferentes que adapten a sus necesidades.

4. Evaluación individual frente a evaluación colectiva: primero yo, después yo y, por último, yo.

El sistema español de evaluación académica optó por dar preferencia a la evaluación individual de los científicos frente a la institucional. De nuevo el origen de esta inclinación hay que buscarla en la introducción de los complementos de productividad docente e investigador en 1989. Concebidos en ese momento como medios de mejorar las retribuciones del profesorado podía ser lógico aplicarlos a nivel individual ya que lo que se perseguía es incentivar el rendimiento personal. No obstante, en ese momento podría haberse pensado en hacer evaluaciones institucionales a nivel de departamentos o centros, pero ni la arquitectura institucional universitaria estaba en ese momento plenamente vertebrada ni existía tradición de evaluaciones grupales. El hecho es que los académicos españoles crecieron en una cultura de evaluación individual donde debían obtener sus plazas por concursos-oposición y ser premiados con gratificaciones basadas en la consecución de méritos personales.

Plenamente convencido de que en el ámbito científico el todo es superior a la parte, que el grupo es preferible al individuo y que compartir es siempre mejor que partir, urge acabar o atenuar el individualismo montaraz que se ha instalado en la academia española que genera efectos perversos en las dinámicas académicas. Hacer que los grupos, departamentos u otras unidades estrechen lazos y compartan objetivos comunes puede mejorar estructuralmente el sistema, hacerlo más eficiente y, sin duda, más habitable. Genera sinergias positivas no solo para el grupo en sí mismo sino para el entorno. Por consiguiente, es hora de introducir evaluaciones institucionales que aúnen los esfuerzos personales en aras de conseguir objetivos más ambiciosos.

No obstante, creo que la evaluación institucional no puede ni debe sustituir a la valoración individual de las competencias de los académicos. Cuestión distinta es si podría ser interesante generar un sistema de evaluación institucional que certifique a nivel de departamentos o grupos el cumplimiento de buenas prácticas en la evaluación científica inspiradas en las declaraciones internacionales sobre las que se ha alcanzado cierto consenso. Así, en el ámbito investigador operacionalizar en una serie de indicadores los principios formulados en *DORA*, *Leiden* o *Hong Kong*, podría repercutir en la implantación de procesos que mejoren las prácticas científicas en las instituciones.

No obstante, soy bastante escéptico sobre como en España se ha abordado la evaluación institucional académica. Véanse si no los sistemas de garantía interna de calidad certificados por la *Aneca* (programa *Audit*). ¿De verdad es un programa que ha tenido éxito por su seguimiento? ¿De verdad que los centros que disfrutan de estos reconocimientos son realmente modélicos en sus procesos de enseñanza? o ¿son mejores que aquellos que no disfrutan de ellos?

Como ya escribí hace unos años

“se corre el riesgo otra vez de generar normativas exhaustivas y quisquillosas que suelen terminar produciendo en los que se someten a evaluación documentos llenos de declaraciones grandilocuentes y vacuas; verborrea que es cortada y pegada de los documentos oficiales sin más, o copiada de otros similares, sin otro objeto que cumplir con los requisitos formales exigidos por la agencia. En fin, textos que nadie se lee más que aquellos que pertenecen a los servicios administrativos de la universidad que deben filtrar los documentos para que se adecuen a los ‘estándares de la agencia’ y, obviamente, los propios evaluadores de la agencia, que deben acreditarlos” (Delgado-López-Cózar, 2016).

“Urge acabar o atenuar el individualismo montaraz que se ha instalado en la academia española que genera efectos perversos en las dinámicas académicas”

En fin, me excuso por mi descreimiento acerca de los efectos prácticos de estos sistemas, pero creo que debo ser franco y no decir lo que se quiere oír si no manifestar lo que se piensa, aunque incomode. A día de hoy, si se mantiene este modelo universitario, sin ligar la financiación a la evaluación de resultados y cumplimiento de objetivos que hagan sentir a las universidades las consecuencias de sus decisiones, lo más conveniente es mantener la acreditación individual de profesorado.

Cuestión distinta es el cambio en los requisitos y criterios, algo a todas luces necesario; o cómo se pueden optimizar los recursos simplificando los procesos de evaluación. En este sentido carece de lógica mantener las mismas acreditaciones a nivel nacional y regional. ¿Qué sentido tiene someterse a la acreditación a la figura de ayudante o contratado doctor simultáneamente por *Aneca* o una agencia autonómica? O uno u otro, pero no los dos. Me parece un gasto innecesario, improductivo y un exceso burocrático que consume recursos de gestión y el precioso tiempo de los académicos. Me inclino por la acreditación nacional por varias razones:

- más masa crítica;
- más lejanía respecto a los sujetos y objetos de evaluación;
- propicia una mayor movilidad en todo el sistema, impidiendo tentaciones de imponer obstáculos territoriales.

No obstante, dado el sistema de organización que nos hemos dotado, tampoco pongo objeciones a que sea a nivel regional siempre que la evaluación se externalice; ya se sabe que los auditores cuanto más lejos de los auditados mejor, para evitar contaminaciones.

5. Contar, contar, contar: criterios cuantitativos en la evaluación académica española. ¿Y la evaluación cualitativa?

Los evaluados y los evaluadores en España, y creo que en todos los países, aman los baremos detallados por distintas razones:

- todo el mundo sabe a qué atenerse, especialmente los evaluados que exigen seguridad y predictibilidad en sus evaluaciones: criterios precisos, concretos y autoevaluables;
- facilita la tarea evaluadora y objetiva la evaluación, minimizando la arbitrariedad;
- alivia el estrés de los evaluadores, pues permiten soslayar su responsabilidad enjuiciadora detrás de parámetros objetivos y cuantificables.

Si miramos con perspectiva histórica, desde 1989 en que se implantaran los sexenios, los distintos sistemas de evaluación académica españoles han ido concretando y precisando al máximo los méritos evaluables, llegando incluso a una cuantificación exacta de los requisitos exigibles para obtener diversos oficios, plazas, becas y subvenciones (número de artículos de revistas, de comunicaciones a congresos, de proyectos obtenidos, de tesis dirigidas, de horas de clase impartidas y así un largo etcétera).

Justamente en este contexto debe entenderse la adopción inmediata de indicadores bibliométricos (número de publicaciones en revistas indexadas en los *JCR* o distintos sucedáneos) con un éxito fulgurante. Han ido penetrando en todos los rincones académicos colonizando todas las áreas de conocimiento; surgidos en las Ciencias Naturales y Experimentales se fueron extendiendo progresivamente a las Ciencias Sociales y Aplicadas, para llegar también a las Humanidades y el Derecho, el último reducto sin rankings bibliométricos (Delgado-López-Cózar, 2019). La contaminación bibliométrica ha sido continua. Las convocatorias anuales de sexenios así como los grandes hitos de la evaluación en *Aneca* (2007, 2017 y 2019) iban especificando cada vez más las exigencias: si primero bastaba con un número de publicaciones (revistas, libros), posteriormente se pedía que estuvieran indexadas en determinados rankings y bases de datos o, por último, que figuraran en determinadas posiciones (cuartiles) señalándose las condiciones en que serían valoradas (obras en colaboración y orden en la autoría) y las puntuaciones exactas a las que equivaldrían.

Esta dinámica debe entenderse a la luz de la situación de la que veníamos. Convertida la universidad, en la época más antigua, en un sistema feudal donde la chistera y el dedo del catedrático de turno eran el supremo criterio de selección o,

“Para los responsables de la política científica española en el momento en que se cimentó nuestro sistema de evaluación, lo más efectivo y rentable fue externalizar de alguna manera la evaluación a través de indicadores, bases de datos e índices que tuvieran reconocimiento internacional. Es por eso por lo que se adoptó el *Factor de Impacto* y los *JCR* como patrón de evaluación (ampliado después a otros índices más omnicomprendivos de todas las disciplinas como *SJR*, *IN-Recs*, *Fecyt* ...)”

en época más moderna, en un reino de taifas donde todo estaba perfectamente repartido, los criterios objetivos aparecieron como los antídotos necesarios para frenar el nepotismo. Para los responsables de la política científica española en el momento en que se cimentó nuestro sistema de evaluación, lo más efectivo y rentable fue externalizar de alguna manera la evaluación a través de indicadores, bases de datos e índices que tuvieran reconocimiento internacional. Es por eso por lo que se adoptó el *Factor de Impacto* y los *JCR* como patrón de evaluación (ampliado después a otros índices más omnicomprendivos de todas las disciplinas como *SJR*, *IN-Recs*, *Fecyt*...). Haciendo esto, las autoridades españolas tenían la convicción de que aseguraban una elevada calidad de los procesos de selección y valoración de las publicaciones e indirectamente un mínimo rigor y relevancia de las propias investigaciones.

De ahí pasamos, siguiendo la ley del péndulo, a un abuso y mal uso de los indicadores bibliométricos, empleados inmisericordemente para todo, por todos y en todos los momentos, sin siquiera atender al tipo de evaluación, contexto y circunstancias (Delgado-López-Cózar; Martín-Martín, 2019), con una simplificación que repugna la razón: un indicador, el factor de impacto de la revista donde un artículo se ha publicado se ha convertido en el patrón oro de la evaluación científica.

Conocida esta deriva que han tenido los sistemas de acreditación, así como los de selección del profesorado en España es difícil dar marcha atrás, e introducir una visión más selectiva y cualitativa. Bien es verdad que en la penúltima reforma de la acreditación en *Aneca* se introdujo dentro de los méritos obligatorios como requisito la selección de las cuatro aportaciones más relevantes del CV del investigador. El paso iba en la buena dirección, pero no se suministraron instrucciones detalladas con ejemplificaciones adecuadas que permitieran a los investigadores una guía para implementarlas adecuadamente.

La filosofía del “todo vale” y “todo cuenta”, introduce estímulos erróneos cuando no fatales pues lleva a los profesores o aspirantes a serlo a intentar “rascar en todo”, lo que los distrae de lo que son las actividades centrales del profesor. Los candidatos se aprestan a amontonar y coleccionar méritos que encajen en todos los casilleros que habilitan las agencias. Estos modelos evaluativos conducen indefectiblemente a la burocratización y falta de flexibilidad provocando que la vida del investigador se convierta en una caza del mérito.

Frente a esta filosofía siempre es mejor solo contar aquello que vale, esto es, menos que más, pues menos es más. Y esto solo se puede hacer enjuiciando (valoración cualitativa) con datos, cuantificando y mostrando las evidencias empíricas.

En el nuevo modelo, lo que debiera importar es fijar cuáles son los logros especialmente significativos que determinarán la acreditación. Por ejemplo, a nivel de catedrático es lógico que el liderazgo sea una cualidad que deben exhibir los candidatos. Dicho liderazgo se manifiesta de muchas maneras (dirección de grupos de investigación, IP de proyectos, dirección de tesis doctorales, presidencia de sociedades científicas de amplio reconocimiento entre la comunidad, etc...). Por tanto, fijada la competencia a exhibir, los candidatos deben justificar con evidencias empíricas que se detenta.

Sin embargo, entiendo que se pueden establecer diferencias según el nivel o categoría profesional al que se opte. Tal vez en los niveles inferiores de la carrera se puede ser más detallado, mientras que en los superiores hay que ser mucho más selectivo. En los comienzos de la carrera el número de méritos es mucho menor y menos variado. Ahí se puede ser algo más exhaustivo, aunque dando preferencia a los indicadores significativos (no se trata de acumular méritos sino de alcanzar los relevantes).

En cualquier caso, a fin de eliminar la carga burocrática que implican los actuales sistemas de acreditación (horas empleadas por los investigadores, tiempos y recursos de los evaluadores y gestores) podría pensarse en establecer unos requisitos mínimos en docencia, investigación y transferencia para poder presentarse. Para niveles funcionariales es fácil indicar el número de sexenios de investigación, quinquenios de docencia y de transferencia requeridos. O bien, seleccionar esos méritos realmente significativos sin los que no se puede siquiera optar. Evidentemente, estos umbrales deben ser modulados según área de conocimiento.

Debe invitarse a que se declare cuáles son los logros significativos del CV, narrando y argumentando su relevancia, pero también justificando con evidencias empíricas y no con meras declaraciones formales cuales son los hitos del CV. Seleccionar un número reducido es imprescindible para enviar el mensaje de que lo que importa es la calidad de lo logrado y no la cantidad. En el ámbito de investigación, lo que importa son las ideas y/o descubrimientos realmente aportados, pero acompañados de evidencias empíricas que los demuestren.

“La filosofía del “todo vale” y “todo cuenta”, introduce estímulos erróneos cuando no fatales pues lleva a los profesores o aspirantes a serlo a intentar “rascar en todo”, lo que los distrae de lo que son las actividades centrales del profesor. Frente a esta filosofía siempre es mejor sólo contar aquello que vale”

Ir a una evaluación fundamentada significa de parte de los evaluadores emitir juicios de valor fundados en evidencias empíricas. De parte de los evaluados se trata de “contar”, esto es, narrar la trayectoria científica destacando cuáles son los hitos de la misma. Demostrar cuáles son las contribuciones del evaluado destacando la coherencia en la investigación y apoyando con evidencias empíricas (indicadores bibliométricos de impacto científico o de repercusión económica, social, según el caso). En definitiva, no se trata de hablar por hablar sino de hablar con datos. En la evaluación fundamentada, los indicadores están al servicio de los méritos, no los méritos al servicio de los indicadores. Los indicadores deben utilizarse para respaldar los logros alcanzados: primero se debe declarar qué se ha logrado y después exponer datos cuantitativos (indicadores bibliométricos entre otros) que los respalden de alguna manera.

Esta óptica cualitativa, la lógica en cualquier proceso de toma de decisiones (juicios, pero fundamentados), es un noble y racional anhelo que tropieza con la tradición española que he intentado esbozar en este ensayo. Implantar la evaluación cualitativa, sin más, eliminando cualquier signo cuantitativo nos puede devolver a épocas pretéritas: a resucitar la arbitrariedad y el nepotismo con su reguero de desafueros e inequidades. A la vista de la todavía escasa cultura de rendición de cuentas y de auditoría pública en España me temo lo peor. Mientras no seamos capaces de convertir estos mecanismos en prácticas habituales de nuestro proceder cotidiano habrá que seguir recurriendo a números que pongan freno a la dedocracia.

Además, el problema, en un país como España, es que dada la todavía escasa masa crítica de académicos en muchas disciplinas se torna imposible evitar los conflictos de interés (filias y fobias, conflictos de interés, escuelas, bandos) que comprometerían una evaluación neutral. En este sentido, los indicadores bibliométricos no deben ser eliminados sino cambiados, usados en su justa medida, en lo que convenga, de acuerdo con el contexto evaluado y en los momentos y circunstancias adecuadas. No deben restringirse a un único indicador sino a varios y variados (multifuentes, multifacetedos y multidimensionales) con capacidad de medir distintos niveles de repercusión (científica, social, económica, política, educativa)

En definitiva, los indicadores bibliométricos deben ser cuidadosamente usados e interpretados y no deben reemplazar el juicio humano (expertos). Deben ayudar en la toma de decisiones y no sustituirla o ser aplicados mecánicamente.

Y no olvidar que los números miden letras, que los números deben estar al servicio de las letras, porque cuando se antepone la medida a lo medido se termina midiendo por medir. Esto es muy peligroso, pues, aún bien empleados los indicadores bibliométricos desatan las más bajas pasiones, no las más altas. Se corre el riesgo de gamificar la actividad científica (Hammarfelt; De-Rijcke; Rushforth, 2016; Orduña-Malea; Martín-Martín; Delgado-López-Cózar, 2016), convertir las mediciones en objetos totémicos que pueden subvertir el orden moral científico lleva a que el medio (la publicación) se convierta en un fin en sí mismo. España es un país paradigmático por el prolongado uso de este tipo de métricas, pero especialmente por el abuso y mal uso que hemos hecho de ellas. Esto ha tenido efectos perversos, comportamientos deshonestos como plagio, publicación duplicada, manipulación de citas, publicar por publicar, cambios en las prácticas de publicación de los científicos, alteraciones en las agendas de investigación... (Feenstra; Delgado-López-Cózar, 2021).

“Implantar la evaluación cualitativa, sin más, eliminando cualquier signo cuantitativo nos puede devolver a épocas pretéritas: a resucitar la arbitrariedad y el nepotismo con su reguero de desafueros e inequidades. A la vista de la todavía escasa cultura de rendición de cuentas y de auditoría pública en España me temo lo peor. Mientras no seamos capaces de convertir estos mecanismos en prácticas habituales de nuestro proceder cotidiano habrá que seguir recurriendo a números que pongan freno a la dedocracia”

“Los indicadores bibliométricos no deben ser eliminados sino cambiados, usados en su justa medida de acuerdo con el contexto evaluado y en los momentos y circunstancias adecuadas. No deben restringirse a un único indicador sino a varios y variados (multifuentes, multifacetedos y multidimensionales) con capacidad de medir distintos niveles de repercusión (científica, social, económica, política, educativa)”

6. Nota

1. Datos y cifras del Sistema Universitario Español. Publicación 2020-2021.
https://www.universidades.gob.es/wp-content/uploads/2022/11/Datos_y_Cifras_2020_21.pdf

7. Referencias

- Ahedo-Gurrutxaga, Igor; Martínez-Palacios, Jone; Ormazabal-Gaston, Andere** (2022). "Letter: Central nodes in the reform of the Spanish Universities Law". *Profesional de la información*, v. 31, n. 3, e310309. <https://doi.org/10.3145/epi.2022.may.09>
- Aguillo, Isidro F.** (2021). "Letter: Respuesta a Delgado-López-Cózar et al.: Factor de impacto: muchas declaraciones y pocos resultados". *Profesional de la información*, v. 30, n. 3. <https://doi.org/10.3145/epi.2021.may.15>
- Arenilla-Sáez, Manuel** (2021). "La necesaria reforma de la Universidad española". *Revista de derecho político*, n. 110, p. 13-46. <https://doi.org/10.5944/rdp.110.2021.30314>
- Codina, Lluís** (2021). "Carta. Por un cambio en la evaluación de la investigación científica: 'no tirar al niño junto con el agua sucia'". *Profesional de la información*, v. 30, n. 4, e300418. <https://doi.org/10.3145/epi.2021.jul.18>
- Delgado-López-Cózar, Emilio** (2016). "La universidad española en el diván". En: Gutiérrez-Fuentes, José-Antonio; Sacristán-del-Castillo, José-Antonio (eds.). *Reflexiones sobre la ciencia en España. Cómo salir del atolladero*. Madrid: Unión Editorial, Fundación Lilly, pp. 99-166. ISBN: 9788472096806 <http://hdl.handle.net/10481/141048>
- Delgado-López-Cózar, Emilio** (2019). *Sexenios 2019: la transferencia se independiza, los rankings de revistas arriban al Derecho y siguen penetrando en las Humanidades y Ciencias Sociales*. Granada, 26 diciembre. <http://hdl.handle.net/10481/58574>
- Delgado-López-Cózar, Emilio; Martín-Martín, Alberto** (2019). "El Factor de Impacto de las revistas científicas sigue siendo ese número que devora la ciencia española: ¿hasta cuándo?". *Anuario ThinkEPI*, v. 13. <https://doi.org/10.3145/thinkepi.2019.e13e09>
- Delgado-López-Cózar, Emilio; Ràfols, Ismael; Abadal, Ernest** (2021). "Letter: A call for a radical change in research evaluation in Spain". *Profesional de la información*, v. 30, n. 3. <https://doi.org/10.3145/epi.2021.may.09>
- Delgado-López-Cózar, Emilio; Ràfols, Ismael; Abadal, Ernest** (2022). "Medidas insuficientes para un cambio en la evaluación de la investigación en España: glosando las nuevas directrices de la Aneca". *Recerca. Revista de pensament i anàlisi*. v. 27, n. 2. <https://doi.org/10.6035/recerca.6308>
- España* (2001). "Ley orgánica 6/2001, de 21 de diciembre, de Universidades". *BOE*, n. 307, 24 diciembre. <https://www.boe.es/boe/dias/2001/12/24/pdfs/A49400-49425.pdf>
- European Commission* (2021). *Towards a reform of the research assessment system: scoping report*. European Commission, Directorate-General for Research and Innovation. <https://data.europa.eu/doi/10.2777/707440>
- European Commission* (2022). *The agreement on reforming research assessment*. European Commission, Directorate-General for Research and Innovation. https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf
- Feenstra, Ramón A.; Delgado-López-Cózar, Emilio** (2021). "The footprint of a metrics-based research evaluation system on Spanish philosophical scholarship: an analysis of researchers perceptions". *Arxiv*, March 22. <https://doi.org/10.48550/arXiv.2103.11987>
- Hammarfelt, Björn; De-Rijcke, Sarah; Rushforth, Alexander D.** (2016). "Quantified academic selves: the gamification of research through social networking services". *Information research*. v. 21, n. 2. <http://InformationR.net/ir/21-2/ISM1.html>
- Orduña-Malea, Enrique; Martín-Martín, Alberto; Delgado-López-Cózar, Emilio** (2016). "Métricas en perfiles académicos: ¿un nuevo juego adictivo para los investigadores?". *Revista española de salud pública*, v. 90, e1-e5. https://www.sanidad.gob.es/biblioPublic/publicaciones/recursos_propios/resp/revista_cdrom/VOL90/C_ESPECIALES/RS90C_EDLC.pdf
- Ràfols, Ismael; Molas-Gallart, Jordi** (2022). "How to reform research evaluation in Spain. Institutional accreditation as a response to the European Agreement on research assessment. Letter". *Profesional de la información*, v. 31, n. 6. <https://doi.org/10.3145/epi.2022.nov.01>
- Rodríguez-Navarro, Alonso** (2022). *Cómo medir el éxito científico. Los errores de España*. Madrid: Aula Magna Proyecto clave McGraw Hill.
- Sosa-Wagner, Francisco** (2016). "La universidad española: entre Europa y una autonomía tramposa". *Revista de administración pública*, n. 200, p. 121-139. <https://doi.org/10.18042/cepc/rap.200.07>