

Carta: La insoportable levedad de la evaluación de proyectos de investigación: ¿necesitamos regularizar la actividad profesional de la evaluación científica?

Letter: The unbearable lightness of the evaluation of research projects: do we need to regularize the professional activity of research evaluation?

Enrique Orduña-Malea

Note: This article can be read in English on:

<http://www.profesionaldelainformacion.com/contenidos/2021/sep/orduna.pdf>

How to cite this article:

Orduña-Malea, Enrique (2021). "The unbearable lightness of the evaluation of research projects: do we need to regularize the professional activity of research evaluation?". *Profesional de la información*, v. 30, n. 5, e300413.

<https://doi.org/10.3145/epi.2021.sep.13>

Carta recibida el 23 de agosto de 2021



Enrique Orduña-Malea

<https://orcid.org/0000-0002-1989-8477>

Universitat Politècnica de València
Departament de Comunicació Audiovisual,
Documentació i Història de l'Art
Camí de Vera, s/n. Edifici 1H
46022 Valencia, España
enorma@upv.es

Resumen

Esta carta tiene como objetivo resaltar graves inconsistencias en la evaluación científica de proyectos de investigación financiados con fondos públicos regionales, destacando fallos estructurales relacionados con la gestión pública de la ciencia que van más allá de los aspectos técnicos bibliométricos. Este problema se ejemplifica a través de un estudio de caso ilustrativo (el programa de ayudas públicas a grupos de investigación emergentes financiado por el gobierno regional de la Comunidad Valenciana, España, NUTS-2), describiendo tres aspectos complementarios de esta convocatoria (el sistema de evaluación, el proceso de evaluación, y la comunicación de la evaluación). Los problemas ilustrados por este estudio de caso conducen a un problema general relacionado con el uso no estratégico de la ciencia. Se propone finalmente la regularización nacional e internacional de la actividad profesional de la evaluación de la investigación en general, y de la bibliometría en particular, como vía para evitar la mala gestión de la ciencia (y de la evaluación científica).

Palabras clave

Evaluación de la investigación; Bibliometría; Proyectos de investigación; Financiación pública de la ciencia; Ciencia de la ciencia; Meta-investigación; Convocatorias de investigación regionales; Gestión de la ciencia.

Abstract

This letter aims to highlight severe inconsistencies in the scientific evaluation of regionally publicly funded research projects, uncovering structural failures related to the public management of science that go beyond technical aspects of bibliometrics. This problem is exemplified through an illustrative case study (the public grants program to emerging research groups funded by the regional government of the Valencian region, Spain, NUTS-2), describing three complementary aspects of this call (the evaluation system, the evaluation process, and the communication of the evaluation). The problems illustrated by this case study lead to a general problem related to the nonstrategic use of science. Finally, the national and international regularization of the professional activity of research evaluation in general, and bibliometrics in particular, is proposed as a route to avoid mismanagement of science (and scientific evaluation).

Keywords

Research evaluation; Bibliometrics; Research projects; Public funding of science; Science of science; Metaresearch; Regional research calls; Science management.

Estimado editor:

En una carta reciente firmada por **Delgado-López-Cózar, Ràfols y Abadal** (2021), y posteriormente respondida por **Agui-Ilo** (2021) y **Codina** (2021), se plantean distintas preocupaciones relacionadas con la evaluación de la investigación en España.

Las preocupaciones sistematizadas y discutidas en esas cartas se centran principalmente, aunque no de forma exclusiva, en la evaluación de individuos (personas investigadoras) a través del trabajo que éstos realizan, así como del impacto (el eco o repercusión) de sus resultados y hallazgos. Dichas evaluaciones son generalmente necesarias para lograr la acreditación (que permite al profesorado universitario acceder a diferentes escalas; es decir, ser promocionado) o un 'tramo' de investigación positivo (lo que conlleva una recompensa económica, así como ciertos privilegios dentro de la universidad; es decir, obtener prestigio).

Si bien la evaluación de las personas es de suma importancia (afecta no sólo a las carreras investigadoras de esos individuos sino también a sus vidas personales), la evaluación de la ciencia es compleja e involucra a un amplio abanico de agentes y aspectos, entre los que destaca la evaluación de proyectos de investigación.

1. El preámbulo: proyectos de investigación

La importancia de la financiación pública de proyectos de investigación está fuera de toda duda, ya que el personal de investigación no puede realizar sus tareas únicamente con los recursos humanos y económicos proporcionados por sus instituciones (especialmente en las disciplinas empíricas).

Los proyectos de investigación pueden surgir tanto desde complejas convocatorias internacionales con grandes presupuestos y en los que participan grupos de varios países y campos de conocimiento, como hasta de simples convocatorias regionales o incluso institucionales en las que participan unos pocos investigadores que gestionan unos fondos limitados.

“ El personal de investigación no puede realizar sus tareas únicamente con los recursos humanos y económicos proporcionados por sus instituciones ”

Las convocatorias públicas regionales son particularmente importantes para los grupos emergentes con currículums incipientes, así como para los grupos consolidados con líneas de investigación enfocadas a nivel local y, por lo tanto, con menos oportunidades de participar con éxito en convocatorias más competitivas.

Como la financiación para la ciencia es limitada, las convocatorias competitivas suelen incluir un proceso de evaluación de la investigación, generalmente centrado en la propuesta específica (es decir, medir y valorar el interés y la necesidad del tema, así como la calidad del diseño del proyecto) o en los currículums del equipo (es decir, medir y valorar la trayectoria del personal de investigación hasta la fecha). Generalmente las evaluaciones incluyen una combinación de ambos aspectos.

La evaluación de proyectos de investigación, especialmente la aplicada a los componentes de los grupos de investigación, hereda la mayoría de los problemas y las deficiencias mencionadas en las cartas citadas, añadiendo además otros problemas de carácter más estructural y sistémico, que constituyen el asunto principal de la presente carta.

Mientras los medios generales y especializados se hacen eco generalmente de las convocatorias más elitistas y con los mayores presupuestos, en los que los aspectos evaluativos se controlan usualmente con mayor cuidado, las irregularidades encontradas en las convocatorias de financiación pública regional pasan desapercibidas, desafortunadamente incluso para el corpus de personas expertas internacionales en Ciencia de la ciencia y evaluación científica.

Las irregularidades en la evaluación de proyectos de investigación en convocatorias autonómicas deberían resaltarse (identificándolas y describiéndolas en detalle) para dar a conocer a la Sociedad la magnitud de las deficiencias que se producen; deficiencias que, por otra parte, quedan ocultas en aquellas cartas con reclamaciones más amplias, cuyo alcance y estructura no permiten el tratamiento de estudios de caso particulares. El objetivo específico de la presente carta es precisamente visibilizar estos estudios de caso regionales.

2. Caso de estudio: grupos emergentes

El Programa de ayudas públicas 2021 a grupos de investigación emergentes financiado por el gobierno regional de la *Comunitat Valenciana* (España, NUTS-2) se utiliza en esta carta como caso de estudio ilustrativo. Este programa se enmarca en las *Subvenciones del Programa de Promoción de la Investigación Científica, el Desarrollo Tecnológico y la Innovación en la Comunitat Valenciana* (Programa I+D+i).

https://innova.gva.es/es/web/ciencia/a-programa-i-d-i1/-/asset_publisher/jMe1UDRYZMHO/content/ix-subvenciones-a-grupos-de-investigacion-emergentes

A continuación se describen tres aspectos complementarios de esta convocatoria: el sistema de evaluación, el proceso de evaluación y la comunicación de la evaluación.

a) El sistema de evaluación

La puntuación final de cada solicitud se basa en la coherencia de la composición del grupo (20%) y en los currículums de los componentes del equipo (80%). El currículum de la persona que participa como investigadora principal representa el 50% de la puntuación final general, mientras que los currículums del resto del equipo representan el 30%. La evaluación de los CVs se basa en la producción científica, la participación en grupos de investigación internacionales, las presentaciones en congresos internacionales, supervisiones de tesis de máster y tesis doctorales, contratos de investigación o proyectos de especial relevancia, y patentes.

Se solicita una propuesta o memoria de investigación, pero ésta no se evalúa ni se considera en la puntuación final. Quizá por ese motivo no existen pautas específicas para cumplimentar adecuadamente estas propuestas. Ni siquiera se indican los apartados que la memoria debería incluir. La plantilla de la memoria utiliza el lenguaje inclusivo de forma incorrecta e incluye varios errores tipográficos.

Es descorazonador ver cómo los gobiernos públicos autonómicos gestionan la ciencia con tanto desinterés, ineficacia y falta de profesionalidad, mientras exigen el máximo al personal de investigación

Esta convocatoria emplea un sistema de evaluación no transparente. No existe información pública sobre cómo se evaluarán los méritos; los criterios de evaluación no se ponen a disposición de las personas solicitantes. La documentación oficial sólo menciona “producción científica”. Sin embargo, se desconoce si el criterio a aplicar será cuantitativo (número de publicaciones), selectivo (número de publicaciones en determinadas revistas o fuentes) o basado en impacto (por ejemplo, número de citas recibidas). No es posible comprobar si se aplican prácticas apropiadas de evaluación de la investigación.

Del mismo modo se desconoce la adecuación de la evaluación con respecto a las diferencias existentes entre disciplinas científicas. La convocatoria admite propuestas de hasta 18 campos¹, los cuales presentan diferencias significativas en la producción científica (e impacto) que deben ser consideradas. Una vez más, no es posible confirmar si se aplican prácticas adecuadas de evaluación de la investigación.

La ausencia de rúbricas de evaluación *ex ante* (antes de la evaluación) no es una práctica recomendada, ya que las rúbricas *post facto* (si las hubiera) no pueden garantizar la transparencia del proceso, dado que podrían haberse adaptado para ajustarse a las decisiones tomadas, limitando la posibilidad de apelación; es decir, las reglas del juego no deben comunicarse a las personas solicitantes cuando ya se han anunciado los resultados. En este caso en particular, las reglas ni siquiera se han comunicado aún.

Finalmente, los solicitantes pueden obtener una puntuación adicional (5% de la puntuación total obtenida) si la persona que ejerce de investigadora principal es mujer. Para evidenciar la necesidad de esta política de género, la convocatoria incluye un cuadro ilustrativo (en su Anexo XV) que presenta el desequilibrio de género por área temática (el porcentaje de hombres y mujeres para cada área y el sesgo de género). Sin embargo, el sesgo de género presenta un amplio rango estadístico (desde 1,50 en biomedicina hasta 36,20 en energía y transporte). Por lo tanto, la puntuación extra fija del 5% para todas las solicitudes, con independencia de su campo de investigación, parece bastante imprecisa, perjudicando a las mujeres en áreas con un sesgo de género más pronunciado, por ejemplo.

b) El proceso de evaluación

El proceso de evaluación es injustificablemente lento. Transcurrieron 125 días desde la fecha límite de la convocatoria de 2020 hasta el inicio del proceso de revisión externa, y 220 días más hasta la publicación de la resolución provisional, que incluía 187 solicitudes.

La lentitud del proceso de evaluación hace que la gestión de las ayudas recibidas sea económicamente ineficiente. La convocatoria de 2020 solicitaba que todos los proyectos se llevaran a cabo a partir del 1 de enero de 2021, admitiendo proyectos con una duración de 12 o 24 meses. Dado que se espera el anuncio de la resolución final para septiembre de 2021, los proyectos comenzarán con un retraso de al menos 8 meses. Aquellos grupos que pretendan terminar en diciembre de 2021 dispondrán de menos de 4 meses para completar el proyecto y justificar los gastos. Esto supondrá un cambio en la ejecución del proyecto en el mejor de los casos, y una incapacidad para llevarlos a cabo en el peor. La pandemia no representa una justificación, ya que esta circunstancia se ha dado en convocatorias anteriores, siendo un problema sistemático y endémico.

Desafortunadamente, estas irregularidades no sólo ocurren en convocatorias de grupos emergentes, como se ilustra en esta carta; también se dan en las convocatorias de grupos consolidados o para la organización y publicidad de congresos internacionales, cuyos solicitantes, a fecha de agosto de 2021, aún desconocen si serán beneficiarios de subvenciones para unos eventos que se deben realizar antes de fin de año.

La falta de regularización de la actividad profesional de evaluación de la investigación refuerza la mala gestión de la ciencia

La falta de transparencia en el proceso de evaluación también debe reseñarse. Por ejemplo, los datos sobre la composición del equipo de revisores y revisoras externas (por ejemplo, el número de personas revisoras por disciplina y género, o sus grados académicos o afiliaciones) no son públicos. De hecho, dado que la evaluación de las

propuestas se basa principalmente en la carrera del personal investigador (considerando el mérito de sus currículums), se debería discutir sobre la verdadera necesidad de disponer de evaluadores externos por campo, si lo único que van a realizar es contar un número de items en cada currículum. En el caso concreto de la evaluación de la producción científica de los solicitantes, sería conveniente saber si el conjunto de evaluadores externos tiene formación en bibliometría, especialmente si se van a utilizar indicadores bibliométricos en la evaluación (otro aspecto que permanece desconocido).

Algunas asociaciones deberían transformarse para llevar a cabo funciones profesionales reguladoras de la evaluación de la investigación

c) La comunicación de la evaluación

Esta convocatoria exhibe adicionalmente malas prácticas profesionales en la comunicación oficial de los resultados. La convocatoria 2020 se resolvió provisionalmente (aunque todavía sin validez legal) el viernes 30 de julio en torno a las 15:00 horas aproximadamente, precisamente en el momento en que los trabajadores públicos y solicitantes comenzaban oficialmente sus vacaciones de verano en España (el mes de agosto es inhábil para los trámites administrativos). Por lo tanto, las reclamaciones no están permitidas hasta septiembre de 2021. Además, no se envió ninguna notificación oficial a los solicitantes para informar de esta resolución:

<https://innova.gva.es/documents/161863198/172397122/RESOLUCI%C3%93%20CONVOCATORIA+ajudes+per+a+2021/1a493e0b-1b72-43cb-8e60-e4b97d754af4>

Por último, pero no por ello menos importante, no se entregó ningún informe de revisión a los solicitantes para justificar la puntuación obtenida (práctica, por cierto, empleada en la evaluación de otros proyectos de investigación). Ni siquiera se ha ofrecido la posibilidad de obtenerlo bajo demanda aún.

Con esta carta se pretende resaltar el desamparo que siente el personal de investigación ante estos procesos de evaluación “de caja negra” llevados a cabo en entornos regionales fuera del foco mediático, donde se utilizan criterios no transparentes y se aplican procesos administrativos ineficientes (incluyendo prácticas dudosas que podrían estar orientadas a tratar de minimizar las reclamaciones de los solicitantes), no cumpliendo con las buenas prácticas internacionales en evaluación de la investigación. Estos hechos se producen año tras año, independientemente del partido político que esté a cargo del gobierno.

Lamentablemente, el estudio de caso particular ilustrado en esta carta seguramente no constituye un ejemplo aislado. Estas prácticas pueden darse en otras regiones tanto de España como de otros países, con mayor o menor intensidad, perjudicando así a los elementos más vulnerables del ecosistema de investigación, como son los jóvenes investigadores y los grupos consolidados que trabajan en líneas de investigación de interés para la industria local.

Resulta descorazonador ver cómo los gobiernos públicos regionales manejan la ciencia con tanto desinterés, ineficiencia y falta de profesionalidad, mientras exigen el máximo al personal investigador.

3. Coda: hacia la profesionalización de la evaluación de la investigación

A través de este caso particular (una convocatoria regional para grupos emergentes) han surgido paradójicamente problemas generales en la evaluación de la investigación (por ejemplo, falta de transparencia y retrasos en la evaluación) que van más allá de la “buena o mala bibliometría” (uso del índice-h o del Factor de Impacto para las evaluaciones a nivel individual del personal de investigación, uso de *Google Scholar* para ciencias sociales y humanidades, o el uso de *altmetrics* para obtener evidencias de impacto alternativas, por nombrar sólo algunos polos de discusión actuales), conduciendo el problema desde la evaluación de la investigación hacia la gestión de la investigación.

El origen del problema proviene precisamente de no situar la investigación en el centro de la maquinaria social y económica (regional) del país, sino como un apéndice alejado del centro de gravedad político, lo que resulta en una carencia de interés estratégico a la hora de gestionar la ciencia de forma eficiente. La falta de regularización de la actividad profesional de la bibliometría en particular, y de la evaluación de la investigación en general, sólo refuerza la mala gestión de la ciencia.

A pesar de la existencia de declaraciones (*DORA*, 2012) y manifiestos (*Hicks et al.*, 2015) con alcance internacional que tratan la evaluación de la investigación, la ausencia de organizaciones con capacidad para ejercer una verdadera presión y “jurisdicción profesional” (*Jappe; Heinze*, 2020) hace que estas buenas prácticas se diluyan en el diseño de políticas públicas de investigación. Además, el hecho de que estas buenas prácticas no sean ni siquiera unánimemente aceptadas por toda la comunidad bibliométrica agrava el problema de la falta de autoridad científica (*auctoritas* de bajo nivel). Este hecho indudablemente provoca la escasa e inconsistente adopción de la *Declaración DORA* y el *Manifiesto de Leiden* en los procedimientos de evaluación de la investigación realizados a nivel mundial por entidades gubernamentales, que en última instancia son los que financian, gestionan y evalúan la investigación (*potestas* de alto nivel).

Sería deseable que algunas agencias y asociaciones nacionales e internacionales se transformasen para llevar a cabo funciones reguladoras en cuanto a la actividad profesional de evaluación de la investigación, como una actividad cuya práctica se aplica de acuerdo a los conocimientos generados por la comunidad científica correspondiente, requiriendo

de una formación específica y especializada. La ausencia de cursos avanzados y posgrados oficiales en esta área podría ser una de las principales causas de la mala evaluación de la investigación, no sólo en España. Debería ser obligatorio que dicho personal especializado y acreditado se involucrase en el diseño de políticas científicas que incluyesen cualquier tipo de proceso de evaluación, junto con especialistas de cada área de conocimiento.

Los gobiernos deben concebir la investigación como el motor económico de su país y como un generador de bienestar sostenible para sus ciudadanos

De la misma forma que la ciudadanía no aceptaría ser vacunada por personas que no fueran profesionales de la enfermería, tampoco deberíamos aceptar que la evaluación de la actividad científica sea realizada por personas sin la formación adecuada, sobre todo si tenemos en cuenta las repercusiones de la actividad investigadora, no sólo para el personal de investigación sino para una sociedad que podría beneficiarse de los hallazgos científicos (la vacuna Covid-19 es solamente un ejemplo mediático reciente, entre muchos otros que permanecen desconocidos para el público en general).

En una época en la que la sostenibilidad y la transparencia están en el punto de mira de los medios, quizá sea el momento de poner en práctica estos valores, especialmente en el área de la gestión de la investigación científica. Para ello, los gobiernos deben comenzar a regularizar esta actividad profesional a partir de la existencia de entidades nacionales (encargadas de alinear las mejores prácticas de evaluación de la investigación con las condiciones sociales, culturales y económicas de cada región y asegurando procesos de evaluación justos), siendo coordinadas por una entidad no gubernamental internacional (encargada de canalizar los hallazgos de la comunidad científica y favorecer el consenso en la evaluación). Todas estas entidades deberían asegurar y proteger los derechos y obligaciones de la comunidad relacionada con el ecosistema de la ciencia (personal de investigación, profesionales y solicitantes de servicios científicos).

Cuando dichas organizaciones regulasen oficialmente la actividad profesional de evaluación de la investigación, se podrían entonces discutir e implementar las técnicas bibliométricas específicas que deben emplearse en el ejercicio oficial de la profesión a la hora de evaluar proyectos, acreditaciones, promociones, periodos de investigación, etc.

Sin embargo, y para que esto ocurra, los gobiernos deben concebir la investigación como el motor económico del país y como un generador sostenible de calidad de vida y bienestar para sus ciudadanos y ciudadanas. De lo contrario, la actividad científica (y su evaluación) será como un iceberg arrancado de los polos y a la deriva (como lo es actualmente) que terminará deritiéndose y golpeando la costa en forma de una ola gigante.

Atentamente,
Dr. Enrique Orduña-Malea

4. Notas

1. Los campos de investigación incluidos son los siguientes:

Biociencias y biotecnología; Biomedicina; Ciencias agrarias y agroalimentarias; Ciencias de la educación; Ciencias físicas; Ciencias matemáticas; Ciencias sociales; Ciencias y tecnologías de materiales; Ciencias y tecnologías medioambientales; Ciencias y tecnologías químicas; Cultura: Filología, Literatura y Arte; Derecho; Economía; Energía y Transporte; Estudios del pasado: Historia y Arqueología; Producción industrial, Ingeniería civil e Ingenierías para la Sociedad; Psicología; Tecnologías de la Información y de las Comunicaciones.

5. Referencias

Aguillo, Isidro F. (2021). "Respuesta a Delgado-López-Cózar *et al.* Factor de impacto: Muchas declaraciones y pocos resultados. Letter". *Profesional de la información*, v. 30, n. 3, e300315.

<https://doi.org/10.3145/epi.2021.may.15>

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; 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, e300309.

<https://doi.org/10.3145/epi.2021.may.09>

DORA. *San Francisco Declaration on Research Assessment* (2012).

<https://sfdora.org/read>

Hicks, Diana; Wouters, Paul; Waltman, Ludo; De-Rijcke, Sarah; Ràfols, Ismael (2015). "Bibliometrics: The Leiden Manifesto for research metrics". *Nature*, n. 520, pp. 429-431.

<https://doi.org/10.1038/520429a>

Jappe, Arlette; Heinze, Thomas (2020). "Jurisdiction of bibliometrics". In: Ball, Rafael (ed.). *Handbook bibliometrics*. Berlin: De Gruyter Saur, pp. 91-98. ISBN: 978 3 11 064227 8

<https://doi.org/10.1515/9783110646610-010>